Hybrid Teams of Humans, Robots and Virtual Agents in a Production Setting

Tim Schwartz*, Michael Feld*, Christian Bürckert*, Svilen Dimitrov‡, Joachim Folz‡,
Dieter Hutter†, Peter Hevesi‡, Bernd Kiefer*, Hans-Ulrich Krieger*, Christoph Lüth†, Dennis Mronga†,
Gerald Pirkl‡, Thomas Röfer†, Torsten Spieldenner*, Malte Wirkus†, Ingo Zinnikus* and Sirko Straube†
German Research Center for Artificial Intelligence, DFKI GmbH
Email: firstname.lastname@dfki.de
*Stuhlsatzenhausweg 3, 66123 Saarbrücken, Germany
†Robert-Hooke-Strasse 1, 28359 Bremen, Germany
‡Trippstadter Strasse 122, 67663 Kaiserslautern, Germany

***Abstract*—This video paper describes the practical outcome of the first milestone of a project aiming at setting up a so-called Hybrid Team that can accomplish a wide variety of different tasks. In general, the aim is to realize and examine the collaboration of augmented humans with autonomous robots, virtual characters and SoftBots (purely software based agents) working together in a Hybrid Team to accomplish common tasks. The accompanying video shows a customized packaging scenario and can be downloaded from http://hysociatea.dfki.de/?p=441.**

I. Introduction

In the context of Industrie 4.0 [1], complex and flexible production requirements of the future may be met by introducing Hybrid Teams into the process that can react on unplanned events by spontaneous reorganization. The accompanying video[1] shows a first setup of a Hybrid Team as part of the project *Hybrid Social Teams for Long-Term Collaboration in Cyber-physical Environments* (HySociaTea) [2]. Besides research on the technical feasibility, another key aspect of the project is the development of (robotic) team-competencies as well as intelligent multi-agent behavior, both of which are also important aspects in purely human teams. The technical systems developed in HySociaTea are mainly meant to be used as assistance systems for humans working in production plants; the robots should therefore be perceived as partners in the overall working process. On the long run, the team organization, as developed and examined in HySociaTea, can be used in different real-world scenarios.

In order to realize this vision, eight different research departments of the three major sites of the German Research Center for Artificial Intelligence (DFKI) are collaborating in this project: The Robotics Innovation Center and Cyber-physical Systems group in Bremen, the Intelligent User Interfaces, Agents & Simulated Reality, and Multilingual Technologies groups in Saarbrücken, and the Augmented Vision, Embedded Intelligence and Knowledge Management groups in Kaiserslautern.

This paper reports the outcome of the first year, in which the basic infrastructure was developed and implemented, and a first demonstration scenario was realized. In the remainder of the paper, we describe the members of the Hybrid Team, the underlying architecture and explain the core elements of the video.

[1]http://hysociatea.dfki.de/?p=441

II. Scenario and Goals

The goal of the team is to accomplish a given manufacturing task by organizing itself according to the individual skills of each team member, respectively. Humans play the central role in the team, since they have the highest cognitive abilities and can react extremely flexible to new situations. Robots typically take over physically demanding tasks and the virtual agents provide a straightforward interface to digitally available information.

The packaging scenario in the video resembles a highly flexible production setting with small batch size, in which the creativity of the worker plays an important role. In the scenario, the customer brings an arbitrary item to the workshop in order to get it packaged in a very customized fashion. This could be, for example, a precious vase that is to be shipped overseas or a technical prototype of a new device that is to be presented to a potential investor.

In such a scenario, the human will be the only team member to a) understand the customer's request and b) to come up with a rough plan on how to accomplish the request and c) to deal with all the problems that might arise during the endeavor. The remaining team members should assist and, if possible, they should do so in a proactive manner, i.e. without being directly instructed.

A. The Setting

A workshop-like environment was set up consisting of a customer area, a storage area and a production area. The customer area consists of a reception desk, which is also the workplace of the technician/manager, who will receive the customer's order. The storage area currently consists of two shelves, containing various materials and tools. Figure 1 shows a sketch of the workshop, marking the positions of the assets (green boxes) and standard position of the team members (red boxes).

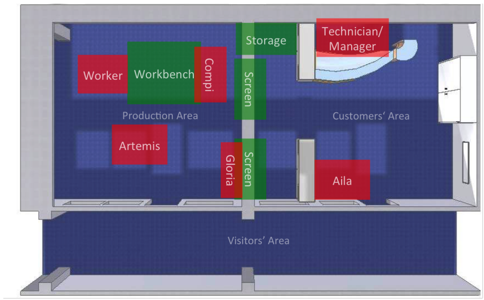

Fig. 1. Plan of the workshop-like environment sketching the different areas, the positions of the workbench, shelves and screens (green boxes), as well as standard positions of the team members (red boxes).

III. Members of the Hybrid Team

The Hybrid Team in the video consists of two humans, three robots, and one virtual character (see Figure 2).

A. Robots

All robots in the video were developed at the Robotics Innovation Center of DFKI in previous or ongoing projects. Software-wise, these robots run on the ROCK framework (RObot Constuction Kit) [3], which is freely available under LGPL[2].

Aila [4] is a humanoid, mobile dual-arm robot, which was originally developed to investigate aspects of mobile manipulation. In the current scenario, Aila serves as a robotic interface to customers visiting the workshop.

Artemis [5] is a space-bot rover that was originally built to autonomously collect objects in an unknown environment in planetary space exploration. In the Hybrid Team Artemis acts as a logistics robot, transporting tools, building-material or objects from and to different locations.

The robot *Compi* [6] is the only stationary robot in the team. Compi can switch between various stages of flexibility or stiffness and acts as a “helping hand” for a worker, e.g. holding objects, while they are being worked on.

B. Virtual Character

Gloria is a virtual character, which is realized using a commercial virtual character SDK called CharActor, provided by Charamel[3]. The CharActor SDK already includes the complete rendering engine, a text-to-speech (TTS) engine including lip-synchronization, and a large library of facial expressions and motions. In the video, Gloria is the worker’s main interface to the system. Gloria “lives” on one of the large screens, which are installed in the workshop. Apart from Gloria and Compi, all other team members can freely move inside the facility.

C. SoftBots

In the context of HySociaTea, SoftBots are purely software based modules without physical or graphical embodiment. These SoftBots typically aggregate data produced by other team members (e.g. raw sensor data, speech acts) and in turn update databases or provide meaningful, refined data. There are several SoftBots in the system, e.g. to keep track of the current location of objects, tools and materials, or to convert numeric position-coordinates into semantic descriptions.

As an example for a more complex SoftBot, we implemented a module that collects information about the worker’s requests for building-material, and that automatically learns which materials are often used together. Gloria uses the information provided by this SoftBot, to pro-actively ask the worker if they also need the additional material.

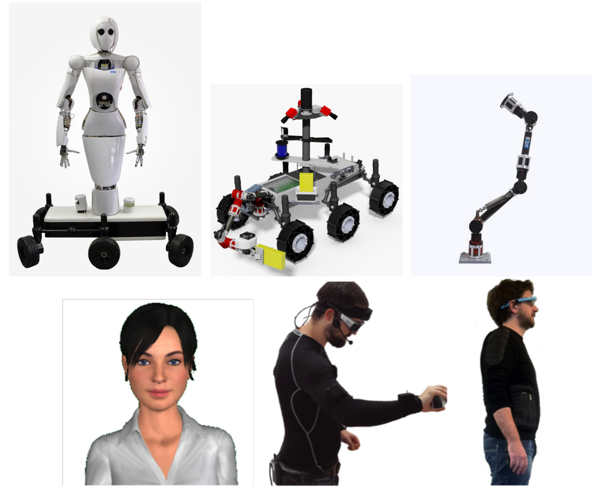

Fig. 2. Members of the Hybrid Team (from top left to bottom right): Aila, the humanoid service robot, Artemis, the logistics robot, Compi the ”helping hand“, Gloria, the virtual character, a worker, wearing a sensor jacket and eye-tracking glasses, and a technician, wearing Google Glass.

D. Augmented Humans

The team in the video includes two humans, a worker and a technician/manager. In order to allow the human team member to interact with the system, they can either use parts of the instrumented environment or various wearables. The worker, who is the main protagonist in the current setup, is heavily instrumented. He wears a sensor jacket [7] that allows to sense the orientation of his head and upper torso, as well as the motion and position of his upper and lower arms. In order to gain information about his position within the room, we developed a positioning system based on magnetic coils [8], [9], which are distributed in the environment. The worker wears an additional sensor to measure the signals of these coils. In addition to that, the worker also wears SMI eye-tracking glasses[4], and a microphone headset for speech interaction. The information of all these wearable sensors is combined in a fusion module that is thus capable to produce data about the human that is similar to what the robotic team members can send about themselves. In contrast to the worker, the technician/manager in the video is minimally instrumented and only wears a Google Glass.

[2] http://www.rock-robotics.org

[3] http://www.charamel.com

[4] http://eyetracking-glasses.com

IV. Behind the Scenes: Communication

A central question when setting up a Hybrid Team is the realization of suitable interfaces. Humans usually use speech, gestures and facial expressions to transfer information. Artificial agents however can use direct data streams to communicate with the system and other artificial team members. On top of that, the Hybrid Team should, at least partially, organize itself and should not be orchestrated by a centralized controlling instance.

To deal with these different levels and requirements, four core modules are realizing the communication within the Hybrid Team that are described in the following: a blackboard for task management, a communication middleware, a dialog planner, and a dual reality module. It is important to stress again that all team members in the video are autonomous, i.e. tasks are not assigned from the system, but chosen by the team member according to individual skills and current capacities to take over new tasks. Therefore, the blackboard exposes current tasks to the team, the middleware distributes the information within the team and the dialog planner enables a human-friendly translation (and access) to the middleware. The dual reality contains a representation of the scene on-site based on the information that is distributed via the central middleware.

A. Event-based Middleware: TECS

Communication between all team members as well as between all submodules and potential sensors and actuators of the instrumented environment is established via an event based middleware that has been developed in the project. TECS is short for Thrift Event-based Communication Service. As the name already implies, it is based on Apache's cross-language services-framework Thrift[5]. TECS uses the Thrift Interface Definition Language (IDL) to describe data-structures, which can then be translated into data objects of various programming languages (e.g. Java, C#, C++, Python etc.) using the Thrift compiler. These objects can then be transmitted via the (Java-based) TECS server as event-based messages. Clients of the TECS server can be addressed either via a broadcast or a publish-subscribe mechanism. The independence from a particular programming language is a very important feature for Hybrid Teams, as the multiple subsystems, especially present in real production scenarios, are usually implemented in various programming languages. We published TECS under Creative Commons (CC BY-NC 4.0)[6] License and it can be downloaded via our website[7].

B. Task Management: Blackboard

As mentioned above, the blackboard is a viable component of the presented architecture of the Hybrid Team. It stores all subtasks that have to be fulfilled in order to accomplish the main goal (packaging an item in our example scenario). In a strict sense, the blackboard is a (rather single minded) SoftBot with a very restricted world model: it only cares about unfinished working tasks. All team members have access to the information on the blackboard. Artificial agents do this via TECS, as the blackboard broadcasts all tasks it receives. Humans have access to the blackboard through a graphical representation, which can be seen on one of the large screens in the workshop or through Google Glasses. In the video, the blackboard is shown behind the virtual agent Gloria on the left screen. Each team member can decide if they are capable to fulfill the task and can then commit themselves to it. This is done with a first-come-first-serve policy, i.e. the fastest team member gets the job. As of now, humans can commit to a task using speech commands (e.g. "I will do task number eight"), through a GUI on a tablet or via swipe gestures on Google Glass. In the next iteration we will also do this by plan- or action-recognition, i.e. humans can just start fulfilling a task and the system will automatically assign the appropriate task to them.

C. Dialog Planner

The dialog planner processes speech input of the users and plans dialog acts for the artificial team members. As of now, Aila, Artemis and Gloria are capable of producing speech output via TTS. The TTS components themselves are located within the respective artificial team members. The dialog planner follows the *information state / update* approach ([10]), albeit in a modified form. The whole formalism and implementation for the planner is built on RDF and description logics, which is uniformly used for specification as well as storage of dialog memory, belief states, user models, etc. The extended functionality of the description logic reasoner HFC [11] makes it possible to equip the collected data with time information, which allows us to use the whole history of dialogs, user and other data for decision making. The management engine itself consists of complex reactive rules which are triggered by incoming data and have access to the complete data history.

D. Dual Reality

A dual reality component can be useful as a management or remote monitoring tool. It is visualizing information that is sent by the team via the central middleware. Using this information, a coordinator (even when not being on-site) can see what is going on and eventually influence the scene. Furthermore, new team members have a direct access to the current and previous states. This dual reality shown here is based on FiVES[8] (Flexible Virtual Environment Server). As of now, the dual reality component listens to position and motor events of the robotic team members as well as to events from the fusion module, and can thus display position and pose of the robots as well as of the worker.

[5] https://thrift.apache.org

[6] https://creativecommons.org/licenses/by-nc/4.0/

[7] http://tecs.dfki.de

[8] http://catalogue.fiware.org/enablers/synchronization-fives

V. The Demonstration Video

This section contains explanations to the accompanying demonstration video[9]. The video is structured into four parts that are shortly outlined in the following.

A. The Team in Action

This main part of the video shows the team performing the outlined packaging task. The video illustrates the interplay of the team when performing the task using communication via TECS, the blackboard and the dialog planner. After a customer has entered the entrance area, an initial greeting is triggered by the worker's speech command. As the customer requests to have an item packaged, the manager pushes a so-called BringTask on the blackboard (using a simple GUI) to get the item delivered from the reception desk to the workbench. Artemis autonomously chooses the BringTask and starts moving. In the upper part of the video and in a short segment later, the Dual Reality view of the moving robot can be seen.

The worker starts to work and pushes new tasks to the team via speech interaction as shown exemplarily in a dialog with Gloria. Again, the robot Artemis is taking over heading to the depot.

B. View At The Workbench

This second video part explains the properties of the robot Compi as well as the present state of interaction with the worker. The worker can either drag the arm to the desired position or adjust the arm's stiffness via a GUI on a tablet to have an optimal handling support for the material at hand.

C. Interaction with Gloria

The third sequence is focusing on the dialog planner and the role of the virtual agent. The human issues a speech command to Gloria, stating that he needs glue. As Gloria realizes that there is no more room on the workbench to place it, she makes suggestions based on what she has observed before. These suggestions depend on registered material usage (frequency and materials often used together, not shown here).

D. Augmented Worker

Finally, examples are shown how the data of the augmented worker is integrated and used in the system. Based on the information from the eyetracker, states and locations of objects are updated in the Dual Reality. Based on the observed locations, the worker (and objects) can be associated with predefined regions of interest. Additionally, the data complements sensory data from the robots to realize collision avoidance and safe coexistence in a shared workspace.

[9] http://hysociatea.dfki.de/?p=441

VI. Conclusions and Future Work

This video documents the first step in the realization of a Hybrid Team consisting of humans, robots, virtual agents and several SoftBots providing information from sensor analysis to the team. Based on the outlined communication architecture, the video shows the formation of the team process based on autonomous decisions of the individual agents involved. The current scenario provides a general setting, which will be extended in the second phase with much more complex interactions and contexts, e.g. distributed decomposition of tasks into subtasks, which will also lead to robot-robot interaction. Accordingly, the video demonstrates the new opportunities that Hybrid Teams will enable in the near future.

Acknowledgment

The research described in this paper has been funded by the German Federal Ministry of Education and Research (BMBF) through the project HySociaTea (grant no. 01IW14001). The authors would like to thank Annemarie Hirth for cutting and editing the video.

References

[1] H. Kagermann, J. Helbig, A. Hellinger, and W. Wahlster, *Recommendations for Implementing the Strategic Initiative INDUSTRIE 4.0: Securing the Future of German Manufacturing Industry; Final Report of the Industrie 4.0 Working Group*. Acatech – National Acadamy of Science and Engineering, 2013.

[2] S. Straube and T. Schwartz, "Hybrid teams in the digital network of the future – application, architecture and communication," *Industrie 4.0 Management*, vol. 2, pp. 41–45, 2016, in press.

[3] S. Joyeux, J. Schwendner, and T. M. Roehr, "Modular software for an autonomous space rover," in *Proceedings of the International Symposium on Artificial Intelligence, Robotics and Automation in Space*. i-SAIRAS, 2014, 8 pages.

[4] J. Lemburg, D. Mronga, A. Aggarwal, J. de Gea Fernández, M. Ronthaler, and F. Kirchner, "A robotic platform for building and exploiting digital product memories," in *SemProM – Foundations of Semantic Product Memories for the Internet of Things*, ser. Cognitive Technologies, W. Wahlster, Ed. Springer, 2013, pp. 91–106.

[5] J. Schwendner, T. M. Roehr, S. Haase, M. Wirkus, M. Manz, S. Arnold, and J. Machowinski, "The Artemis rover as an example for model based engineering in space robotics," in *Workshop Proceedings of the IEEE International Conference on Robotics and Automation 2014 (ICRA-2014)*. IEEE, 2014, 7 pages.

[6] V. Bargsten and J. de Gea Fernández, "COMPI: Development of a 6-DOF compliant robot arm for human-robot cooperation," in *Proceedings of the 8th International Workshop on Human-Friendly Robotics*. Technische Universität München (TUM), 2015.

[7] S. S. Mourkani, G. Bleser, N. Schmitz, and D. Stricker, "A low-cost and light-weight motion tracking suit," in *Proceedings of the International Conference on Ubiquitous Intelligence and Computing*. IEEE, 2013, pp. 474–479.

[8] G. Pirkl and P. Lukowicz, "Robust, low cost indoor positioning using magnetic resonant coupling," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 431–440.

[9] G. Pirkl, P. Hevesi, J. Cheng, and P. Lukowicz, "mbeacon: Accurate, robust proximity detection with smart phones and smart watches using low frequency modulated magnetic fields," in *Proceedings of the 10th EAI International Conference on Body Area Networks*. ICST, 2015, pp. 186–191.

[10] S. Larsson and D. R. Traum, "Information state and dialogue management in the TRINDI dialogue move engine toolkit," *Natural Language Engineering*, vol. 6, no. 3&4, pp. 323–340, 2000.

[11] H.-U. Krieger, "An efficient implementation of equivalence relations in OWL via rule and query rewriting," in *Proceedings of the IEEE International Conference on Semantic Computing (ICSC'13)*. IEEE, 2013, pp. 260–263.