Summary

During the workshop, eight topics were suggested and discussed, some of them directly related to each other, like *cognitive offload* and *shared workload*, while *implementation methods for industry* could be seen as rather orthogonal to most of the other topics.

This document is an attempt to extract and summarise the main points of the discussion, that would lead into central research questions and challenges to be integrated into a form of roadmap for **Human AI Teaming** research in Sweden.

System features and abilities: Trust, reliability, failures

The high level question *How can we deal with silent failures?* was marked the most as one that should be taken into consideration. Together with topics in the area of *Trust and Transparency*, which also received quite many importance markers, a line of research in Human AI Teaming could thus be (should be?) the establishment of reliable, transparent and trustworthy mechanisms for teamwork between humans and AI systems. In both areas, sub topics touched upon the problem of "understanding what went wrong and why", hence, the above mentioned mechanisms would have to rely on the ability of AI systems to perform introspection and to handle the possibility of a situation offering several ways of interpretation. At the same time, humans teaming up with AI systems must have the opportunity to observe and interpret the actions of these AI systems, which is where transparency comes into play.

Teaming paradigms: Offloading (work load, cognitive load), team structure, emergence

The four topics of *Cognitive Offload*, *Shared Workload*, *Hybrid Cognitive Systems*, and *Emergence* can all be seen as the basis for the discussion of teaming paradigms. Questions that arise deal with the distribution of authority (Is the AI system a team mate or a tool to the human?), with the distribution of actual work load in terms of Team Resource Management (Is the task something physical or cognitive, and how can it best be handled? Should it be handled in a distributed way or in direct collaboration?), and with the theoretically grounded formation behind the team structure (Should the team emerge around its task or is it predefined? Who are the team members, are they determined up-front or "found on the fly"?).

Orthogonal topic 1: Individualised AI

A topic that is relevant across the others is the question of *Individualised AI* (even specialised regarding tasks or areas of application). Several of the sub topics for this overall question form questions better summarised under trust and transparency (see above), but there is then the open question of whom or what an AI should or could be individualised for.

Orthogonal topic 2: Implementation methods in industry

Another orthogonal question is that of *Implementation methods for the industry*, as there is still a need to establish use cases and scenarios that are applicable and relevant for the industry, where AI

systems could be assessed regarding a "Human readiness level" rather than the otherwise common TRL scale.

From:
https://hait.cs.lth.se/ - **Human AI Teaming**

Permanent link:
https://hait.cs.lth.se/topics:roadmap_stubb

Last update: **2022-12-23 13:29**