HumanE AI Net:

The HumanE AI Network

Grant Agreement Number: 952026
Project Acronym: HumanE AI Net

Project Dates: 2020-09-01 to 2024-08-31
Project Duration: 48 months

D2.1: Micro-Project Results on Learning, Reasoning, Perception and Interaction for the 2nd Project Period

Author(s): James Crowley, John Shawe-Taylor, Antti Oulasvirta
Contributing partners: INRIA, Univ Grenoble Alps, UCL, Aalto Univ.
Date: May 31, 2023
Approved by: Paul Lukowicz
Type: Report (R)
Status: Draft
Contact: James Crowley

Dissemination Level		
PU	Public	x

DISCLAIMER

DOCUMENT INFO

Authors

Authors	Institution	e-mail
James Crowley	INRIA	James.Crowley@inria.fr
John Shawe-Taylor (JST)	UCL	j.shawe-taylor@ucl.ac.uk
Antti Oulasvirta	Aalto Univ	antti.oulasvirta@aalto.fi

Document History

Revision		
Date	**Lead Author(s)**	**Comments**
28/02/2023	JLC	Empty template
31/03/2023	JLC	First Draft based on john's contributions. Request for input from missing Micro-Project reports.
06/04/2023	JLC	Included contributions from Micro-project to WP2 as example for other work packages.
12/04/2023	JLC	Draft presented to WP leaders, Contributions requested for WP 1 and 3.
17/04/2023	JLC	Added information about MPs for each WP. Circulated to WP leaders for WP 1, 2 and 3 for descriptions of Contributions of each MP to WP.
20/04/2023	AO	Sent request for missing status reports sent to Micro-project leaders.
26/04/2023	AO	Update of WP3 related Micro-Projects. Relaunched Micro-project leaders for missing activity reports.
12 May 2023	AO	Updated reports on contributions from Micro-project to WP3
13 May 2023	JST	Updated reports from missing Micro-Projects related to WP1. Updated report on
14 May 2023	JLC	Updated report with missing Micro-Project results reports. Edited for consistent formatting.
14 May 2023	Virginia Dignum	Internal Review
15 May 2023	JLC	Final Version

Table of Contents

Introduction

This deliverable reports on the Micro-Projects that have responded to the research challenges raised in Humane AI Net work-packages 1, 2 and 3, during the second project period from 1 Oct 2021 through 28 Feb 2023. It includes descriptions of Micro-Projects that were underway at the beginning of the second project period (started before Oct 2021), micro-projects that started and completed during the second period, as well as micro-projects that began during the second project period and have not yet competed.

This report describes how these micro-projects respond to research challenges raised in the original research agenda for work packages 1, 2 and 3. This is followed by a summary of the new micro-project procedures that have taken effect from January 2023 to respond directly to the revised Human AI Net research agenda released during 202. The report concludes with a formatted copy of the final reports from each micro-project, as reported by the project teams using an on-line web form.

In order to help organize a quick overview of micro projects and their contributions to work packages, we have adopted a notation for each micro-project based on its starting date and the workpackages to which it contributes. Each micro-project is identified with a label: yyyy-mm-WPnnn[-m] where yyyy refers to the year of the start date (2021, 2022, 2023 or 2024), mm refers to the month of the of the start date of the Micro-Project (01 to 12), and nnn refers to the work packages to which the Micro-Project contributed. In cases where more than one microproject share a start date and contribute to the same workpackages, and additional field, m is added. Thus for, example, a Micro-Project labeled 2022-01-WP123-1 can be seen as one of a set of micro-projects starting in January 2022 and contributing to research objectives found in work packages 1, 2, and 3.

Our assessment is that the micro-project approach has been very successful in initiating collaborative exploration of new research ideas at TRL1 and 2, often with outcomes that can be followed with larger research efforts. However, a few research areas remained to be addressed, and not all partners had been able to profit from the flexibility for rapid exploration of new research ideas made possible by the micro-project concept.

Work package leaders address these issues in regular meetings, by reviewing the progress of the overall research against the work programme and identify missing elements as well as particularly promising emerging directions. These were then used to incite calls for new MP proposals addressing these topics, but at the same time leaving open the option for proposals addressing different topics.

Following the development of a new strategic research agenda in 2022, the workpackage leaders have decided to adopt a new, more focused procedure for inciting and soliciting Micro-Projects. Starting in 2023, the consortium will issue regular calls for micro projects to address specific focused research challenges with new ideas and approaches (TRL 1, 2, or 3). New calls will be initiated by the workpackage leaders every 4 months. Any proposal initiated by a consortium partner, that meets minimum criteria for a micro project, and proposes to explore plausible new approach to the research challenge will be accepted. The acceptance procedure is very lightweight with acceptable proposals authorized to begin work in the month following submission.

A first call for such proposals was issued in January 2023. Results will be reported in deliverable D3.1 at the end of the third period.

The following summarizes the contributions to the research agendas for workpackages 1, 2 and 3 provided by Micro-Project carried out during the second project period. A companion report (5.2) reports on results for work packages 4, 5, 6, and 7.

Work Package 1: Human-in-the-Loop Machine Learning, Reasoning and Planning

Coordinator: John Shawe-Taylor, UCL

Research Objectives for WP 1

The work package 'Learning, Reasoning and Planning with Human in the Loop' is concerned with the development of core technologies enabling more inclusive AI solutions that allow both understanding and interaction with humans. Such developments will underpin much of the work in other work packages that implement and demonstrate how humans can be empowered by AI in ways that enrich rather than sideline. The set of Micro-Projects (MPs) that have been proposed cover the wide spectrum of tasks envisaged in the work package. The first of these was the linking of symbolic and sub-symbolic learning evident both explicitly and implicitly in several MPs, hence providing evidence for the value of this integration at a variety of levels. The second task of learning with and about narratives is also an explicit theme in one MP, as well as appearing as a thread in other MPs. Again, there is evidence that the 'narratives' perspective can motivate the development of new representations as well as guide the design of systems that are potentially more understandable to humans. The third task was the development of continuous and incremental learning in joint human/AI systems that have also been prominent in a number of Micro-Projects. The final two technical tasks, Auto ML and 'quantifying model uncertainty' have received less attention in this period. The final task is the consolidation and coordination of the research agenda which has been addressed in the ongoing work on a coherent program of research for human-centric AI, particularly in the form of the report of the Dagstuhl workshop.

Contributions from Micro-Projects:

2021-01-WP13: Coping with the variability of human feedback during interactive learning through ensemble reinforcement learning

Proposed by: Mehdi Khamassi, (Sorbonne University)
Other Partners: Petros Maragos (ATHINA Research Center)

Start Date: 1 January 2021
Duration: 15 Months
Completion Date 4 April 2023

Contribution to Objectives Work Package 1:

This 6-month project entails robot online behavioral adaptation during interactive learning with humans. Specifically, the robot shall adapt to each human subject's specific way of giving feedback during the interaction. Feedback here includes reward, instruction and demonstration, and can be regrouped under the term "teaching signals". Following the HumanE AI vision, interactive learning puts the human in the loop, prompting human-aware robot behavioral adaptation.

The micro-project directly contributes to one of the objectives of WP1 (T1.3) to enable "continuous incremental learning in joint human/AI systems" by 'exploiting rich human feedback'.

Publications

- Rémi Dromnelle, Erwan Renaudo, Benoît Girard, Petros Maragos, Mohamed Chetouani, Raja Chatila, Mehdi Khamassi (2022). Reducing computational cost during robot navigation and human-robot interaction with a human-inspired reinforcement learning architecture. International Journal of Social Robotics, doi: 10.1007/s12369-022-00942-6, https://link.springer.com/article/10.1007/s12369-022-00942-6

Tangible Results

- Open source code: https://github.com/DromnHell/meta-control-decision-making-agent
- Publication: Preprint made open on HAL – https://hal.sorbonne-universite.fr/hal-03829879

2021-03-WP123: Memory Aware Conversational AI to aid virtual Team-Meetings

Proposed by: Catholijn Jonker, TUDelft

Other Partners:
1. Catharine Oertel (Technical University Delft),
2. Andras Lorincz (Eotvos Lorand University)

Start Date: 2 March 2021
Duration: 24 Months
Completion Date 2 March 2023

Contribution to Objectives Work Package 1:

This micro-project principally addresses WP3 but the corpus generated may be of interest in developing new approaches in WP1.

Publications

- Tsfasman, M., Fenech, K., Tarvirdians, M., Lorincz, A., Jonker, C., & Oertel, C. (2022). Towards creating a conversational memory for long-term meeting support: predicting memorable moments in multi-party conversations through eye-gaze. In ICMI 2022 - Proceedings of the 2022 International Conference on Multimodal Interaction (pp. 94-104). (ACM International Conference Proceeding Series). Association for Computing Machinery (ACM). https://doi.org/10.1145/3536221.3556613

Tangible Results

- The corpus is already available by request - contact Maria Tsfasman (m.tsfasman@tudelft.nl) and Catharine Oertel (c.r.m.m.oertel@tudelft.nl). It is planned that the corpus will be released by the end of 2024.
- For any other questions and inquiries feel free to contact Maria Tsfasman (m.tsfasman@tudelft.nl) and/or Catharine Oertel (c.r.m.m.oertel@tudelft.nl).
- https://www.youtube.com/watch?v=kTS8XwUVNnI

2021-04-WP1: Discovering Temporal Logic patterns as binary supervised learning,

Proposed by: Chiara Ghidini, (FBK)

Other Partners:

Federico Chesani (University of Bologna)
Sergio Tessaris (Free University of Bozen-Bolzano - External Partner)

Start Date: 14 April 2021
Duration: 5 Months
Completion Date 14 September 2021

Contribution to Objectives Work Package 1:

The Micro-Project has principally addressed task T1.1 of WP1: "Linking symbolic and sub-symbolic learning", in creating interpretable representations for learning patterns that also relate to task T1.2 “learning with and about narratives”, in that the patterns have a temporal logic structure that accords with natural understanding of events.

Publications

- Chesani, F., Francescomarino, C. D., Ghidini, C., Grundler, G., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Shape your process: Discovering declarative business processes from positive and negative traces taking into account user preferences. In Almeida, J. P. A., Karastoyanova, D., Guizzardi, G., Montali, M., Maggi, F. M., and Fonseca, C. M., editors, Enterprise Design, Operations, and Computing - 26th International Conference, EDOC 2022, Bozen-Bolzano, Italy, October 3-7, 2022, Proceedings, volume 13585 of Lecture Notes in Computer Science, pages 217–234. Springer.
- Chesani, F., Francescomarino, C. D., Ghidini, C., Grundler, G., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Shape your process: Discovering declarative business processes from positive and negative traces taking into account user preferences. In Almeida, J. P. A., Karastoyanova, D., Guizzardi, G., Montali, M., Maggi, F. M., and Fonseca, C. M., editors, Enterprise Design, Operations, and Computing - 26th International Conference, EDOC 2022, Bozen-Bolzano, Italy, October 3-7, 2022, Proceedings, volume 13585 of Lecture Notes in Computer Science, pages 217–234. Springer.
- Chesani, F., Francescomarino, C. D., Ghidini, C., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Process discovery on deviant traces and other stranger things. IEEE Transactions on Knowledge and Data Engineering, pages 1–17. DOI: https://doi.org/10.1109/TKDE.2022.3232207

Tangible Results (available on the AI4Europe platform

- Loan Approval1: dataset. https://drive.google.com/drive/folders/15BwG4PJq8iIMh9Sr9dpMXAYBY-qp7QDE?usp=sharing
- Loan Approval2: dataset. https://drive.google.com/drive/folders/1fcJ8itzdMbNOjEAeV6nUEeI5B6__aB_c?usp=sharing
- Discovery Framework: program/code https://zenodo.org/record/5158528
- Experiments: https://github.com/stessaris/negdis-experiments/tree/v1.0

2021-4-WP123 Multimodal Perception and Interaction with Transformers

Proposed by: James Crowley, Inria

Other Partners:

1. Andras Lorincz (Eotvos Lorand University - ELTE),
2. Dominique Vaufreydaz, Fabien Ringeval (Univ Grenoble Alpes),

3. Camille Guinaudeau, Marc Evrard (Uni Paris Saclay,
4. Marko Grobelnik (Jozef Stefan Institut-JSI),
5. Pavel Pecina (Charles University)

Start Date: 1 April 2021
Duration: 6 Months
Completion Date 31 October 2021

Contribution to Objectives Work Package 1:

This micro-project has aided and encouraged the use of a transformers and self-attention for multimodal modal interaction by Humane AI Net researchers, by identifying relevant tools and benchmark data sets, by providing tutorials and training materials for education, and by identifying research challenges for multimodal perception and interaction with Transformers.

Tangible Results (available on the AI4Europe platform

- A survey of tools and datasets for a multimodal perception with transformers (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/TransformerTools.pdf)
- A tutorial on the use of transformers for multimodal perception. (http://crowley-coutaz.fr/jlc/Courses/ACAI2021/Multimodal-Transformer-Tutorial.html)
- Report on challenges for the use of transformers for multimodal perception and interaction. (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/ReseachChallengesDataSets.pdf)

Publications

1) Y. Wang, X. Shen, S. Hu, Y. Yuan, J. L. Crowley, D. Vaufreydaz, Self-Supervised Transformers for Unsupervised Object Discovery using Normalized Cut. IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2022, pp14543-14553, New Orleans, Jun 2022.
2) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin and J. L Crowley, "Emotion Recognition with Pre-Trained Transformers Using Multimodal Signals", 10th International Conference on Affective Computing and Intelligent Interaction (ACII), Oct 2022.
3) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin, J. L. Crowley. Transformer-Based Self-Supervised Learning for Emotion Recognition. 26th International Conference on Pattern Recognition (ICPR 2022), Aug 2022, Montreal, Canada.

2021-09-WP15: A Graph-Based Drift-Aware Data Cloning Process,

Proposed by: Joao Gama, INESC TEC

Other Partners: Giuseppe Manco (CNR), Holger Hoos (Leiden University)

Start Date: 1 Sept, 2021
Duration: 7 Months
Completion Date 31 March, 2022

Contribution to Objectives Work Package 1:

The goal is to devise a data generation methodology that, given a data sample, can approximate the stochastic process that generated it. The methodology can be useful in many contexts where we need to share data while preserving user privacy. This relates to task T1.1 linking symbolic and sub-symbolic reasoning as well as task T1.2 learning with and about narratives in that the generation process is interpretable and adheres to a human interpretable model.

Publications

- Luciano Caroprese, Francesco Sergio Pisani, Bruno Veloso, Matthias König, Giuseppe Manco, Holger H. Hoos, and João Gama. Modelling Concept Drift in Dynamic Data Streams for Recommender Systems (under evaluation)

2021-04-WP136: Combining symbolic and sub-symbolic approaches - Improving neural Question-Answering-Systems through Document Analysis for enhanced accuracy and efficiency in Human-AI interaction.

Proposed by: Haris Papageorgiou (Athena RC)

Other Partners: Georg Rehm (DFKI)

Start Date: 15 April 2021
Duration: 9.5 Months
Completion Date 31 January 2022

Contribution to Objectives Work Package 1:

This micro-project has addressed task T1.1 of WP1: "Linking symbolic and sub-symbolic learning" most directly specifically with the purpose of enabling richer interactions between humans and AI systems. The Micro-Project has therefore also addressed task T1.3, “development of continuous and incremental learning in joint human/AI systems”.

Publications

- Pappas Dimitris, Lyris Ioannis, Kountouris George and Papageorgiou Haris: "A Neurosymbolic Question Answering System Combining Structured and Unstructured Biomedical Knowledge", Proceedings of the 3rd Conference on AI for Humanity and Society (AI4HS), Stockholm, 2022

Tangible Results (available on the AI4Europe platform

- Video: "Combining symbolic and sub-symbolic approaches - Improving neural Question-Answering-Systems through Document Analysis for enhanced accuracy and efficiency in Human-AI interaction"
- Demonstrator: A neurosymbolic Question Answering System on Covid-19 and SARS-CoV-2

2021-05-WP13: Interactive Reinforcement Learning for Humorous Agents

Proposed by: Brian Ravenet (CNRS)

Other Partners: Rui Prada (INESC-ID)

Start Date: 1 May 2021
Duration: 4 Months
Completion Date 31 August 2021

Contribution to Objectives Work Package 1:

The main result of this project are the creation of an intelligent agent capable of playing a game – Cards Against Humanity- that involves matching sentences with humorous comebacks. This addressed task T1.3 of “development of continuous and incremental learning in joint human/AI systems”.

Tangible Results (available on the AI4Europe platform

- Not applicable.

2021-05-WP16: Feasibility analysis of hardware acceleration for AML,

Proposed by: Fernando Martin Maroto, (Algebraic AI)

Other Partners: Christian Weis (Technische Universität Kaiserslautern)

Start Date: 1 May 2021
Duration: 6 Months
Completion Date 31 October 2021

Contribution to Objectives Work Package 1:

This micro-project has investigated the use of specialist hardware for implementation of Algebraic machine learning. This analysis is required in order to make this approach applicable to 'Learning, Reasoning and Planning with Human in the Loop' in particular to task T1.3: the development of continuous and incremental learning in joint human/AI systems.

Publications

- Fernando Martin-Maroto and Gonzalo G. de Polavieja. Algebraic Machine Learning. arXiv:1803.05252, 2018.
- Fernando Martin-Maroto and Gonzalo G. de Polavieja. Finite Atomized Semilattices. arXiv:2102.08050, 2021.
- Fernando Martin-Maroto and Gonzalo G. de Polavieja. Semantic Embeddings in Semilattices. arXiv:2205.12618, 2022

Tangible Results

- The report and the details of the FPGA based prototype are available for review. Some results are under a patent process and are not yet available to the public

2022-06-WP1, AI-ASSISTANT TO MITIGATE CONFIRMATION BIAS IN COOPERATIVE BAYESIAN OPTIMIZATION

Proposed by: Samuel Kaski, Aalto

Other Partners: Frans Oliehoek (Delft University of Technology: TU Delft)

Start Date: 1 June 2022
Duration: 4 Months
Completion Date 30 September 2022

Contribution to Objectives Work Package 1:

This micro-project contributes to developing methodologies that allow humans to be interactively involved "in the loop".

Here, the loop is a cooperative Bayesian optimization game where the goal is to optimize a 2D black-box function. At each iteration, the AI chooses the first coordinate, and then the user observes and opts for the second. Finally, the function is queried, and the result is shown to both parties. The researcher can control agents' characteristics, making it suitable for studying confirmation bias and imperfect knowledge. The project investigates how a planning AI agent can alleviate BO regret due to the human agent's biases and imperfect information allocation. The aim is to build a planning AI agent to aid the user in the optimization task, where no single party has full decision-making power. This

contributes most specifically to task T1.3, the development of continuous and incremental learning in joint human/AI systems.

Publications

- Cooperative Bayesian Optimization for Imperfect Agents, Khoshvishkaie A., Mikkola P., ... and Kaski S., submitted to ECML-2023 A tutorial on the use of transformers for multimodal perception.

Tangible Results (available on the AI4Europe platform)

- The code will be publicly available upon acceptance to the conference mentioned above.

2022-09-WP12-2, Multi-Relational Contextual Reasoning for Complex Scene Generation for Autonomous Vehicle Data

Proposed by: Loris, Bozzato, FBK

Other Partners:
Thomas Eiter (Technical University of Vienna), Daria Stepanova (BOSCH)

Start Date: 1 September 2022
Duration: 7 Months
Completion Date 31 March 2023

Contribution to Objectives Work Package 1:

In this work, we provide a prototype specialized to generating similar and problematic scenes in the domain of Autonomous Driving.

This work fits well into Task 1.1 of WP1: "Linking symbolic and subsymbolic learning", since it uses a symbolic approach to enable the use of domain knowledge in order to advance the performance of a sub-symbolic model.

Furthermore, it also loosely fits into Task 1.5 of WP1: "Quantifying model uncertainty", since it can quantify how similar the generated new inputs are to the original ones..

Publications

- Loris Bozzato, Thomas Eiter, Rafael Kiesel and Daria Stepanova (2023). Contextual Reasoning for Scene Generation (Technical Report). https://arxiv.org/abs/2305.02255

Tangible Results (available on the AI4Europe platform)

- Prototype implementation: https://github.com/raki123/MR-CKR

(The prototype has been also submitted as an AI Asset to the AI4Europe)

2021-06-WP136: Adaption of ASR for Impaired Speech with minimum resources (AdAIS),

Proposed by: Mireia Diez Sanchez, (BUT, Brno University of Technology)

Other Partners: Tim Polzehl (TUB, TECHNISCHE UNIVERSITÄT BERLIN)

Start Date: 15 June 2021
Duration: 5 Months
Completion Date 15 November 2021

Contribution to Objectives Work Package 1:

In this microproject, we pursued enabling access to AI technology to those who might have special needs when interacting with AI: Automatic Speech Recognition made accessible for people with dysarthria. We showed that, using the adapter module, fMLLR and xvectors are complementary to each other, and proved the effectiveness of the approach outperforming existing SoTA on UASpeech dysartric speech ASR.

AdAIS addressed topics related to the tasks of WP1 Learning, Reasoning and Planning with Human in the Loop, in particular T1.1 Linking symbolic and sub-symbolic learning.

Publications

- Publication: M. K. Baskar, T. Herzig, D. Nguyen, M. Diez, T. Polzehl, L. Burget, J. Černocký, “Speaker adaptation for Wav2vec2 based dysarthric ASR”. Proc. Interspeech 2022, 3403-3407, doi: 10.21437/Interspeech.2022-10896

Tangible Results (available on the AI4Europe platform

- Link to publication: https://www.isca-speech.org/archive/pdfs/interspeech_2022/baskar22b_interspeech.pdf
- Open source tool for training ASR models for dysarthic speech: https://github.com/creatorscan/Dysarthric-ASR The repository contains: A baseline recipe to train a TDNN-CNN hybrid model based ASR system, this recipe is prepared to be trained on the TORGO dataset. And an end-to-end model using ESPnet framework prepared to be trained on UASpeech dataset.

2021-09-WP1: Neural-Symbolic Integration: explainability and reasoning in KENN,

Proposed by: Frank van Harmelen, VU University

Other Partners: Luciano Serafini (FBK)

Start Date: 1 September 2021
Duration: 7 Months
Completion Date 31 March 2022

Contribution to Objectives Work Package 1:

As a result of using background knowledge from a knowledge we can train the neural network with many fewer training examples. Since KENN is based on fuzzy logic, a major bottleneck was the choice of the appropriate configuration of the logic (choice of norms and co-norms), since earlier work from Amsterdam had showed that some of the classical fuzzy logic configurations would perform very poorly in a machine learning setting. As a result of the collaborations, we have developed so called Fuzzy Refinement Functions. Such "refinement functions" are functions that change the truth value computed by a fuzzy logic operator in order to improve the gradient behaviour, while still maintaining the desired logical combinatorics. We have implemented such refinement functions in an algorithm called Iterative Local Refinement (ILR). This contributes directly to task T1.1 linking symbolic and subsymbolic computation.

Publications

- Refining neural network predictions using background knowledge, Alessandro Daniele, Emile van Krieken, Luciano Serafini & Frank van Harmelen, Machine Learning (2023) https://link.springer.com/article/10.1007/s10994-023-06310-3

Tangible Results

- Publication at https://link.springer.com/article/10.1007/s10994-023-06310-3
- Code and data at https://github.com/DanieleAlessandro/IterativeLocalRefinement

2022-09-WP12-1: Gray-box approach to narrative analysis,

Proposed by: Szymon Talaga, University of Warsaw

Other Partners: James Crowley (University of Grenoble Alpes)

Start Date: 1 September 2022
Duration: 5 Months
Completion Date 31 January 2023

Contribution to Objectives Work Package 1:

The Micro-Project fits into a broader HumanE-AI objective of developing common ground concepts providing better representations shared by humans and machines alike. Narrative-oriented tools for automated text analyses can empower human analysts as, arguably, the narrative framework provides a more natural and meaningful context for people without formal training in linguistics and/or computer science for reasoning about textual data. This directly addresses task T1.2 “learning with and about narratives”.

Tangible Results

- https://github.com/sztal/segram

Work Package 2: Multimodal Perception and Modeling

Work Package Coordinator: James L. Crowley, INRIA/Grenoble INP and UGA

Research Objectives for WP 2:

The objective of WP2 is to provide a foundation for multimodal perception and modeling needed for natural interaction and collaboration between humans and intelligent systems.

Humans interact with the world using five major senses: sight, hearing, touch, smell, and taste. Almost all interaction with the environment is naturally multimodal, as audio, tactile or paralinguistic cues provide confirmation for physical actions and spoken language interaction. To interact and collaborate with people, intelligent systems must be able to perceive and model humans, human actions, and behaviors, human attention and awareness, human emotions, human language and human social interaction, as well as real-world human environments. Multimodal interaction seeks to fully exploit these parallel channels for perception and action to provide robust, natural interaction and collaboration with intelligent systems.

WP2 builds on recent advances in multimodal perception and modeling. Specifically, we define research challenges for natural multimodal interaction, document and augment the existing tools and data sets needed to meet such challenges and demonstrate the use of recent advances in machine learning for building systems.

Work on multimodal perception and modeling is organized around the following activities: Learning of multimodal models grounded in physical reality (T2.1), Multimodal perception and modeling of actions, activities and tasks (T2.2), Multimodal perception of awareness, emotions, and attitudes (T2.3), Perception of Social Signals and Social Dynamics (T2.4), Distributed Collaborative Perception and Modeling (T2.5), Dealing with lack of labeled training data (T2.6), and Assembling benchmark datasets (T2.7).

Contributions from Micro-Projects

2021-4-WP123 Multimodal Perception and Interaction with Transformers

Proposed by: James Crowley, Inria

Other Partners:

1. Andras Lorincz (Eotvos Lorand University - ELTE),
2. Dominique Vaufreydaz, Fabien Ringeval (Univ Grenoble Alpes),
3. Camille Guinaudeau, Marc Evrard (Uni Paris Saclay,
4. Marko Grobelnik (Jozef Stefan Institut-JSI),
5. Pavel Pecina (Charles University)

Start Date: 1 April 2021
Duration: 6 Months
Completion Date 31 October 2021

Contribution to Objectives Work Package 2:

This micro-project has aided and encouraged the use of a transformers and self-attention for multimodal modal interaction by Humane AI Net researchers, by identifying relevant tools and benchmark data sets, by providing tutorials and training materials for education,

and by identifying research challenges for multimodal perception and interaction with Transformers.

Tangible Results (available on the AI4Europe platform

- A survey of tools and datasets for a multimodal perception with transformers (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/TransformerTools.pdf)
- A tutorial on the use of transformers for multimodal perception. (http://crowley-coutaz.fr/jlc/Courses/ACAI2021/Multimodal-Transformer-Tutorial.html)
- Report on challenges for the use of transformers for multimodal perception and interaction. (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/ReseachChallengesDataSets.pdf)

Publications

1) Y. Wang, X. Shen, S. Hu, Y. Yuan, J. L. Crowley, D. Vaufreydaz, Self-Supervised Transformers for Unsupervised Object Discovery using Normalized Cut. IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2022, pp14543-14553, New Orleans, Jun 2022.
2) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin, J. L. Crowley. Transformer-Based Self-Supervised Learning for Emotion Recognition. 26th International Conference on Pattern Recognition (ICPR 2022), Aug 2022, Montreal, Canada.
3) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin and J. L Crowley, "Emotion Recognition with Pre-Trained Transformers Using Multimodal Signals", 10th International Conference on Affective Computing and Intelligent Interaction (ACII), Oct 2022.

2021-5-WP26: Neural Mechanism in Human Brain Activity During Weight Lifting
Start Date: 2021-05-01
Duration: 8 months
End Date: 2022-01-31

Submitted by: Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr
Partners:
1. TUBITAK BILGEM, Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr
2. DFKI Kaiserslautern, Hamraz Javaheri, Hamraz.Javaheri@dfki.de

In this project, it was investigated whether EEG (electroencephalography) signal can be used for detecting the motion as well as the variable weights a person is lifting. To do this, an experimental paradigm has been designed and EEG data have been acquired during performing biceps flexion-extension motions for different weight categories: lifting with no weight (empty), medium, and heavy lifting. The outcomes of the project can be applied in industrial exoskeleton applications as well as physical rehabilitation of stroke patients.Results contribute to T2.2 Multimodal perception and modeling of actions, activities and tasks.

2022-3-WP2 Natural Imitation of Dance Moves and Human Gestures with a Humanoid Robot

Submitted by: Shivesh Kumar, shivesh.kumar@dfki.de
Partners:
1. DFKI Kaiserslautern, Shivesh Kumar, shivesh.kumar@dfki.de
2. INRIA Paris, Justin Carpentier, justin.carpentier@inria.fr

Start Date: March 2022
Duration: 9 months
End Date: November 2022

The objectives of this Micro-Project are to create a software toolkit that makes it possible to achieve realistic human-like motions that can lead to a feeling of trust and comfort towards a robot. The project is based on a generic formalization of robot dancing which maks it possible to use musical features for choreography generation. The DFKI team will develop and evaluate this package using the open-source software Pinocchio developed by the INRIA Paris team with its recently introduced proximal formulation of the constrained dynamics. This will contribute to tasks T2.1, T2.2 , T2.3 and T2.4.

2022-4-WP246 Polarization with the Friedkin-Johnsen model over a dynamic social network

Submitted by: Elisabetta Biondi, elisabetta.biondi@iit.cnr.it
Partners:
1. Consiglio Nazionale delle Ricerche (CNR), Elisabetta Biondi, elisabetta.biondi@iit.cnr.it
2. Central European University (CEU), Janos Kertesz, kerteszj@ceu.edu, Gerardo Iniguez, IniguezG@ceu.edu

Start Date: 2022-04-01
Duration: 4
End Date: 2022-12-31 (delay due maternity leave of Elisabetta Biondi).

The objective of this Micro-Project is to develop methods to develop metrics for evaluating polarization of Human social networks. The Friedkin-Johnsen model is a popular model in opinion dynamics, validated on real groups, and well-investigated from the opinion polarization standpoint. However, this model is static. The goal of this Micro-Project is to design a variant of the Friedkin-Johnsen model that captures the dynamic nature of social networks, based on a novel definition of global polarization that combines network features and opinion distribution. This can contribute to task 2.4 Perception of Social Signals and Social Dynamics.

2022-7-WP234 MSSW: Multilingual SynSemClass for the Semantic Web
Proposed by: Jan Hajic, Charles Univ,
Partners
1. Charles Univ, Jan Hajic
2. DFKI, Thierry deClerck

Start Date: July 2022
Duration: 6 months
End Date: Dec 2022

SynSemClass is a dataset created in a previous Humane AI Net Micro-Project called META-O-NLU. The objective of this micropropject is to convert SynSemClass to a Linguistic Linked Open Data (LLOD) datas, connecting it to the huge amount or interlinked data already available. Linguistic Linked Open Data is a generic term for a set of mutually connected language resources, using ontological relations. The connections between concepts and between concepts and their expression in natural language make them

suitable for both research and industrial applications in the area of content analysis, natural language understanding, inferencing and other tasks.

In the present task, the concrete deliverable will be an LLOD for SynSemClass, connecting it to the huge amount or interlinked data already available. A partner is involved in the Prêt-à-LLOD H2020 project, making this project synergistic in nature and multiplicative in terms of results in previous projects. Partners are also involved in the COST Action "European network for Web-centered linguistic data science" (NexusLinguarum).

2022-9-WP12: Graybox methods for augmenting human-driven narrative analyses

Submitted by: Szymon Talaga, stalaga@uw.edu.pl
Partners:
1. Univ. Warsaw: Andrejz Nowak, Szymon Talaga, stalaga@uw.edu.pl
2. Institut Polytechnique de Grenoble, James Crowley, james.crowley@univ-grenoble-alpes.fr

Start Date: 2022-09-01
Duration: 5 Months
End Date: 2023-01-31

This Micro-Project sets a foundation for developing a new approach to narrative analysis providing a gray-box (at least partially explainable) NLP model tailored for facilitating work of qualitative text/narrative analysts. We have conducted a proof-of-concept study combining existing standard NLP methods (e.g. topic modeling, entity recognition) with qualitative analysis of narratives about smart cities and related technologies and use this experience to conceptualize our approach to narrative analysis, in particular with respect to problems which are not easily solved with the existing tools. This current initial research micro-project will be followed by a Micro-Project dedicated to formalizing our approach to narrative analysis and developing its open-source implementation (for Python)

2022-9-WP2 Multi-Relational Contextual Reasoning for Complex Scene Generation for Autonomous Vehicle Data

Submitted by: Loris Bozzato, bozzato@fbk.eu
Partners:
1. Fondazione Bruno Kessler, Loris Bozzato, ghidini@fbk.eu
2. Technical University of Vienna, Thomas Eiter, eiter@kr.tuwien.ac.at
3. BOSCH Deutschland (External Partner), Stepanova Daria

Start Date: 1 September 2022
Duration: 7 Months
Completion Date 31 March 2023

Description:

MR-CKR is a symbolic reasoning framework for Multi-Relational Contextual Knowledge Repositories that was previously developed by the partners. This Micro-Project has sought to test the applicability of the MR-CKR framework to the task of generating challenging inputs for a machine learning model. Given a set of diagnoses describing contexts in which the model performs poorly we generated new inputs that are (i) in the described contexts and (ii) as similar as possible to a given starting input. (i) allows us to train the network in a

targeted manner by feeding it exactly those cases that it struggles with. (ii) ensures that the new input only differs from the old one in those aspects that make the new input problematic for the model. Thus, allowing us to teach the model to recognize aspects relevant for the answer.

This contributes to T2.2 Multimodal perception and modeling of actions, activities and tasks.

2022-9-WP235 Multilingual and Multimodal conversational agent combined with search engine models

Start Date 1 Sept 2022
Duration: 6
End Date: 28 April 2023 (on-going)

Submitted by: Eric Blaudez, eric.blaudez@thalesgroup.com
Partners:
1. Thales Eric Blaudez, eric.blaudez@thalesgroup.com
2. Unibo Paolo Torrini, p.torroni@unibo.it
3. LISN Christophe Servan, c.servan@qwant.com

Description:

The micro-project provides an demonstration of the hierarchical Framework for collaboration decribed in the Humane-AI Net revised strategic workplan, by constructing amultimodal and multilingual conversational agents focused on search.

The framework is based on hierarchical levels of abilities:

- Reactive (sensori-motor) Interaction: Interaction is tightly-coupled perception-action where actions of one agent are immediately sensed and interpreted as actions of the other. Examples include greetings, polite conversation and emotional mirroring
- Situated (Spatio-temporal) Interaction Interactions are mediated by a shared model of objects and relations (states) and shared models for roles and interaction protocols.

On this Micro-Project we focused on the 2 first levels (Reactive and Situational) and designed the global framework architecture. The results are to be demonstrated in a Proof of Concept (PoC).

Work Package 3: Human AI Collaboration and Interaction

Coordinator: Antti Oulasvirta, Aalto University

Research Objectives for WP 3:

The objective of WP3 "Human AI Interaction and Collaboration" is to establish new methodological and conceptual basis for human-AI collaboration. In particular, the goal is to develop HCI and AI methodology for social basis for human-AI partnership, especially group cognition and emotional expression. For AI to understand people, it needs to both be able to infer intentions and emotions from observations as well as make its own intentions understandable to human partners via grounding, emotional expression, and explanation. We believe that these capabilities need to be to some extent be engineered into AI, in order to ensure more natural behavior from first interaction and to reach a desirable level of controllability and transparency. However, they need to be made interactive for users to control and understand.

The intelligent systems studied in WP3 may be in the form of embodied agents, be them physical robots, animated characters, physical interactive objects, smart environments, or simply software systems. Humans and systems may interact through visual displays, physical devices, acoustic signals, printed text, spoken language, or other modalities. This WP will work to advance and implement the HumanE AI Net vision of allowing such interaction to take the form of synergetic collaboration and co-creation leveraging the new Human in the Loop learning, reasoning and planning methods of WP 1 and the advances in perception and world modeling from WP2.

WP3 builds on human-computer interaction, especially on theories of human-machine communication, while integrating them with relevant advances in machine learning methods. Work on human AI interaction and collaboration is organized around the following tasks:

T3.1 Foundations of Human-AI interaction and Collaboration;
T3.2 Human-AI Interaction/collaboration paradigms;
T3.3 Reflexivity and Adaptation in Human AI collaboration;
T3.4 User Models and Interaction History;
T3.5 Visualization Interactions, and Guidance;
T3.6 Language-based and Multilingual Interaction;
T3.7 Conversational, Collaborative AI; and
T3.8 Trustworthy Social and Sociable interaction.

Special objective in Project Year 2:

Recent breakthroughs in AI have shown proficiency in interactions with the natural world and with language, however coordination and collaboration with human partners is an open challenge. Our refined objective focuses on methods, theories, studies, and techniques for common ground. Common ground refers to shared (common) beliefs and goals related to a shared activity. When there is no common ground, repairs and compensations may be needed. Participants must correct misunderstandings or take time to re-establish common ground. Toward Project Year 2, we invited proposals considering both technical and human aspects of grounding. Startred in the Venice meeting in May

2022, and elaborated in the Stockholm workshop in November 2022, topics started in 2023 included:

- Exploiting context-awareness for grounding
- Grounding with Large Language Models
- Pragmatics, including linguistic and embodied aspects
- Affordances for grounding
- Co-adaptive processes in grounding
- Storytelling and narratives
- Information retrieval
- Speech-based and multimodal interaction with AI
- Cultural factors affecting grounding
- Empirical measurements of grounding
- Design processes for grounding
- Special application areas with specific requirements for grounding, such as translation, games, explainable AI, etcetera.

Our Micro-Projects for this period cover one or more of the aforementioned challenges.

Overview of results

The new theme of interactive grounding has seen first results already from the Micro-Projects: A summary of main findings so far:

- There are many subconscious effects in play in human-AI interaction. Social psychology, which has studied many such effects but in the context of human-human interaction, can shed light to those. Partners in the HumaneAI Net found out that synchronization of movements between AI and humans can help improve trust between. (Bartkowski et al. CHI'23)
- Conversational context is part of the common ground between AI and human collaborators. With LLMs like ChatGPT, this context builds up from the prompts and responses between the partners. However, it is unclear how humans perceive the context. This depends partially on what remember and how they construe conversational contexts. Scholars in HumaneAI Net collected a rich empirical dataset to understand this and to build better AI partners (Jonker et al.)
- An empirical study found that AI assistants that support novices in complex design tasks (here: 3D fabrication), could better improve common ground if they understood users' commands from users' point-of-view and allowed referencing objects using familiar vocabulary (submission by Müller et al.)
- XR (extended reality) technology presents a nascent opportunity for new forms human-AI interaction with special demands for interactive grounding. In this domain, AI is presently mostly used for content creation. (A review of AI for XR authored by Hirzle et al., CHI'23)
- Agency presents a challenge to human-AI collaborations; there is emerging European network of scholars focusing on this topic (Shin et al., a workshop at CHI'23)
- In some professional application domains, common ground is multimodal and needs to be interactively shaped. Schetinger et al. are studying the design of storyboards (e.g., for movies) as a case.

Selected top papers and workshop contributions from WP3 from this period:

- Teresa Hirzle, Florian Müller, Fiona Draxler, Martin Schmitz, Pascal Knierim, and Kasper Hornbæk. 2023. When XR and AI Meet – A Scoping Review on Extended Reality and Artifcial Intelligence. In **Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)**, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 45 pages. https://doi.org/10.1145/3544548.3581072
- Wieslaw Bartkowski, Andrzej Nowak, Filip Ignacy Czajkowski, Albrecht Schmidt, and Florian Müller. 2023. In Sync: Exploring Synchronization to Increase Trust Between Humans and Non-humanoid Robots. In **Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)**, April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3544548.3581193
- Tewari M and Lindgren H (2022), Expecting, understanding, relating, and interacting - older, middle-aged and younger adults' perspectives on breakdown situations in human–robot dialogues. **Front. Robot. AI** 9:956709. doi: 10.3389/frobt.2022.956709.
- Kilic K, Weck S, Kampik T, Lindgren H. Argument-Based Human-AI Collaboration for Supporting Behavior Change to Improve Health. to appear in **Front. AI, 2023**.
- Joongi Shin, Janin Koch, Andrés Lucero, Peter Dalsgaard, Wendy E. Mackay. Integrating AI in Human-Human Collaborative Ideation. **Extended Abstracts (workshop) CHI 2023 - SIGCHI conference on Human Factors in computing system**s, Apr 2023, Hamburg, Germany. pp.1-5. ⟨hal-04023507⟩
- Tsfasman, M., Fenech, K., Tarvirdians, M., Lorincz, A., Jonker, C., & Oertel, C. (2022). Towards creating a conversational memory for long-term meeting support: predicting memorable moments in multi-party conversations through eye-gaze. In **ICMI 2022 - Proceedings of the 2022 International Conference on Multimodal Interaction** (pp. 94-104). (ACM International Conference Proceeding Series). Association for Computing Machinery (ACM). https://doi.org/10.1145/3536221.3556613

Contributions from Micro-Project to Work Package 3

2021-3-WP3: Exploring the impact of Agency on Human-Computer Partnerships

Contributed to understanding of agency in human-computer partnerships. In a review paper provided a synthesis of the concept, identified which parameters relevant for the description of a system's agency, and proposed techniques to modify them.

2021-3-WP3: Exploring the impact of Agency on Human-Computer Partnerships

Contributed to understanding of agency in human-computer partnerships. In a review paper provided a synthesis of the concept, identified which parameters relevant for the description of a system's agency, and proposed techniques to modify them.

2021-3-WP123: Memory Aware Conversational AI to aid virtual Team-Meetings

Studied human recollection of team meetings and how conversational AI could use this information to create better team cohesion in virtual settings. Used the results to create memory aware conversational AI that can leverage such data to increase team cohesion.

2022-3-WP3, Optimal Alerting

Studied how AI assistants could better alert humans by estimating their internal states from observations. Contributed to a computational theory called POSG, a multi-agent

framework for human-AI interaction developed between CNRS and Aalto University: https://jgori-ouistiti.github.io/CoopIHC/

2022-4-WP3: Making for Everyone: Interactive, voice-based support for the design of 3D objects for digital fabrication

Studied an AI-powered system that assists users in creating 3D objects for digital fabrication. The key contribution to WP3 is a natural language processing (NLP) solution for grounding. It enables users to describe objects using their natural language (e.g., "A green rectangular box."). This supports the goal of making personal digital fabrication accessible for everyone.

2022-5-WP3: In Sync: Synchronization in interaction between Humans and Embodied AI Systems

Demonstrated empirically that AI systems that synchronize with their human partners are better trusted by them.

2022-7-WP3: A Scoping Review on Artificial Intelligence and Extended Reality

Reviewed literature on XR (extended reality) to understand opportunities for human-AI interaction and demand for interactive grounding.

2022-9-WP23: Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems

Studied context exchange in conversational interaction with AI, supporting interactive grounding in two senses - among dialog sides (for example, technologies like example automatic speech transcription rarely use the other dialog side information to adapt itself) and among dialog system components (the speech synthesis rarely uses dialog context to produce more fluent or expressive speech).

2022-10-WP3 3D Semantic Label Transfer and Matching in Human-Robot Collaboration

Contributed a semantic segmentation algorithm that allows robots to estimate the object perception of a human partner, which can be used for grounding in human-robot interaction.

2023-1-WP3: Storyboarder: combining Play Script and Image Generation

Demonstrated multimodal grounding techniques for LLMs in the task of generating storyboards that combine text and image.

New Microproject Procedures from January 2023

Following publication of the modified Humane AI Net Strategic Research agenda, documented in HAI Net Deliverable D6.1, during the consortium general assembly in October 2022 it was decided to implement new procedures concerning proposal and funding of micro-project. The procedure have come into force in January 2023.

Starting January 2023, the consortium will publish maintain a current list of research challenges that are to be updated 3 or 4 times year. Consortium partners may propose micro-projects responding to the current call at any time. Changes to the call are discussed in regular meetings of each workpackage, prior to publication of the new challenges. The administrative rules of the micro projects are as before (at least two partners, ideally 2-6 months duration, with 2-4 PMs per partner and obligation to produce a tangible result to be made available through the AI4Europe platform. Deviations from these guidelines may be approved by the project management committee with suitable justification.

The HumaneAI Net website must now be used for submission proposals for micro-projects. Each microproject should include visits between the sites. Partners who have run out for funds can now apply for micro-projects from our "reserve fund". Preference for use of reserve funds will be given to projects that involve external partners those that involve industry

For the external group travel funds (including subsistence for longer stays) will be paid by the consortium coordinator. No person months can be paid for the external partners.

All proposals that meet the following evaluation Criteria will be automatically approved:

- Alignment with call objectives and the specific research direction
- Feasibility, innovation and societal relevance of the proposed approach
- Measurable impact potential of the solution on theories, methods, and societal or economic/business impact
- Quality of the proposed collaboration and partnership

The following calls for proposals have been published in January 2023 for workpackages 1, 2, and 3:

Establishing Common Ground for Collaboration with AI Systems (WP1 and WP2)

The focus of this call for micro-projects is on 'Collaborative Artificial Intelligence' as described in Section 3.1 of the updated Humane AI Research Agenda (HAI Net Deliverable D6.1.pdf). We are interested in micro-projects and clusters of micro-projects that seek to provide practical demonstrations, tools, or new theoretical models for AI systems that can collaborate with and empower individuals or groups of people to attain shared goals.

Recent progress in machine learning has provided a variety of powerful new enabling technologies for intelligent systems. We seek to harness these advances in areas such as large language models, generative systems, adversarial learning, self-supervised learning, visual object detection and natural language understanding to provide new foundation for establishing common ground for collaboration with AI systems.

We are interested in AI systems that can both communicate and understand descriptions of situations, goals, intentions or operational plans in order to establish shared

understanding for collaboration. Descriptions may be expressed by sounds, motion, mechanical forces, visual displays, natural language, or any other communication mode, but must be expressed in a manner that is comprehensible to a human partner. We are particularly interested in theories and demonstrations of systems that can explain their internal models by providing additional information to justify statements and answer questions such as who, what, where, when, why and how.

Examples of targeted outcomes could include tools, demonstrations or theoretical models for topics such as, but not limited to

1) Systems that use sounds and visual displays and/or mechanical forces to guide and assist human operations in dynamically changing environments;

2) Systems that use natural language or other interactions to obtain a model of the goals and intentions of a partner in order to provide information, explanations, warnings, or suggest possible courses of actions;

3) Systems that can interpret narrative descriptions of events in order to verify facts, answer questions to provide explanations for events or provide additional information.

The ambition should be to enable interaction between AI systems and humans that is two-way in order to construct shared representations, and include the possibility of defining new representations for perceptions, actions or events. This has the advantage that the shared representation does not necessarily need to be finalised before interaction begins, but can be adapted in response to the exchange of information analogous to the way that humans often synchronise their understanding with explanations of different terms, concepts, or situations. Micro-projects that explore this or other aspects of interactive alignment are encouraged even if they are only aiming at modest complexities of representation. In general, we encourage applications that aim to demonstrate modest but measurable steps towards the more ambitious goals!

We are keen that Micro-projects should demonstrate tangible progress towards the more ambitious goals described above and so encourage applicants to provide concrete measures of progress towards Collaborative AI that their project will monitor. At the same time in order to keep the projects aligned with the larger goals we also encourage projects to document how the work addresses those goals and if appropriate we also encourage you to include in your submission applications for follow-on MPs that can build on the work of the initial MP if that proves successful. We would consider giving pre-approval to follow-on projects conditional on the initial project successfully meeting its objectives.

Interactive Grounding (WP3)

Recent breakthroughs in AI have shown proficiency in interactions with the natural world and with language, however coordination and collaboration with human partners is an open challenge.

This topic focuses on methods, theories, studies, and techniques for common ground. Common ground refers to shared (common) beliefs and goals related to a shared activity. When there is no common ground, repairs and compensations may be needed. Participants must correct misunderstandings or take time to re-establish common ground.

This call invites proposals considering both technical and human aspects of grounding. Based on the Stockholm workshop, topics in 2023 include but are not limited to:

- Exploiting context-awareness for grounding
- Grounding with Large Language Models
- Pragmatics, including linguistic and embodied aspects

- Affordances for grounding
- Co-adaptive processes in grounding
- Storytelling and narratives
- Information retrieval
- Speech-based and multimodal interaction with AI
- Cultural factors affecting grounding
- Empirical measurements of grounding
- Design processes for grounding
- Special application areas with specific requirements for grounding, such as translation, games, explainable AI, etcetera.

We invite micro-projects covering one or more of the aforementioned challenges. Micro-projects with a focus on industrial applications and societal use cases are also welcome. All proposals must make clear how they contribute to the theme of the call: interactive grounding.

Summary of Micro Projects During the Second Period

Nine micro-projects initiated during the project first period had completion dates after the 1 Sept 2021 and have results reported below. Ten micro-project were started and completed during the second project period and are reported below. Two micro-projects are reported to be underway at the end of the second project period and will report results in the deliverable covering the third project period.

As explained above, In order to help organize a quick overview of micro projects and their contributions to work packages, we have adopted the following notation yyyy-mm-WPnnn as a label for Micro-Projects, where yyyy refers to the year of the start date of the Micro-Project (2021, 2022, 2023 or 2024), mm refers to the month of the of the start date of the Micro-Project (01 to 12), and nnn refers to the work packages to which the Micro-Project contributed. In cases where more than one Micro-Project contributing to the same Work packages start in the same month, we add an extra field -mm.

Micro Projects Started in 1st year and Completed after 1 September 2021

Nine micro-projects initiated during the project first period were completed after the 1 Sept 2021 and have results reported below.

2021-01-WP13: Coping with the variability of human feedback during interactive learning through ensemble reinforcement learning

Start Date: January 2021
Duration: 15 Months
Completion Date April 2023

Proposed By: Mehdi Khamassi, mehdi.khamassi@sorbonne-universite.fr

Partners:

1. Sorbonne University, Mehdi Khamassi, mehdi.khamassi@sorbonne-universite.fr
2. ATHINA Research Center, Petros Maragos, maragos@cs.ntua.gr

Results:

This 6-month project entails robot online behavioral adaptation during interactive learning with humans. Specifically, the robot shall adapt to each human subject's specific way of giving feedback during the interaction. Feedback here includes reward, instruction and demonstration, and can be regrouped under the term "teaching signals". For example, some human subjects prefer a proactive robot while others prefer the robot to wait for their instructions; some only tell the robot when it performs a wrong action, while others reward correct actions, etc. The main expected outcome was a new ensemble method of human-robot interaction which can learn models of various human feedback strategies and use them for online tuning of reinforcement learning so that the robot can quickly learn an appropriate behavioral policy.

We designed a new ensemble learning algorithm, combining model-based and model-free reinforcement learning, for on-the-fly robot adaptation during human-robot interaction. The algorithm includes a mechanism for the robot to autonomously detect changes in a human's reward function from its observed behavior, and a reset of the ensemble learning accordingly. We simulated a series of human-robot interaction scenarios to test the robustness of the algorithm. In scenario 1, the human rewards the robot with various

feedback profile: stochastic reward; non-monotonic reward; or punishing for error without rewarding correct responses. In scenario 2, the humans teaches the robot through demonstrations, again with different degrees of stochasticity and levels of expertise from the human. In scenario 3, we simulated a human-robot cooperation task for putting a set of cubes in the right box. The task includes abrupt changes in the target box. Results show the generality of the algorithm.

Humans and robots are doomed to cooperate more and more within the society. This micro-project addresses a major AI challenge to enable robots to adapt on-the-fly to different situations and to different more-or-less naive human users. The solution consists in designing a robot learning algorithm which generalizes to a variety of simple human-robot interaction scenarios. Following the HumanE AI vision, interactive learning puts the human in the loop, prompting human-aware robot behavioral adaptation.

The micro-project directly contributes to one of the objectives of WP1 (T1.3) to enable "continuous incremental learning in joint human/AI systems" by 'exploiting rich human feedback'. It also directly contributes to one of the objectives of WP3 (T3.3) to enable reflexivity and adaptation in Human AI collaboration.

Publications:

Rémi Dromnelle, Erwan Renaudo, Benoît Girard, Petros Maragos, Mohamed Chetouani, Raja Chatila, Mehdi Khamassi (2022). Reducing computational cost during robot navigation and human-robot interaction with a human-inspired reinforcement learning architecture. International Journal of Social Robotics, doi: 10.1007/s12369-022-00942-6

Links to Tangible results:

- **Open source code**: https://github.com/DromnHell/meta-control-decision-making-agent
- **Publication**: Preprint made open on HAL – https://hal.sorbonne-universite.fr/hal-03829879

2021-03-WP123: Memory Aware Conversational AI to aid virtual Team-Meetings

Start Date: 2 March 2021
Duration: 24 Months
Completion Date 2 March 2023

Submitted by: Catholijn Jonker, c.r.m.m.oertel@tudelft.nl

Other Partners:
1. Technical University Delft, Catharine Oertel, c.r.m.m.oertel@tudelft.nl (4 PM)
2. Eotvos Lorand University, Andras Lorincz, lorincz@inf.elte.hu (2 PM)

Contributes to: WP1, WP2, WP3, T3.2

Results:

The MEMO corpus was collected, which contains 45 group discussions around the topic of COVID-19. A total of 15 groups were formed, consisting of 3 to 6 participants who took part in 3 group discussions, with a 3-4 day gap between sessions. A total of 59 individuals with diverse backgrounds took part in the study. Before and after each session participants completed a series of questionnaires to determine which moments they recalled from their conversations, along with their personality traits, values and perceptions.

To capture conversational memory, we collected first-party free-recall reports of the most memorable moments from the discussion immediately after the interaction and again 3-4 days later. For the shorter-term memories, participants also mapped the moments to a particular interval in the video of their discussion, which were used for the ground-truth conversational memory annotations.

For each participant, personality and value profiles were recorded in the pre-screening survey, along with demographic information to identify their social group affected by COVID-19. Pre-session questionnaires also assessed participants' mood before each session. Post-session questionnaire included questions about mutual understanding, personal attitude and perceived social distance. The perception of the discussion and the group as a whole was also monitored in the post-session questionnaire with variables such as Task & Group Cohesion, Entitativity, Perceived Interdependence, Perceived Situation Characteristics, Syncness, and Rapport.

The following automatic annotations were extracted on the corpus:

Transcripts - Transcripts were generated with automatic speech recognition methods and were manually reviewed and corrected where needed. Transcript timestamps are available at the utterance level as well as word-level textgrid files for each recording. Speaker diarization is also available.

Eye gaze and head pose - automatically annotated with EyeWare software, the annotation itself will be provided, but the code uses proprietary API. This includes gaze targets collected through screenshots of participants' screenviews.

Prosody - eGeMAPS feature set was extracted from the default eGeMAPS configuration in OpenSmile

Body pose - Body pose (upper body only) and hand pose when visible were estimated with the models available in the MediaPipe software

Facial action units - Facial action units were estimated for participants using the OpenFace **Software**

A **Paper** describing the corpus and the annotations in more detail is in preparation. Additionally, the collected annotations are to be packaged in an appropriate manner for ease of use for future researchers.

The resulting corpus and corresponding manual and automatic annotations address the goals of work package 3 by providing a new and unique dataset for obtaining a better understanding of conversational memory and the multimodal signals associated with identifying memorable moments which may be used to provide virtual agents the ability to identify topics and episodes which can enhance the interaction between human and machine.

Publications: :

1. Tsfasman, M., Fenech, K., Tarvirdians, M., Lorincz, A., Jonker, C., & Oertel, C. (2022). Towards creating a conversational memory for long-term meeting support: predicting memorable moments in multi-party conversations through eye-gaze. In ICMI 2022 - Proceedings of the 2022 International Conference on Multimodal Interaction (pp. 94-104). (ACM International Conference Proceeding Series). Association for Computing Machinery (ACM). https://doi.org/10.1145/3536221.3556613

Links to Tangible results:

- The corpus is available by request - contact Maria Tsfasman (m.tsfasman@tudelft.nl) and Catharine Oertel (c.r.m.m.oertel@tudelft.nl). It is planned that the corpus will be released by the end of 2024.
- For any other questions and inquiries feel free to contact Maria Tsfasman (m.tsfasman@tudelft.nl) and/or Catharine Oertel (c.r.m.m.oertel@tudelft.nl).
- https://www.youtube.com/watch?v=kTS8XwUVNnI

2021-03-WP3: Exploring the impact of Agency on Human-Computer Partnerships

Start Date: 2022-03-10
Duration: 4
End Date: 2022-09-01

Submitted by: Janin Koch, Janin.Koch@inria.fr

Partners:
1. Ludwig-Maximilians-Universität München, Albrecht Schmidt
2. Københavns Universitet, Kasper Hornbaek
3. Stichting VU, Koen Hindriks
4. Umeå University, Helena Lindgren

Contributes to: WP3: T3.1, T3.3

Results:

The project's objective was to examine the theoretical and practical contributions of agency to successful human-computer partnership. Understanding agency is critical for establishing effective collaboration in human-centered AI research.

The goal of this project was to 1) produce an overview of existing HCI and AI literature on agency, 2) hold a workshop to brainstorm and categorize agency components, define metrics and experimental protocols, and 3) create interactive demos demonstrating various forms of human and system agency.

We conducted individual and collaborative brainstorming sessions with all participants to create an initial overview of current literature in order to establish a common starting point (1). We talked about potential overlaps in our work and how such perspectives influence our current work.

We will hold a workshop at CHI'23 on 'Integrating AI in Human-Human Collaborative Ideation' to examine the role AI can play in such interactive environments in order to identify distinct dimensions and measures of agency within human-ai interaction (2) [Shin et al. , 2023].

Umeå has also investigated how conversations between a human and a socially intelligent robot in a home environment can influence perceptions of agency [Tewari and Lindgren, 2022] and the importance of goal setting in such a scenario [Lindgren and Weck, 2022; Kilic et al., 2023] (3).

While there is currently project underway, we will report our findings in the second half of this year.

Publications:

1) Tewari M and Lindgren H (2022), Expecting, understanding, relating, and interacting - older, middle-aged and younger adults' perspectives on breakdown situations in human–robot dialogues. Front. Robot. AI 9:956709. doi: 10.3389/frobt.2022.956709.
2) Kilic K, Weck S, Kampik T, Lindgren H. Argument-Based Human-AI Collaboration for Supporting Behavior Change to Improve Health. to appear in Front. AI, 2023.
3) Lindgren H and Weck S. 2022. Contextualising Goal Setting for Behaviour Change – from Baby Steps to Value Directions. In 33rd European Conference on Cognitive Ergonomics (ECCE2022), October 4–7, 2022, Kaiserslautern,Germany. ACM, New York, NY, USA. https: //doi.org/10.1145/3552327.3552342
4) Joongi Shin, Janin Koch, Andrés Lucero, Peter Dalsgaard, Wendy E. Mackay. Integrating AI in Human-Human Collaborative Ideation. CHI 2023 - SIGCHI conference on Human Factors in computing systems, Apr 2023, Hamburg, Germany. pp.1-5. ⟨hal-04023507⟩

Links to Tangible results:

- https://www.frontiersin.org/articles/10.3389/frobt.2022.956709/full
- https://www.frontiersin.org/articles/10.3389/frai.2023.1069455/full
- https://dl.acm.org/doi/abs/10.1145/3552327.3552342 (Tb released in the proceedings of CHI)

2021-04-WP1: Discovering Temporal Logic patterns as binary supervised learning

Proposed by: Chiara Ghidini, (FBK)

Start Date: 14 April 2021
Duration: 5 Months
Completion Date 14 September 2021

Proposed By: Chiara Ghidini, ghidini@fbk.eu

Partners:

1. University of Bologna, Federico Chesani, fchesani@gmail.com
2. Free University of Bozen-Bolzano (external Partner), Sergio Tessaris, tessaris@inf.unibz.it

Results:

The microproject has produced three main results:

- A two-step approach for the discovery of temporal-logic patterns as a binary supervised learning problem, that is starting from a set of "positive traces" (execution traces whose behaviour we want to observe in the discovered patterns), and a set of "negative" traces (execution traces whose behaviour we do not want to observe in the discovered patterns). In detail, in the first step, sets of patterns (possible models) that accept all positive traces and discard as much as possible of the negative ones are discovered. In the second step, the model(s) optimizing one criterion, as for instance the generality or the simplicity, are selected among the possible discovered models.

- Two synthetic labelled ("positive" and "negative") event log datasets used for the synthetic evaluation of the proposed approach.

- A paper describing the approach, as well as the approach evaluation.

The microproject has produced three main results:

- A two-step approach for the discovery of temporal-logic patterns as a binary supervised learning problem, that is starting from a set of "positive traces" (execution traces whose behaviour we want to observe in the discovered patterns), and a set of "negative" traces (execution traces whose behaviour we do not want to observe in the discovered patterns). In detail, in the first step, sets of patterns (possible models) that accept all positive traces and discard as much as possible of the negative ones are discovered. In the second step, the model(s) optimizing one criterion, as for instance the generality or the simplicity, are selected among the possible discovered models.

- Two synthetic labelled ("positive" and "negative") event log datasets used for the synthetic evaluation of the proposed approach.

- A paper describing the approach, as well as the approach evaluation.

Publications:

1. Chesani, F., Francescomarino, C. D., Ghidini, C., Grundler, G., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Shape your process: Discovering declarative business processes from positive and negative traces taking into account user preferences. In Almeida, J. P. A., Karastoyanova, D., Guizzardi, G., Montali, M., Maggi, F. M., and Fonseca, C. M., editors, Enterprise Design, Operations, and Computing - 26th International Conference, EDOC 2022, Bozen-Bolzano, Italy, October 3-7, 2022, Proceedings, volume 13585 of Lecture Notes in Computer Science, pages 217–234. Springer.
2. Chesani, F., Francescomarino, C. D., Ghidini, C., Grundler, G., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Shape your process: Discovering declarative business processes from positive and negative traces taking into account user preferences. In Almeida, J. P. A., Karastoyanova, D., Guizzardi, G., Montali, M., Maggi, F. M., and Fonseca, C. M., editors, Enterprise Design, Operations, and Computing - 26th International Conference, EDOC 2022, Bozen-Bolzano, Italy, October 3-7, 2022, Proceedings, volume 13585 of Lecture Notes in Computer Science, pages 217–234. Springer.
3. Chesani, F., Francescomarino, C. D., Ghidini, C., Loreti, D., Maggi, F. M., Mello, P., Montali, M., and Tessaris, S. (2022). Process discovery on deviant traces and other stranger things. IEEE Transactions on Knowledge and Data Engineering, pages 1–17. DOI: https://doi.org/10.1109/TKDE.2022.3232207

Links to Tangible results:

1. Loan Approval1: dataset. https://drive.google.com/drive/folders/15BwG4PJq8ilMh9Sr9dpMXAYBY-qp7QDE?usp=sharing
2. Loan Approval2: dataset. https://drive.google.com/drive/folders/1fcJ8itzdMbNOjEAeV6nUEel5B6__aB_c?usp=sharing
3. **Discovery Framework**: program/code https://zenodo.org/record/5158528
4. **Experiments**: https://github.com/stessaris/negdis-experiments/tree/v1.0

2021-04-WP123 Multimodal Perception and Interaction with Transformers

Proposed by: James Crowley, Inria

Other Partners:

1. Andras Lorincz (Eotvos Lorand University - ELTE),
2. Dominique Vaufreydaz, Fabien Ringeval (Univ Grenoble Alpes),
3. Camille Guinaudeau, Marc Evrard (Uni Paris Saclay,

4. Marko Grobelnik (Jozef Stefan Institut-JSI),
5. Pavel Pecina (Charles University)

Start Date: April 2021
Duration: 6 Months
Completion Date October 2021

Transformers and self-attention (Vaswani et al., 2017), have become the dominant approach for natural language processing (NLP) with systems such as BERT (Devlin et al., 2019) and GPT-3 (Brown et al., 2020) rapidly displacing more established RNN and CNN structures with an architecture composed of stacked encoder-decoder modules using self-attention.

This micro-project will survey tools and data sets for experiments for demonstrating the potential use of transformers for multimodal perception and multimodal interactions. We will define research challenges and performance metrics for multimodal perception tasks such as audio-visual narration of scenes, cooking actions and activities, audio-visual deictic (pointing) gestures, and perception and evocation of engagement, attention, and emotion. We will provide tutorials on the use of transformers for multimodal perception and interaction.

Expected Outputs:

- Performance targets for a phased set of research challenges of increasing difficulty.
- A survey of tools and benchmark data for experiments to explore use of embeddings, encoder-decoders, self-attention architectures and related problems associated with applying transformers to different modalities.
- Concept demonstrations for simple examples of multimodal perception.

Tangible Results (available on the AI4Europe platform

- A survey of tools and datasets for a multimodal perception with transformers (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/TransformerTools.pdf)
- A tutorial on the use of transformers for multimodal perception. (http://crowley-coutaz.fr/jlc/Courses/ACAI2021/Multimodal-Transformer-Tutorial.html)
- Report on challenges for the use of transformers for multimodal perception and interaction. (http://crowley-coutaz.fr/jlc/HumanE-AI-Net/TransfomerMicroProject/ReseachChallengesDataSets.pdf)

Publications

1) Y. Wang, X. Shen, S. Hu, Y. Yuan, J. L. Crowley, D. Vaufreydaz, Self-Supervised Transformers for Unsupervised Object Discovery using Normalized Cut. IEEE International Conference on Computer Vision and Pattern Recognition, CVPR 2022, pp14543-14553, New Orleans, Jun 2022.
2) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin and J. L Crowley, "Emotion Recognition with Pre-Trained Transformers Using Multimodal Signals", 10th International Conference on Affective Computing and Intelligent Interaction (ACII), Oct 2022.
3) J. Vazquez-Rodriguez, G. Lefebvre, J. Cumin, J. L. Crowley. Transformer-Based Self-Supervised Learning for Emotion Recognition. 26th International Conference on Pattern Recognition (ICPR 2022), Aug 2022, Montreal, Canada.

Related WP: WP1, WP2, WP3

Connection of Results to Work Package Objectives:

This Micro-Project has aided and encouraged the use of a transformers and self-attention for multimodal modal interaction by Humane AI Net researchers, by identifying relevant

tools and benchmark data sets, by providing tutorials and training materials for education, and by identifying research challenges for multimodal perception and interaction with Transformers.

Bibliography

1) Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. and Polosukhin, I. (2017). Attention is all you need. arXiv preprint arXiv:1706.03762
2) Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805.
3) Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., and Amodei, D. (2020). Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

2021-05-WP3 Memory Aware Conversational AI to aid virtual Team-Meetings

Start Date: 1 May 2021
Expected Duration: 6 Months
Completion Date: 31 November 2021

In this micro-project, we propose investigating human recollection of team meetings and how conversational AI could use this information to create better team cohesion in virtual settings.

Specifically, we would like to investigate how a person's emotion, personality, relationship to fellow teammates, goal and position in the meeting influences how they remember the meeting. We want to use this information to create memory aware conversational AI that could leverage such data to increase team cohesion in future meetings.

To achieve this goal, we plan first to record a multi-modal dataset of team meetings in a virtual setting. Second, administrate questionnaires to participants in different time intervals succeeding a session. Third, annotate the corpus. Fourth, carry out an initial corpus analysis to inform the design of memory-aware conversational AI.

This micro-project will contribute to a longer-term effort in building a computational memory model for human-agent interaction.

Expected Outputs:

- A corpus of repeated virtual team meetings (6 sessions spaced, 1 week each)
- manual annotations (people's recollection of the team meeting etc.)
- automatic annotations (e.g., eye-gaze, affect, body posture etc.)
- A paper describing the corpus and insights gained on the design of memory-aware agents from initial analysis

2021-05-WP26: Neural Mechanism in Human Brain Activity During Weight Lifting

Start Date: 1 May 2022
Proposed Duration: 4 months
End Date: 31 Jan 2023

Submitted by: Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr

Partners:

1. TUBITAK BILGEM, Sencer Melih Deniz, sencer.deniz@tubitak.gov.tr (4 PM)
2. DFKI Kaiserslautern, Hamraz Javaheri, Hamraz.Javaheri@dfki.de (4 PM)

Contributes to: WP2, WP6, T2.2, T2.3, T6.5, T6.6

Description:

In this project, it was investigated whether EEG (Electro-Encephalo-Graphy) signal can be used for detecting the motion as well as the variable weights a person is lifting. To do this, an experimental paradigm has been designed and EEG data have been acquired during performing biceps flexion-extension motions for different weight categories: lifting with no weight (empty), medium, and heavy lifting. The outcomes of the project can be applied in industrial exoskeleton applications as well as physical rehabilitation of stroke patients.

Results:

Features in EEG data generating difference for each lifted weight of category have been investigated. EEG data via different two EEG headsets have been collected from various participants while they lift different categories of load, namely empty, medium and heavy, in this project. Then, EEG data have been analyzed to realize if different category of weigths result in difference in EEG data by applying different deep learning methods together with different machine learning methods. According to the obtained results, it can be said that that EEG signals can be successfully used as a method to predict different loads during dynamic bicep curl motion. Therefore, this result could result more researches to develop rehabilitation systems robust to dynamic changes in weight. Moreover, information regarding weight change could contribute to a better estimation of fatigue condition to be used in sports and training applications. Finally, it has been evaluated that the approach to predict different categories of lifted weight could be used in further optimizations in industrial applications for which usage of exoskeleton can be given as an example.

Results of Micro-Project in which TUBITAK BILGEM and Deutsches Forschungszentrum für Künstliche Intelligenz (DFKI) Kaiserslautern collaborated was presented at the IEEE-EMBS International Conference on Biomedical and Health Informatics jointly organised with the IEEE-EMBS International conference on Wearable and Implantable Body Sensor Networks organized in Ioannina, Greece between 27-30 September 2022. Also it was published with the title “Prediction of Lifted Weight Category Using EEG Equipped Headgear" in 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics Conference Proceedings.

Publications:

"Prediction of Lifted Weight Category Using EEG Equipped Headgear", published in 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics.

Links to Tangible results:

Paper: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9926744

Dataset: https://www.ai4europe.eu/research/research-bundles/neural-mechanism-human-brain-activity-during-weight-lifting?category=ai_assets

2021-07-WP13 Interactive Reinforcement Learning for Humorous Agents

Start Date: July 2021
Duration: 6 Months
Completion Date: Dec 2021

This project aims at investigating the construction of humor models to enrich conversational agents through the help of interactive reinforcement learning approaches.

Our methodology consists in deploying an online platform where passersby can play a game of matching sentences with humorous comebacks against an agent.

The data collected from these interactions will help to gradually build the humor models of the agent following state of the art Interactive Reinforcement Learning techniques.

We plan to work on this project for 4 months, resulting in an implementation of the platform, a first model for humor-enabled conversational agent and a publication of the obtained results and evaluations.

Expected Outputs:

- Online game for collecting humorous interaction data
- Humor models for conversational agents
- Paper in International Conference of Journal related to AI and AI in Games

Actual Outputs:

- Dataset - 1712 jokes, rated on a scale of 1 to 9 in terms of joke level, originality, positivity, entertainment, whether it makes sense and whether it is family-friendly, (dataset)
- Online Game - A game of matching sentences with humorous comebacks against an agent (like the game Cards Against Humanity), (program/code)

Connection of Results to Work Package Objectives:

The micro-project produced for HumaneAI-net a dataset of annotated associations between black and white cards following the game design of Cards Against Humanity. By doing so, the micro-project led to the creation of a unique dataset of humorous associations between concepts, annotated in terms of different humor styles by the participants of the experiment. The preliminary analysis on how the dataset can be leveraged to build different humor models for conversational agents is particularly relevant for the tasks T3.3 and 3.4 of WP3. Additionally, the Micro-Project aims at exploring how to refine the humor models through an interactive learning approach, particularly relevant for the task T1.3 of WP1.

2021-06-WP136 Adaption of ASR for Impaired Speech with minimum resources (AdAIS)

Start Date: June 15, 2021
Actual Duration: 5 Months
Completion Date: Oct 15, 2021

This Micro-Project will study the adaptation of automatic speech recognition (ASR) systems for impaired speech. Specifically, the micro-project will focus on improving ASR systems for speech from subjects with dysarthria and/or stuttering speech impairment types of various degrees. The work will be developed using either German “Lautarchive” data comprising only 130 hours of untranscribed doctor-patient German speech conversations and/or using English TORGO dataset. Applying human-in-the-loop methods we will spot individual errors and regions of low certainty in ASR in order to apply human-originated improvement and clarification in AI decision processes.

Contributing to tasks:

WP1 Learning, Reasoning and Planning with Human in the Loop
T1.1 Linking symbolic and sub-symbolic learning

WP3 Human AI Interaction and Collaboration
T3.1 Foundations of Human-AI interaction and Collaboration
T3.6 Language-based and multilingual interaction
T3.7 Conversational, Collaborative AI

WP6 Applied research with industrial and societal use cases
T6.3 Software platforms and frameworks
T6.5 Health related research agenda and industrial use cases

Expected Outputs:

- Paper for ICASSP 2021 and/or Interspeech 2022

Micro Projects started and Completed During 2nd period

Ten micro-project were started and completed during the second project period and are reported below.

2022-03-WP2: Natural Imitation of Dance Moves and Human Gestures with a Humanoid Robot

Start Date: 1 March 2022
Proposed Duration: 9 months
End Date: 30 Nov 2022

Submitted by: Shivesh Kumar, shivesh.kumar@dfki.de

Partners:
1. DFKI Bremen ,Melya Boukheddimi, Shivesh Kumar, shivesh.kumar@dfki.de (5 PM)
2. INRIA Paris, Justin Carpentier, justin.carpentier@inria.fr (4 PM)

Contributes to: WP2

Results:

In order to address the topic, two main contributions were introduced in this project: The first one is the proposition of a generic formalization of robot dancing which allows us to use musical features for choreography generation. Optimal dance trajectories were computed using direct optimal control. From this formalization we derive three different methods of dance generation, that differ in the level of flexibility, human involvement, and automatization. The methods are: imitated, improvised, and automatic choreography generation. The imitated and improvised choreographic methods, are based on beat timing extraction. The automatic choreography generation method,uses the additional music features volume and vocal melody.The results are validated on 4 different music pieces in simulation using the dynamic simulator MuJoCo as well as in experiments on the real robot RH5 Manus. This work was published in the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2022, and was selected as a finalist for best entertainment and amusement paper.

The second one focuses on the ability of exploiting the full capabilities of a robot through motion generation, with the aim of achieving motions that are more human-like and that can lead to a certain trust and comfort feeling of the human towards the robot acting in its environment. To this purpose, we proposed a first study on resolving all the loop-closure

constraints of the series-parallel hybrid robot RH5 Manus within the trajectory optimization process. To this end, we use the open-source software Pinocchio developed by the INRIA Paris team with its recently introduced proximal formulation of the constrained dynamics. This approach allows us to converge to an optimal solution according to the least squares principle, even in the context of singularities. Among the optimization methods available in the literature, the differential dynamics programming (DDP) approach was used to generate optimal trajectories with respect to the constrained dynamics.

Results are presented in simulation as well as experiments on the real robot. This work is significant for humanoid robots based on electric actuation where one must seek to push the robot to its limits to achieve human like agility.

This work has been submitted to the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023 and is under review.

Publications :

1. Melya Boukheddimi, Daniel Harnack, Shivesh Kumar, Rohit Kumar, Shubham Vyas, Octavio Arriaga, Frank Kirchner, Robot Dance Generation with Music Based Trajectory Optimization, In IEEE/RSJ *International Conference on Intelligent Robots and Systems (IROS) 2022*, (IROS-2022), IEEE, Nov/2022.
2. Melya Boukheddimi , Rohit Kumar , Shivesh Kumar , Justin Carpentier , and Frank Kirchner, Investigations into Exploiting the Full Capabilities of a Series-Parallel Hybrid Humanoid using Whole Body Trajectory Optimization, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023, (IROS-2022), IEEE, Nov/2023. (submitted)

Links to Tangible results:

1. Melya Boukheddimi, Daniel Harnack, Shivesh Kumar, Rohit Kumar, Shubham Vyas, Octavio Arriaga, Frank Kirchner, Robot Dance Generation with Music Based Trajectory Optimization, In IEEE/RSJ *International Conference on Intelligent Robots and Systems (IROS) 2022*, (IROS-2022), IEEE, Nov/2022.
2. Melya Boukheddimi , Rohit Kumar , Shivesh Kumar , Justin Carpentier , and Frank Kirchner, Investigations into Exploiting the Full Capabilities of a Series-Parallel Hybrid Humanoid using Whole Body Trajectory Optimization, In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023, (IROS-2022), IEEE, Nov/2023. (submitted)

Videos :

https://www.youtube.com/watch?v=aN_v39p17tg
https://www.youtube.com/watch?v=MA42YUg3e8E

2022-04-WP3: Making for Everyone: Interactive, voice-based support for the design of 3D objects for digital fabrication

Start Date: May 2022
Duration: 4 Months
End Date: Aug 2022

Submitted by: Florian Müller, florian.mueller@um.ifi.lmu.de

Partners:

1. LMU Munich, Florian Müller, florian.mueller@um.ifi.lmu.de (3 PM)
2. University of Bari (External Partner), Giuseppe Desolda, giuseppe.desolda@uniba.it

Contributes to: WP3, T3.2

Results:

Manufacturing tools like 3D printers have become accessible to the wider society, making the promise of digital fabrication for everyone seemingly reachable. While the actual manufacturing process is largely automated today, users still require knowledge of complex design applications to not only produce ready-designed objects, but also adapt them to their needs or design new objects from scratch. To lower the barrier for the design and customization of personalized 3D models, we imagine an AI-powered system that assists users in creating 3D objects for digital fabrication. Reaching this vision requires a common understanding - a common ground - between the users and the AI system.

As a first step, in this micro project, we explored novices' mental models in voice-based 3D design by conducting a high-fidelity Wizard of Oz study with 22 participants without skills in 3D design. We asked the participants to perform 14 tasks revolving around some basic concepts of 3D design for digital modeling, like the creation of objects, the manipulation of objects (e.g., scaling, rotating, and/or moving objects), and the creation of composite objects. We performed a thematic analysis of the collected data assessing how the mental model of novices translates into voice-based 3D design.

We found that future AI assistants to support novice users in voice-based digital modeling must: manage the correction the users do during and after the commands to fix certain errors; deal with vague and incomplete commands by automatically completing the commands with sensible defaults or by asking the users for clarification; consider the prior novices knowledge, for example, about the use of undo e redo functions; provide only a simplified set of operations for creating simple and composite 3D objects; design a workflow similar to what novices would do if they were building real objects, for example, providing wizard procedures that guide novices in designing composite 3D models starting from the bottom; provide different commands to select 3D objects; understand and execute chained commands; understand commands that are relative to the users' point of view; grant multiple ways to refer to the axes, for example, by using their names, colors and user direction; favor explicit trigger words to avoid unintentional activation of the voice assistant; embrace diversity in naming approaches since novices often use other words to refer to 3D objects.

Publications:

Paper was not accepted at CHI 23. Currently under submission to INTERACT 23.

Links to Tangible results:

The transcribed and coded data we collected in our study, together with the codebook. We plan to make this data available to the community through a publication after the paper is published.

https://syncandshare.lrz.de/getlink/fiEFHiEQVQYtHDj5ZSWBdp/

2022-04-WP246: Polarization with the Friedkin-Johnsen model over a dynamic social network

Start Date: April 2022
Duration: 10 months (delay due maternity leave)
End Date: Dec 2022.

Submitted by: Elisabetta Biondi, elisabetta.biondi@iit.cnr.it

Partners:

1. Consiglio Nazionale delle Ricerche (CNR),Elisabetta Biondi, elisabetta.biondi@iit.cnr.it
2. Central European University (CEU), Janos Kertesz, kerteszj@ceu.edu, Gerardo Iniguez, IniguezG@ceu.edu

Contributes to: WP2-T2.5, WP4, WP6

Description:

The Friedkin-Johnsen model is a very popular model in opinion dynamics, validated on real groups, and well-investigated from the opinion polarization standpoint. Previous research has focused almost exclusively on static networks, where links between nodes do not evolve over time. In this micro-project, we want to fill this gap by designing a variant of the Friedkin-Johnsen model that embeds the dynamicity of social networks. Furthermore, we will design a novel definition of global polarization that combines network features and opinion distribution, to capture the existence of clustered opinions. We have analyzed the polarization effect of the new dynamic model, and identify the impact of the network structure.

Results:

Human social networks are very complex systems and their structure has an essential impact on opinion dynamics. However, since my main goal is to study the impact of the opinion dynamics model per se, we decided to deal with two different social network typologies: a Erdős–Rényi (ER) and a stochastic block model (SBM).

Design of the Friedkin-Johnsen (FJ) dynamic model.

We have implemented a rewiring policy that has been extensively studied in discrete opinion diffusion models. This involves substituting edges that connect nodes with different opinions with other edges. We have adapted this scheme to work with the FJ model's opinions, which are within the range of [-1,1], in both the asynchronous and synchronous versions. According to two parameters θ (the disagreement threshold) and p_rew (the rewiring probability):

- With probability 1-p_rew the FJ is applied
- With probability p_rew, if i and j disagree, i.e. $|x_i-x_j| > \theta$, the edge (i,j) is replaced with an edge (i,k) where k agrees with i, i.e. $|x_i-x_j| <= \theta$.

The above algorithm was specifically designed and implemented for the ER graph. However, in the case of the SBM, I have limited the potential candidates for rewiring to nodes within a maximum of two hops distance. This decision was made to prevent the block structure from becoming entirely irrelevant. The rationale behind this choice is based on the triadic closure mechanism, which suggests that individuals are more inclined to choose new acquaintances among the friends of their friends.

Design of the polarization metric.

The design of the polarization metric involved developing a definition for identifying highly polarized networks. We defined a highly polarized network as one in which there are two distinct opinions that are clustered into two tightly connected communities. To achieve this, we needed to consider both the network structure and the distribution of opinions. Therefore, we decided to use two different metrics to measure these aspects: bimodality for the opinion distribution and homogeneity for its correspondence with the network structure.

Bimodality.

The bimodality coefficient was used to measure the extent to which a distribution is bimodal. It is calculated using the skewness and kurtosis values and represents how similar the distribution is to one with two modes.

Homogeneity

To measure the homogeneity of the opinion distribution with the network structure, we examined the local distribution of nodes' opinions. We looked at whether each node's opinion was similar to those of its neighbors, which would suggest that it was in line with the overall opinion distribution over the network. The final homogeneity value was close to zero if the distribution of opinions was close to linear.

Experimental evaluation.

We have developed a Python simulator that can compute the dynamic FJ (rewiring included), and polarization metrics over time based on the given network and initial opinions. To test the model, we ran simulations on a small network comprising 20 nodes and compared the outcomes of the FJ with rewiring to those without rewiring. For the ER network, we used a vector of uniformly distributed opinions over [-1,1] as the initial opinions. However, for the SBM networks, we employed a different configuration, where the initial opinions were uniformly extracted over the intervals [-0.5,0-0.1] and [0.1,0.5], depending on whether the nodes belonged to one or the other block.

In conclusion, this Micro-Project involves the design of a dynamic version of the FJ model for synchronous and asynchronous cases. Additionally, we have developed a new definition of polarization that considers both the distribution of opinions and the network topology. To assess the model's effectiveness, we conducted simulations on two different network types: an ER network and an SBM network. Our findings indicate that the rewiring process has significant effects on polarization, but these effects are dependent on the initial network.

Links to Tangible results:
Github link of the code of the simulator for the new dynamic model:
https://github.com/elisabettabiondi/FJ_rewiring_basic.git

2022-04-WP35: Human-machine collaboration for content analysis in context of Ukranian war

Start Date: June 2022
Duration: 9 months
End Date: Feb 2023

Submitted by: Virginia Dignum, Umeå University (UMU)

Partners
1. Umeå University (UMU), Nina Khairova, nina.khairova@umu.se (6 PM)
2. Consiglio Nazionale delle Ricerche (CNR), Carmela Comito, carmela.comito@icar.cnr.it
3. Università di Bologna (UNIBO), Andrea Galassi, p.torroni@unibo.it

Contributes to: WP3, WP5, T5.6, T3.6

Description:

In this project, which we work with a Ukranian academic refugee, to combine methods for semantic text similarity with expert human knowledge in a participatory way to develop a

training corpus that includes news articles containing information on extremism and terrorism.

Results

1) Collection and curation of two event-based datasets of news about Russian-Ukrainian war.

The datasets support analysis of information alteration among news outlets (agency and media) with a particular focus on Russian, Ukrainian, Western (EU and USA), and international news sources, over the period from February to September 2022. We manually selected some critical events of the Russian-Ukrainian war. Then, for each event, we created a short list of language-specific keywords. The chosen languages for the keywords are Ukrainian, Russian, and English.

Finally, besides the scraping operation over the selected sources, we also gather articles using an external news intelligence platform, named Event Registry which keeps track of world events and analyzes media in real-time. Using this platform we were able to collect more articles from a larger number of news outlets and expand the dataset with two distinct article sets. The final version of the RUWA Dataset is thus composed of two distinct partitions.

2) Development of an unsupervised methodology to establish whether news from the various parties are similar enough to say they reflect each other or, instead, they are completely divergent and therefore one is likely not trustworthy. We focused on textual and semantic similarity (sentence embeddings methods such as Sentence-BERT), comparing the news and assess if they have a similar meaning. Another contribution of the proposed methodology is a comparative analysis of the different media sources in terms of sentiments and emotions, extracting subjective points of view as they are reported in texts, combining a variety of NLP-based AI techniques and sentence embeddings techniques. Finally, we applied NLP techniques to detect propaganda in news article, relying on self-supervised NLP systems such as RoBERTa and existing adequate propaganda datasets.

Preliminary Qualitative results:

When the events concern civilians all sources are very dissimilar. But Ukraine and Western are more similar. When the event is military targets, Russian and Ukraine sources are very dissimilar from other sources, there is more propaganda in Ukraine and Russian ones.

Contribution to the objectives of HumaneAI-net WPs:

The micro-project has been realized together with a refugee Ukranian academic, addressing, thus, WP5 goals by ensuring an AI system operating within a moral and social framework, in verifiable and justified ways. We focused on methods ensuring responsible design of AI Systems and compliance to ethical, trust, fairness, public perception and societal principles.

Links to Tangible results:

Github repository of datasets and software:
https://github.com/fablos/ruwa-dataset

2022-05-WP3: In Sync: Synchronization in interaction between Humans and Embodied AI Systems

Start Date: May 2022

Proposed Duration: 4 months
End Date: August 2022
Submitted by: Florian Müller, florian.mueller@um.ifi.lmu.de

Partners

1. LMU Munich, Florian Müller, florian.mueller@um.ifi.lmu.de Institution (4 PM)
2. University Warsaw, Andrzej Nowak, andrzejn232@gmail.com (4 PM)

Contributes to: WP3: T3.2

Description:

When we go for a walk with friends, we can observe an interesting effect: From step lengths to arm movements - our movements unconsciously align; they synchronize. Prior research in social psychology found that this synchronization is a crucial aspect of human relations that strengthens social cohesion and trust. In this micro project, we explored if and how this effect generalizes beyond human-human relationships. We hypothesized that synchronization can enhance the relationship between humans and AI systems by increasing the sense of connectedness in the formation of techno-social teams working together on a task.

Results

To evaluate the feasibility of this approach, we built a prototype of a simple non-humanoid robot as an embodied representation of an AI system. The robot tracks the upper body movements of people in its vicinity and can bend to follow human movements and vary the movement synchronization patterns. Using this prototype, we conducted a controlled experiment with 51 participants exploring our concept in a between-subjects design. We found significantly higher ratings on trust between people and automation in an established questionnaire for synchronized movements. However, we could not find an influence on the willingness to spend money in a trust game inspired by behavioral economics. Taken together, our results strongly suggest a positive effect of synchronized movement on the participants' feeling of trust toward embodied AI representations.

List of pubications,

1. Wieslaw Bartkowski, Andrzej Nowak, Filip Ignacy Czajkowski, Albrecht Schmidt, and Florian Müller. 2023. In Sync: Exploring Synchronization to Increase Trust Between Humans and Non-humanoid Robots. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3544548.3581193

Links to Tangible results:

Paper:

https://syncandshare.lrz.de/getlink/fiAUnSaqgQJyEdcw5XJqpN/in_sync_final.pdf

Video:

- **Short**: https://syncandshare.lrz.de/getlink/fiRjwbk1AoYxKujaEaZ5ax/in_sync_video_short.mp4
- **Full**: https://syncandshare.lrz.de/getlink/fiGEX3bGahhbzrChUiXqvL/in_sync_video_full.mp4
- **Repository**: https://github.com/wbartkowski/In-Sync-Robot-Prototype

2022-06-WP1: AI-Assistant to Mitigate Confirmation Bias in Cooperative Bayesian Optimization

Start Date: 1 June 2022

Duration: 4 Months
Completion Date 30 September 2022

Proposed by: Samuel Kaski, Samuel.kaski@aalto.fi, Aalto University

Partners:

1. Aalto University, Samuel Kaski, samuel.kaski@aalto.fi
2. Delft University of Technology (TU Delft), Frans Oliehoek. F.A.Oliehoek@tudelft.nl

Results:

In human-AI cooperation, autonomous agents work interdependently to achieve a common goal. Agents cooperate with their limited information processing capabilities, and various factors such as agent biases, heuristics, and imperfect knowledge can hinder cooperation. To achieve effective human-AI cooperation, AI agents should account for the human's information processing capabilities and biases and be adaptable to changing needs and preferences. We are interested in how the AI agent can complement the capabilities of the human user and assist them when making a joint decision. This micro-project contributes to developing methodologies that allow humans to be interactively involved "in the loop" (WP1).

Here, the loop is a cooperative Bayesian optimization game where the goal is to optimize a 2D black-box function. At each iteration, the AI chooses the first coordinate, and then the user observes and opts for the second. Finally, the function is queried, and the result is shown to both parties. The researcher can control agents' characteristics, making it suitable for studying confirmation bias and imperfect knowledge. The project investigates how a planning AI agent can alleviate BO regret due to the human agent's biases and imperfect information allocation. The aim is to build a planning AI agent to aid the user in the optimization task, where no single party has full decision-making power.

In this mini-project, we conducted an experiment with a synthetic user for various scenarios. We assumed the user decision process comprises two hierarchical steps: updating the belief and taking action based on the belief. In the user model, these steps are regulated by model parameters, determining the conservatism in belief updates and exploration in opting for actions. Regarding that AI-assistant has a well-specified user model, we formulate the assistance decision problem as an MDP with unknown parameters ? and ?. We adjusted the Bayes-adaptive Monte Carlo planning methods to our problem to find the best policy for AI.

The reward in our planning AI is a weighted sum of two parts. Intuitively, one is responsible for ensuring that the AI favours actions for which the user can choose a promising second coordinate, and the other is responsible for reducing the risk that the user will act suboptimally, especially when the user model fails to predict the user well.

We compared our planning AI's performance to a greedy AI (GreedyAI), which only tries to optimize the function based on its updated knowledge without considering the user. We also considered random query optimization (RandOpt) and a BO algorithm with the same acquisition function as GreedyAI uses but with full access to the data as logical lower and upper bounds for the optimization performance, respectively. The results demonstrate the planning AI can assist the user in optimizing the function significantly better (measured as BO regret) than the GreedyAI and RandOpt, and relatively close to the logical upper bound under some conditions, even with its imperfect information.

Interestingly, a well-designed reward makes the cooperation effective even when the user follows a relatively high explorative policy. Investigating in-depth reveals that the planning

AI lets the user explore the function adequately and reduces the chance of getting stuck at local optima.

In summary, the findings indicate that integrating a user model into a planning AI agent can mitigate potential biases of the user, enabling the team to avoid local optima and achieve better planning outcomes in sequential optimization games. By anticipating the user's future behaviour, the agent can better guide the user towards optimal query.

Publications:

Cooperative Bayesian Optimization for Imperfect Agents, Khoshvishkaie A., Mikkola P., ... and Kaski S., submitted to ECML-2023

Links to Tangible results:

The code will be publicly available upon acceptance to the conference mentioned above.

2022-7-WP3 A Scoping Review on Artificial Intelligence and Extended Reality

Start Date: July 2022
Duration: 9 months
End Date: Mar 2023

Submitted by: Teresa Hirzle, tehi@di.ku.dk

Partners
1. UCPH, Teresa Hirzle, kash@di.ku.dk
2. LMU, Florian Müller, albrecht.schmidt@ifi.lmu.de
3. Saarland University (External partner), Martin Schmitz
4. Universität Innsbruck (External partner), Pascal Knierim

Contributes to: WP3

Results:

We conducted a scoping review covering 311 papers published between 2017 and 2021. First, we screened 2619 publications from 203 venues to cover the broad spectrum of XR and AI research. For the search, we I inductively built a set of XR and AI terms. The venues include research from XR, AI, Human-Computer Interaction, Computer Graphics, Computer Vision, and others. After a two-phase screening process, we reviewed and extracted data from 311 full papers based on a code book with 26 codes about the research direction, contribution, and topics of the papers, as well as the algorithms, tools, datasets, models, and data types the researchers used to address research questions on XR and AI. The extracted data for these codes form the basis for our predominantly narrative synthesis. As a result, we found fve main topics at the intersection of XR and AI: (1) Using AI to create XR worlds (28.6%), (2) Using AI to understand users (19.3%), (3) Using AI to support interaction (15.4%), (4) Investigating interaction with intelligent virtual agents (IVAs) (8.0%), and (5) Using XR to Support AI Research (2.3%). The remaining 23.8% of the papers apply XR and AI to an external problem, such as for medical training applications (3.5%) or for simulation purposes (3.0%). Finally, we summarise our findings in 13 research opportunities and present ideas and recommendations for how to address them in future work. Some of the most pressing issues are a lack of generative use of AI to create worlds, understand users, and enhance interaction, a lack of generalisability and robustness, and a lack of discussion about ethical and societal implications.

In terms of the call topics, we analysed whether XR can serve as a tool to establish and enhance interactive grounding in human-AI interaction. Here, we found that there is a lack of understanding user experience during human-AI interaction using XR technology. Typically, AI is used for content creation and to enhance interaction techniques. We did, however, not find a lot of papers that use XR to support human-AI interaction. There are some works that look into artificial agents and how an interaction with them can be realised through XR. However, most of these works do not yet work in real=time and are mostly based on mock-up scenes.

List of publications
1. Teresa Hirzle, Florian Müller, Fiona Draxler, Martin Schmitz, Pascal Knierim, and Kasper Hornbæk. 2023. When XR and AI Meet - A Scoping Review on Extended Reality and Artifcial Intelligence. *In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (CHI '23), April 23–28, 2023, Hamburg, Germany. ACM, New York, NY, USA, 45 pages. https://doi.org/10.1145/3544548.3581072

Links to Tangible results:

Paper: https://thirzle.com/pdf/chi23_xrai_scoping_review_hirzle.pdf
Reviewed Papers and Coding Spreadsheet:
https://thirzle.com/supplement/chi23_xrai_scoping_review_hirzle.zip

2022-09-WP12: Graybox methods for augmenting human-driven narrative analyses

Start Date: 2022-09-01
Duration: 4 months
End Date: 2023-01-31

Submitted by: Szymon Talaga, stalaga@uw.edu.pl

Partners:

1. Univ. Warsaw, Szymon Talaga, stalaga@uw.edu.pl (11 PM)
2. Institut Polytechnique de Grenoble, James Crowley, james.crowley@univ-grenoble-alpes.fr

Contributes to: WP1: T1.3, WP2

Description:

This Micro-Project has laid the groundwork for developing a new approach to narrative analysis providing a gray-box (at least partially explainable) NLP model tailored for facilitating work of qualitative text/narrative analysts. We conducted a proof-of-concept study combining existing standard NLP methods (e.g. topic modeling, entity recognition) with qualitative analysis of narratives about smart cities and related technologies and use this experience to conceptualize our approach to narrative analysis, in particular with respect to problems which are not easily solved with the existing tools. This current initial research micro-project will be followed by a Micro-Project dedicated to formalizing our approach to narrative analysis and developing its open-source implementation (for Python)

Results

The aim of the project was to develop a software package (for Python) providing easy to use and understand (also for researchers not trained in computer science or linguistics) tools for extracting narrative information (active and passive actors, the actions they perform as well as descriptions of both actors and actions, which together define events) and organizing them in rich hierarchical data structures (data model is implicitly graphical) from which subsequently different sorts of descriptive statistics can be generated depending on particular research questions. Crucially, for this to be practically possible, a legible and efficient framework for querying the produced data is needed.

The above goal fits into a broader HumanE-AI objective of developing common ground concepts providing better representations shared by humans and machines alike. In particular, the contribution of the project to work on aligning machine analyses with human perspective through the notion of narratives is twofold. Firstly, narrative-oriented tools for automated text analyses can empower human analysts as, arguably, the narrative framework provides a more natural and meaningful context for people without formal training in linguistics and/or computer science for reasoning about textual data. Secondly, the development of the software for narrative analysis is naturally intertwined with conceptual work on the core terms and building blocks of narratives, which can inform subsequent work on more advanced approaches.

Importantly, the software is developed as a graybox model, in which core low-level NLP tasks, such as POS and dependency tagging, are performed by a blackbox statistical model, and then they are transformed to higher order grammar and narrative data based on a set of transparent deterministic rules. This is to ensure high explainability of the approach, which is crucial for systems in which the machine part is supposed to be a helper of a human analyst instead of an implicit leader.

Currently, the core modules of the package responsible for the grammatical analysis are mostly ready (but several improvements are still planned). This includes also a co-reference resolution module. Moreover, the core part of the semantic module, which translates grammatical information to more semantic constructs focused on actors, actions and descriptions, is also ready. What is still missing are an interface exposing methods for end users allowing easy access and analysis of rich data produced by the package as well as a principled and convenient query framework on which the interface should be based. This is the main focus of the ongoing and future work. The second missing part is the documentation, but this part is best finished after the interface is ready.

Thus, even though the package in the current state can seem a little rough from the perspective of an end user, its quality and usefulness will increase steadily as new updates are delivered.

Links to Tangible results: https://github.com/sztal/segram

2022-09-WP2: Multi-Relational Contextual Reasoning for Complex Scene Generation for Autonomous Vehicle Data

Start Date: 1 September 2022
Duration: 7 Months
Completion Date 31 March 2023

Submitted by: Loris Bozzato, bozzato@fbk.eu

Partners:
1. Fondazione Bruno Kessler, Loris Bozzato, ghidini@fbk.eu (3 PM)
2. Technical University of Vienna, Thomas Eiter, eiter@kr.tuwien.ac.at (3 PM)
3. BOSCH Deutschland (External Partner), Stepanova Daria

Contributes to: WP1: T1.1, T1.5, WP2: T2.2

Results:

Our work considers the applicability of the MR-CKR framework to the task of generating challenging inputs for a machine learning model. Here, MR-CKR is a symbolic reasoning framework for Multi-Relational Contextual Knowledge Repositories that we previously developed. Contextual means that we can (defeasibly) derive different conclusions in different contexts given the same data. This means that conclusions in one context can be invalidated in a more specific context. Multi-Relational means that a context can be "more specific" with respect to different independent aspects, such as regionality or time.

The general idea of generating challenging inputs for a machine learning model is the following: We have limited data on which we can train our model, thus, it is like that the model does not cover all eventualities or does not have enough data in specific contexts to lead to the correct result. Obtaining more data is often very difficult or even infeasible however. We introduce a new approach to solving this problem. Namely, given a set of diagnoses describing contexts in which the model performs poorly we generate new inputs that are (i) in the described contexts and (ii) as similar as possible to a given starting input.

(i) allows us to train the network in a targeted manner by feeding it exactly those cases that it struggles with. (ii) ensures that the new input only differs from the old one in those aspects that make the new input problematic for the model. Thus, allowing us to teach the model to recognize aspects relevant for the answer.

This fits very well with the capabilities of MR-CKR: on the one hand, we have different contexts in which the inputs need to be modified to suit a different diagnosis of failure of the model. On the other hand, we can exploit the different relations by having one relation that specifies that inputs are more modifiable in one context than another and another relation that describes whether one diagnosis is a special case of another. Additionally, it allows us to incorporate global knowledge such that we can only modify inputs in such a manner that the result is still "realistic", i.e., satisfies the axioms in the global knowledge.

In this work, we provide a prototype specialized to generating similar and problematic scenes in the domain of Autonomous Driving. This work fits well into Task 1.1 of WP1: "Linking symbolic and subsymbolic learning", since we use a symbolic approach to enable the use of domain knowledge in order to advance the performance of a subsymbolic model. Furthermore, it also loosely fits into Task 1.5 of WP1: "Quantifying model uncertainty", since we can quantify how similar the generated new inputs are to the original ones.

2022-11-WP3 Use of dialog context to boost ASR/NLG/TTS and improve the overall quality of voice dialog systems

Start Date: 1 Nov 2022

Proposed Duration: 4 Months

End Date: 28 Feb 2023

Submitted by: Brno University of Technology, Petr Schwarz, schwarzp@fit.vutbr.cz

Partners

1. Brno University of Technology, Petr Schwarz, schwarzp@fit.vutbr.cz
2. Charles University, Ondrej Dusek, odusek@ufal.mff.cuni.cz

Contributes to WP3

Description:

This project brings us data, tools, and baselines that enable us to study and improve context exchange among component and dialog sides (AI agent and human) in voice dialog systems. A better context exchange allows us to build more accurate automatic speech transcription, better dialog flow modeling, more fluent speech synthesis, and more powerful AI agents. The context exchange can be seen as an interactive grounding in two senses - among dialog sides (for example, technologies like example automatic speech transcription rarely use the other dialog side information to adapt itself) and among dialog system components (the speech synthesis rarely uses dialog context to produce more fluent or expressive speech). The individual project outputs are summarized below

Results

1) Audio data collection software based on the Twilio platform and WebRTC desktop/mobile device clients. The purpose is to collect audio data of communication between agents (company, service provider, for example, travel info provider) and users. This software enables us to collect very realistic voice dialogs that have high-quality audio

(>= 16kHz sampling frequency) on the agent side and low telephone-quality audio on the user side. The code is available here: https://github.com/oplatek/speechwoz

2) We have established a relationship with Paweł Budzianowski (Poly.AI) and Izhak Shafran (Google). Paweł created the MultiWoz database – an excellent dialog corpus (https://arxiv.org/abs/1810.00278) that we use for the text-based experiment. We decided to collect our audio data similarly. Izhak organized DSTC11 Speech Aware Dialog System Technology Challenge (https://arxiv.org/abs/2212.08704) and created artificial audio data for MultiWOZ through speech synthesis, reading, and paraphrasing. Both provided us with the necessary advice for our data collection.

3) Speech dialog data – the data collection platform preparation and data collection are very time-consuming. The data collection is in progress and will be released before June 26th, 2023.

4) Initial experiments with context exchange between dialog sides (user and agent) were performed. These experiments show a nice improvement in the component of automatic speech recognition side. The results will be re-run with the collected data and published when the collection is finished.

5) Initial experiments with training instance weighting for response generation – which brings context to dialog system response generation, were performed. Experiments were based on the AuGPT system, previously developed at CUNI. The code is available here: https://github.com/knalin55/augpt. Instance weighting increases the re-use of context, compared to normal training, and can go even beyond natural occurrences in data. Simple weighting (threshold) seems better than designing a complex instance weight (in terms of automated metrics, limited manual evaluation is not conclusive). Cross entropy loss works better than unlikelihood loss, where dialogue success may be reduced.

6) There is ongoing work on building a team for JSALT research summer workshop.

https://www.clsp.jhu.edu/2023-jelinek-summer-workshop,

https://jsalt2023.univ-lemans.fr/en/index.html.

This is a prestigious workshop organized by John Hopkins University every year. This year it is supported and co-organized by the University of Le Mans. Our topic is the Automatic design of conversational models from observation of human-to-human conversation (https://jsalt2023.univ-lemans.fr/en/automatic-design-of-conversational-models-from-observation-of-human-to-human-conversation.html). The topic passed a scientific review by more than 40 world-class researchers in AI in Baltimore, USA, in December 2022, and was selected for this workshop out of 15 proposals together with three others. The workshop topic builds on the outcome of this Micro-Project and will reuse the collected data.

Links to Tangible results:

- (WIP) Code for audio data collection: https://github.com/oplatek/speechwoz
- (WIP) Code for end-to-end response generation: https://github.com/knalin55/augpt
- (WIP) Report for end-to-end response generation: https://docs.google.com/document/d/1iQB1YWr3wMO8aEB08BUYBqiLh0KreYjyO4EHnb395Bo/edit
- Workshop links: https://www.clsp.jhu.edu/2023-jelinek-summer-workshop, https://jsalt2023.univ-lemans.fr/en/index.html
- Workshop proposal: https://docs.google.com/document/d/19PAOkquQY6wnPx_wUXIx2EaInYchoCRn/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true

- Workshop team (not finalized yet): https://docs.google.com/spreadsheets/d/1EsHZ-_OREkvf8ODiN7759OYHqSb6MBAX/edit?usp=sharing&ouid=105764332572733066001&rtpof=true&sd=true

(WIP) = Work in Progress

On Going Micro Projects at the end of period 2.

Micro-projects completed after Feb 2023.

2022-09-WP235: Multilingual and Multimodal conversational agent combined with search engine models.

Start Date 2022-09-01
Proposed Duration: 6
End Date 2023-04-28 (on-going)

Submitted by: Eric Blaudez, eric.blaudez@thalesgroup.com

Partners:
1. Thales Eric Blaudez, eric.blaudez@thalesgroup.com
2. Unibo Paolo Torrini, p.torroni@unibo.it
3. LISN Christophe Servan c.servan@qwant.com

Contributes to: WP2, WP3, WP5

Description:

The micro-project provides a demonstration of the hierarchical Framework for collaboration decribed in the Humane-AI Net revised strategic workplan, by constructing amultimodal and multilingual conversational agents focused on search.

The framework is based on hierarchical levels of abilities:

- Reactive(sensori-motor) Interaction: Interaction is tightly-coupled perception-action where actions of one agent are immediately sensed and interpreted as actions of the other. Examples include greetings, polite conversation and emotional mirroring
- Situated (Spatio-temporal) Interaction Interactions are mediated by a shared model of objects and relations (states) and shared models for roles and interaction protocols.

On this Micro-Project we focused on the 2 first levels (Reactive and Situational) and designed the global framework architecture. The results are to be demonstrated in a Proof of Concept (PoC).

Results:

Components development:

The development of the components is on going. The mini project deals with three modules: the Human Machine Interface (HMI) to propose an interaction interface, the reactive module to recognize emotions and the situational module to manage the dialog situation. The framework is developed with the strategy proposed by [1]; in this project, we focus on reactive and situational part.

Human Machine interface [Status: DONE]

The HMI is a simple Flask application connected to Framework REST Service.

Reactive module [Status: ON GOING, framework integration almost finished]

The objective of the reactive module was to recognize emotions in conversation, so as to enable the other modules to drive the conversation accordingly. We thus addressed an Emotion Recognition in Conversation (ERC) task. The first step was a literature review and a survey of the available datasets and architectures. We chose the MELD [3] dataset as it is considered a benchmark for ERC and one of the few datasets available for building our ERC module. We considered different options for capturing the context, including the current utterance and its corresponding dialogue history were analyzed. As an architecture, we chose EmoBERTa [1], due to its strong performance and availability of implementation. We trained EmoBERTa using MELD splits for training, validation and test dataset. We carried out several experiments to establish baselines and examine the impact of different context representations. Moreover, since MELD conversations differ from those that may arise in the application domain envisaged for our chatbot, since we do not have any validation datasets available, we explored the generalization capability of the model by transfer learning and few-shot learning. In particular, we considered SetFit [4] as a few-shot learning technique and DailyDialog [2] for transfer learning validation. The experiments conducted on EmoBERTa have yielded promising results, indicating that this architecture is a suitable starting point for constructing the reactive module. However, the task of generalization and adaptability of the model is challenging, and ongoing experiments are being carried out to address these issues. References

Situational Module

This module is responsible of the “Interactions mediated by shared models of objects (entities) and relations and shared models of roles and interaction protocols”. Knowledge graph are used to represent the situation, the relations and keep the context of the dialog.

- The document Analysis module has been developed in AI4EU project and is adapted and used as submodule for the miniproject. The module provides a data processing and a knowledge extraction functions by using Natural Language Processing. It builds a structured representation of the data (knowledge graph, named entities, keywords, summaries …) and provide semantic search. This module is in constant evolution; it will provide at the end of year a capability to manage more than 100 languages.
- The LISN part aims to create a spoken language understanding (SLU) part and the dialogue manager part. LISN proposed a first proof of concept based Speech recognition on RASA [1] in order to propose a full dialogue framework. For this part, at LISN, we focus on the SLU part of the module. The SLU module is based on the DIET-Classifier model (DIET stands for Dual Intent and Entity Transformer) [2] for both performing intent detection and concept detection. For the next part, we will move to another framework since we observed some limitation in the use of the RASA framework. We will also use another model to perform both intent and concept detection, by using the JointBERT model [3]. Preliminary results are encouraging.
- Thales & LISN works on Situational context representation based on (Temporal)-Knowledge graph and is based on Endsley model [4].

Bibliography:

[1] Crowley, James & Coutaz, Joëlle & Grosinger, Jasmin & Vázquez-Salceda, Javier & Angulo, Cecilio & Sanfeliu, Alberto & Iocchi, Luca & Cohn, Anthony. *IEEE Pervasive*, Vol 22, No. 1, Mar 2023.

Reactive Module (UniBO)

[1] Kim, T., Vossen, P.: Emoberta: Speaker-aware emotion recognition in conversation with roberta. CoRR abs/2108.12009 (2021)

[2] Li, Y., Su, H., Shen, X., Li, W., Cao, Z., Niu, S.: Dailydialog: A manually labelled multi-turn dialogue dataset. In: IJCNLP(1). pp. 986–995. Asian Federation of Natural Language Processing (2017)
[3] Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E., Mihalcea, R.: MELD: A multimodal multi-party dataset for emotion recognition in conversations. In: ACL (1). pp. 527–536. Association for Computational Linguistics (2019)
[4] Tunstall, L., Reimers, N., Jo, U.E.S., Bates, L., Korat, D., Wasserblat, M., Pereg, O.: Efficient few-shot learning without prompts. CoRR abs/2209.11055 (2022)

Situational Module (Thales & LISN)
[1] Bocklisch, T., Faulkner, J., Pawlowski, N., and Nichol, A. (2017). Rasa: Open source language understanding and dialogue man-agement. In the proceedings of the first NIPS workshop on Conversational AI.
[2] Bunk, T., Varshneya, D., Vlasov, V., and Nichol, A. (2020). Diet: Lightweight language understanding for dialogue systems. arXiv preprint arXiv:2004.09936.
[3] Chen, Q., Zhuo, Z., and Wang, W. (2019). Bert for joint intent classification and slot filling. arXiv preprint arXiv:1902.10909.
[4] M. R. Endsley and D. J. Garland (Eds.). Situation Awareness Analysis and Measurement. Mahwah, NJ: Lawrence Erlbaum, 2000.

Links to Tangible results:
Jira & Confluence: https://humane-dialog.atlassian.net/jira/software/projects/HAD/pages
Github [PRIVATE]: project: https://gitlab.com/humane-ai-chatbot/chatbot-fmk
submodules:
T-KEIR: https://github.com/ThalesGroup/t-keir.git
erc-unibo-module: https://github.com/helemanc/erc-unibo-module

2023-01-WP3: Storyboarder: combining Play Script and Image Generation

Start Date: Feb 2023
Proposed Duration: 6
End Date: July 2023

Submitted by: Victor Chitolina Schetinger, victor.schetinger@tuwien.ac.at

Partners:
1. Charles University, Rudolf Rosa, rosa@ufal.mff.cuni.cz
2. IADE – Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia
3. (External Partner), Edirlei Lima, edirlei.lima@universidadeeuropeia.pt

Contributes to: WP3

Results:

In the short time since our start we were able to achieve the results we proposed, namely the generation of storyboards through the use of text and image generative models combined. Due to the fast development of these fields in the past months, however, the quality of the results is not up to the state of the art. We plan to use the remainder time of the project to improve them, explore adjacent research directions and work on a publication plan.

Links to Tangible results:

https://ufallab.ms.mff.cuni.cz/cgi-bin/rosa/theaitre-demo/demo_images.py