AAAI Technical Report SS-12-02
Designing Intelligent Robots: Reintegrating AI

Designing Intelligent Robots for Human-Robot Teaming in Urban Search & Rescue *

Geert-Jan M. Kruijff, Francis Colas, Tomáš Svoboda, Jurriaan van Diggelen, Patrick Balmer, Fiora Pirri, Rainer Worst

Abstract

The paper describes ongoing integrated research on designing intelligent robots that can assist humans in making a situation assessment during Urban Search & Rescue (USAR) missions. These robots (rover, microcopter) are deployed during the early phases of an emergency response. The aim is to explore those areas of the disaster hotzone which are too dangerous or too difficult for a human to enter at that point. This requires the robots to be "intelligent" in the sense of being capable of various degrees of autonomy in acting and perceiving in the environment. At the same time, their intelligence needs to go beyond mere task-work. Robots and humans are interdependent. Human operators are dependent on these robots to provide information for a situation assessment. And robots are dependent on humans to help them operate (shared control) and perceive (shared assessment) in what are typically highly dynamic, largely unknown environments. Robots and humans need to form a *team*. The paper describes how various insights from robotics and Artificial Intelligence are combined, to develop new approaches for modeling human robot teaming. These approaches range from new forms of modeling *situation awareness* (to model distributed acting in dynamic space), *human robot interaction* (to model communication in teams), *flexible planning* (to model team coordination and joint action), and *cognitive system design* (to integrate different forms of functionality in a single system).

Introduction

Urban Search & Rescue (USAR) is a domain where robots can make a difference (Murphy et al. 2008). Robots may be able to enter disaster sites which are otherwise too dangerous or too difficult for humans to get to. Once there, robots can gather information about the situation, providing human operators with video feeds, maps, and sensor data. Using this information, humans might be able to make a better situation assessment, to aid emergency management.

There is a reason for why we phrase the above using can's and maybe's. Disaster areas are difficult environments to operate in, for humans and for robots. These are hardly robot-friendly places. Inevitably a deployment will experience what Woods et al (Woods et al. 2004) termed "***(Robin) Murphy's Law****: any deployment of robotic systems will fall short of the target level of autonomy, creating or exacerbating a shortfall in mechanisms for coordination with human problem holders*." If this were just a problem statement, more autonomy could be a possible solution; See also (Birk and Carpin 2006). But what is really at stake is the coordination between the robots and the humans involved. It is not 'just' the task-work, it is the combination of task-work with the interaction, the team-work, which we need to address; See also (Murphy 2004).

In this paper, we describe ongoing research on designing intelligent robots to help address that issue. We focus on four interrelated questions.

Q1 How to model a notion of situation awareness which (a) bridges the gap between a robot's quantitative, and a human's qualitative sense of space, (b) facilitates use by a geographically distributed team, and (c) provides the basis for understanding and planning individual or joint action (Q4)?

Q2 How to model the impact of situations in task- and teamwork which influence human performance, given that (a) humans typically perform under stress in USAR missions, and (b) stress alters interaction patterns (Q3) ?

Q3 How to model user-adaptive human-robot communication, to adjust how, what, and when a robot communicates given an awareness of the current operative situation (Q1) and its effects on human performance (Q2)?

Q4 How to model morphology-adaptive planning and execution, to guide and adjust how a robot plans and executes its own actions under different circumstances (Q1)?

We follow a user-centric design methodology in developing these approaches. Various rescue services and -organizations are involved throughout all phases of the development. Each year, we focus on physically realistic use

*This paper describes research done under the EU-FP7 ICT 247870 NIFTi project. For more about NIFTi, please visit `http://www.nifti.eu`. The paper was written as a team effort. For DFKI, GJ Kruijff, Mira Janíček, Shanker Keshavdas, Hendrik Zender, Benout Larochelle, and Harmish Khambhaita. For ETH Zürich, Francis Colas, François Pomerleau, Ming Liu. For CTU, Tomáš Svoboda, Tomáš Petriček, Karel Zimmerman. For Fraunhofer, Rainer Worst, Thorsten Linder, Slava Tretyakov, Hartmut Surmann. For ROMA, Fiora Pirri, Mario Gianni, Arnab Sinha, Panos Papadakis. For TNO, Jurriaan van Diggelen, Mark Neerincx, Nanja Smets, Tina Mioch.

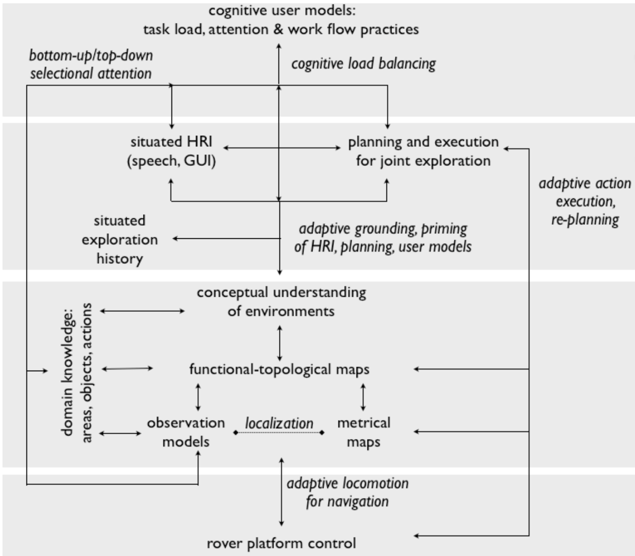


Figure 2: System architecture for a single robot

cases, in which we experiment with and evaluate our approaches. Fig. 1 illustrates one such use case, namely a tunnel accident involving a lorry, load, and multiple cars. The photos were taken at several end user training sites.

In this paper, we focus primarily on aspects of system design. Fig. 2 shows the schema of the overall architecture. The bottom half of the schema primarily concerns the continuous building and maintaining of robot-internal representations of the environment, and a robot's own internal control state. Hybrid maps are used to represent the environment. These maps combine metrical and topological structure to build up a qualitative level of representation. At that level, object and landmark observations can be grounded, as well as conceptual inferences about afforded actions (e.g. where to be to look for victims inside a car).

The resulting conceptual, grounded understanding of the environment is used by higher-level processes such as flexible planning and execution monitoring, situated dialogue processing, and cognitive user modeling. The relation between these higher-level processes, and the robot-internal model of the environment, is bi-directional. Higher-level processes anchor their interpretations and inferences in the environment model (bottom-up), while at the same time their projections can drive lower-level attentional- and behavioral processes (top-down). Given this bi-directionality, or interdependency between functionality, there is no strict separation between an "AI layer" and a "robotics layer." Functionality from AI and robotics is used across the board, combining probabilistic and logical forms of inference, to deal with uncertainty and incompleteness in observing, acting, interacting, and understanding while humans and robots jointly explore a complex environment.

In the system design, the human dimension of human-robot teaming is more than just a single "box," an add-on component. The human perspective is pervasive throughout the representations the robot builds. The conceptual understanding of the environment provides a human-like view on the environment, and the inference of spatially grounded affordances results in robot behavior that is transparent to a human operator ("this is where I would go if I were to look inside a car."); See also (Khambhaita et al. 2011). When it comes to human-robot interaction and planning, humans are explicitly modeled as actors, and action and interaction are planned in ways that conform to human operational practice. Finally, all of these processes interact with a dedicated process which continuously estimates the current task-load of human actors, their "stress," to provide an explicit model that can inform how a robot decides to act and interact.

Fig. 3 illustrates a how different components in the system architecture interact, dealing the command to go to a particular car. The system uses a mixture of ROS and CAST (Hawes and Wyatt 2010), to integrate components.

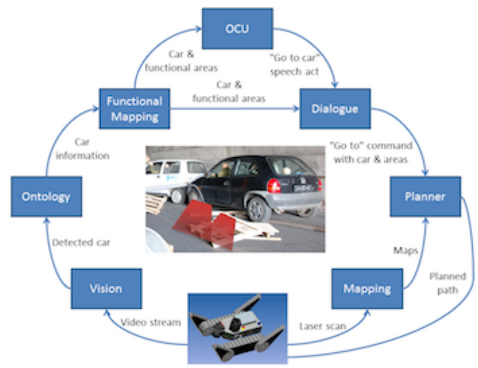


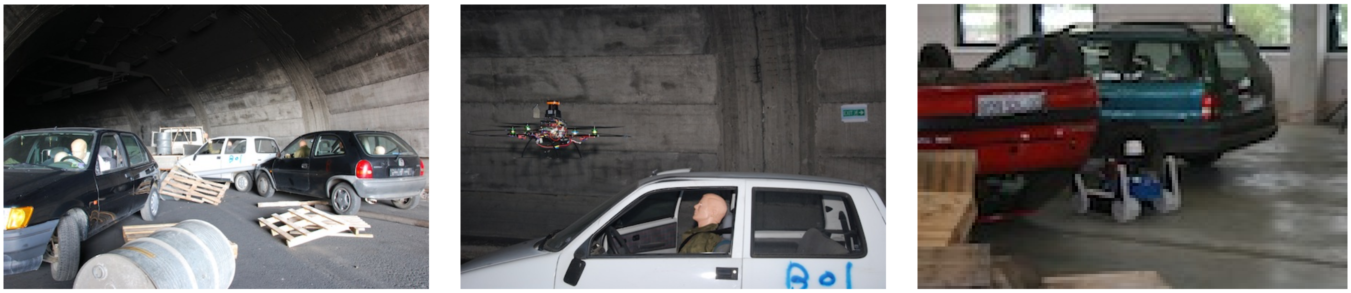
Figure 3: Interaction among components

Intelligence in Situation Awareness

The mission of the human-robot team is to explore a disaster area, to provide enough information to make a situation assessment. A human-robot team consists of at least one rover (UGV) and a microcopter (UAV), several humans located at a remote control post, and possibly one or more human operators in-field. The team is thus geographically dispersed. For situation awareness this requires the approach to be able to integrate different perspectives on the environment, (e.g. UAV, UGV, and descriptions from an in-field operator), and to facilitate different perspectives and needs; See also (Salmon et al. 2009). In the section below we focus primarily on the bottom-up construction of hybrid maps, up to a conceptual-functional description, thus dealing with question (1) from the introduction. This level of description is combined with functionality for interaction and planning, as discussed in later sections; See also Fig. 2.

3D metric and topological mapping

We address the mapping problem using several abstraction layers. First we try to build an accurate metric representation of the environment based on the 3D rolling laser sensor mounted on our robot. Based on this metric representation,



(a) (b) (c)

Figure 1: NIFTi tunnel accident use case: (a) Sample setup; (b) UAV as roving sensor; (c) UGV

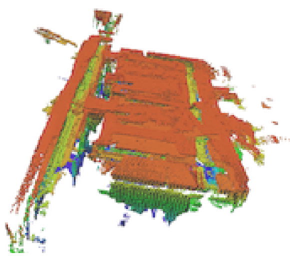


Figure 4: 3D map of two office rooms and a corridor. Warmer colors indicate more elevated obstacles.

we then segment the navigable space of the environment into coherent areas linked in a navigation graph.

Over the last two decades, metric mapping and localization have been addressed simultaneously as they rely on each other to proceed. Simultaneous Localization and Mapping (SLAM) is usually solved by approximating the maximum a posteriori probability over the joint distribution of the map and the pose history of the robot. Rao-Blackwellized particle filtering allows for efficient computation (Grisetti, Stachniss, and Burgard 2007). For 2D environments, several software packages exist that implement efficient 2D mapping based on 2D laser data, such as GMapping [1] or the Karto mapping library[2].

Going 3D requires both to have an efficient 3D representation of the environment and to be able to estimate the 6 degrees-of-freedom pose of our robot. The representation of the map is made using fast and flexible octrees (Wurm et al. 2010). Fig. 4 shows an example of such a 3D map. It has been taken in an office environment using the continuously rolling laser. To avoid part of the distortions, the 3D point clouds are registered into the map only when the robot is static. Preliminary results show that in most cases the distortion when the robot is moving is not too large, but the localization may jump from local optima and induce point cloud deformation due to the pose interpolation.

The 6 degrees-of-freedom pose estimate is based on a robust 2D map when the robot lies in a mostly 2D part of the environment. To handle 3D environment, we rely on fast and efficient 3D pose estimation (Pomerleau et al. 2011).

[1] http://www.ros.org.wiki/gmapping

[2] http://www.ros.org/wiki/karto

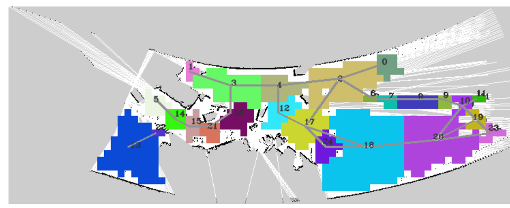


Figure 5: Topological segmentation of the tunnel environment. The navigation graph is shown in grey.

For the topological segmentation, we take as input the map of the environment. Previously we performed topological extraction based on spectral clustering and mutual information (Liu, Colas, and Siegwart 2011). In order to better handle changes in the map, both due to exploration and due to actual changes, we implement now a incremental topological segmentation. Fig. 5 depicts the result of this new method in the tunnel environment.

Integration of mapping and perception

Having a rich 3D point cloud and knowing the robot position relative to it may essentially improve results on some notoriously difficult computer vision problems. Image based detection of rear parts of cars in the tunnel accident use case works relatively well (Zimmermann, Hurych, and Svoboda 2011), see Fig. 6.

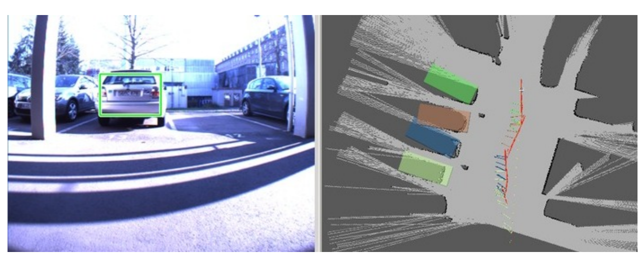


Figure 6: Car detection using visual features. 3D position is computed from multiple detections at different viewpoints, and odometry (red line in 2D map, created by GMapping)

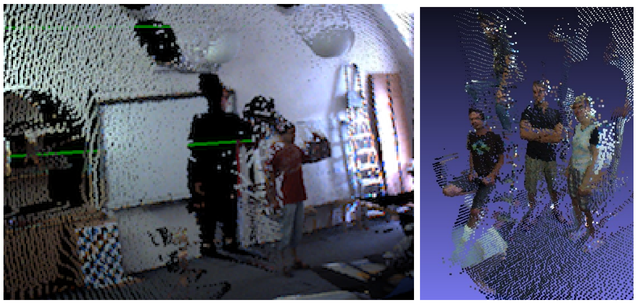


Figure 7: 3D point clouds from the rotating laser scaner colored by using an image from the Ladybug3 omnidirectional camera. The date are one-shot, i.e. both laser and image data are taken from one viewpoint. The displayed views are rendered from a different angle to visualize depth of the scene.

Unfortunately, estimating the 3D positions of cars proved to be much more difficult, especially the orientation. In order to address the issue of 3D instability we attach 2D features to the 3D (laser) maps. An example of assigning image colors to the 3D point clouds is shown in Fig. 7. More than just image colors may be assigned to the 3D points. The 2D object detector essentially creates a probabilistic map over the image. The detector can be trained on various car poses. The detector responses will be attributed to 3D points. The 3D information brings also the absolute scale which also allows for discarding many false alarms.

The result of connecting visual perceptions with the 2D- and 3D map representations we construct, is that we now obtain grounded observations of objects in the scene. We use these object observations to perform further inferences about the environment.

Functional mapping

Functional Mapping is a form of spatial inference. Given an object, and an action to be performed, functional mapping infers areas around the object, where the action can be performed relative to the object. This is a combination of logical inference over associated ontologies for objects and their internal structure, and for actions; and geometric inference. In the tunnel accident, functional mapping for example infers that being in a particular position relative to a car window facilitates looking into that car. The projection of the (functional) areas into space is based on real-time map data and the observed 3D pose of the object. Functional mapping thus combines top-down inferencing, from apriori knowledge of expected objects and bottom-up inferencing from real-time observations.

Inferring functional areas serves several purposes. First of all, when fire-fighters explore a disaster site, they themselves move between functional areas to make their observations (Khambhaita et al. 2011). We observed the same behavior when fire-fighters tele-operated robots to explore an accident, as shown in Table 1. Making the robot follow similar behavior makes that behavior *transparent* to an operator working with the robot. Secondly, we use the inference of functional areas to determine optimal vantage points for the robot to perform an observation.

Part.	% Observation time	% Observation time in functional areas	%Time in functional areas of objects	
			Vehicles	Threats
1	38.17	66.7	86.67	13.33
2	53	97.6	0	100
3	48	65.3	41.96	58.04

Table 1: Time each operator spent observing the environment (as % of run time), what % of that was spent in functional areas projected from objects, divided between types of objects. (Dated: January 2011)

In the tunnel accident use case, functional mapping uses a pre-defined ontology containing information on car models and the 3D positions of the windows on each model, and another ontology with specifications of the robot itself and the sensors present on it. Both ontologies are OWL/RDF-based, with classes which are based largely on WordNet (Fellbaum 1998). We use the forward chaining engine *HFC* (Krieger 2011) for inference over the ontologies. HFC is equipped with a standard OWL-DL rule set and a number of custom rules for drawing *default conclusions*. The inference over the car ontology yields a topological structure, with 3D locations for the windows; see Fig. 8.

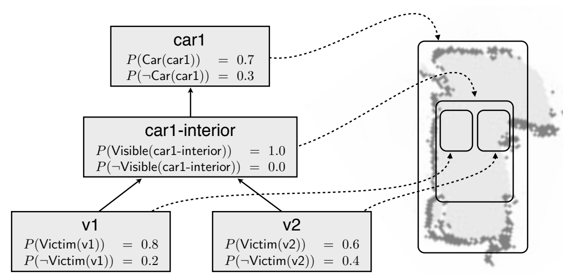


Figure 8: Inference of structure for observed car

The (grounded) inferences of car structure are then combined with inferences over the robot morphology and sensor performance (optimal distance, field of view), to determine optimal vantage points for looking inside the car. Fig. 9 (left) illustrates the maximal and minimal bounds in the robot's position for the robot's camera to observe a minimum patch size corresponding to the size of a human face for face-detection. All this information is then spatially projected onto the real time map, converted to vantage point planning poses, and visualized in the GUI, see Fig. 9 (right).

Intelligence in Team Interaction

Human-robot interaction is regarded one of the major bottlenecks in rescue robotics (Murphy 2004; Murphy et al. 2008). Tele-operating a robot is highly demanding on a human operator. More autonomy appears to be a way out of this. But more autonomy also raises the need for humans and robots to coordinate action. And that requires human-robot communication. Unfortunately, most models of human-robot

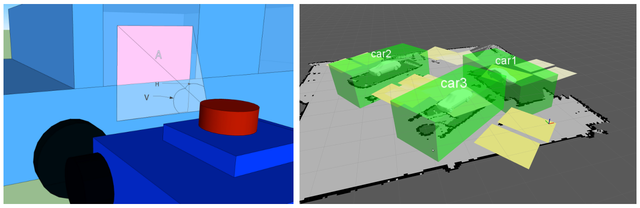


Figure 9: (l.) If detected Area A>*p*, patch size for reliable face detection, robot position gets included in functional area. (r.) Functional areas during pilot runs in Dortmund

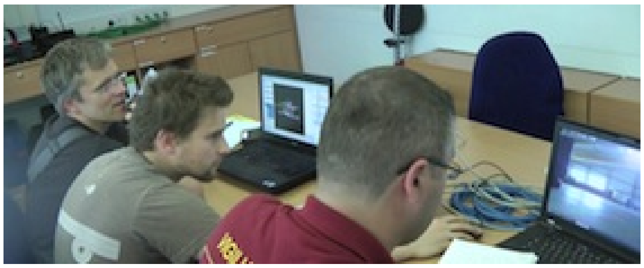


Figure 10: Human team at remote control post: UGV pilot, Mission Director, Mission Specialist

communication have so far been relatively limited in their (possible) use of spoken dialogue, one of the most natural means for humans to interact. Furthermore, these models typically do not ground communication in the social structure, to explain why actors are interacting (or need to), and what information is to be exchanged (or not). Below we briefly discuss an advanced model of situated dialogue processing that makes it possible to model spoken interaction in collaborative activities like human-robot teaming in USAR. This helps address issues in questions 2 and 3 mentioned in the introduction.

Interaction in collaborative activities

Like any form of human-robot interaction, communication between humans and robots in a human-robot team can be proximal or remote (Goodrich and Schultz 2007). Fig. 10 shows an example of remote interaction between humans located in a control post, and two in-field robots (one of them, the UAV, operated by a pilot in-field). People communicate with one another face-to-face and using hand-held radios, and they can interact with the robots using a multi-modal GUI that includes spoken dialogue (Larochelle et al. 2011).

Dialogue in human-robot dialogue is typically about the environment, and the tasks to perform. We have been working on a model of dialogue processing that explicitly places dialogue in such a situated, social context (Kruijff, Janíček, and Lison 2010). Dialogue is interpreted relative to a set of multi-agent beliefs and intentions that are situated in space and time (Lison, Ehrler, and Kruijff 2010), and social structure (Kruijff and Janíček 2011). This interpretation is a continuous process involving the composition of a representation of (linguistically expressed) meaning, the formation of hypotheses for how referring expressions can be anchored in context are formed, and the inference of possible intentions as explanations of how the utterance meaning could fit into the dialogue flow. This process of forming intention, intension, and extension (or denotation) is capable of dealing with uncertainty and incompleteness in interpretation, by combining logical and probabilistic forms of inference. The result of this process is an intention (or, more precisely, an abductive proof) which indicates how to use the provided information to update the robot's situated multi-agent beliefs and intentions. Following up on this, the robot formulates an appropriate intention in response, selects the actions to realize the intention (again, formulated as an abductive proof to see how best to anchor these in existing beliefs and intentions), and then carries them out.

The model we adopt is based in earlier approaches to collaborative dialogue, e.g. (Grosz and Sidner 1986) and recently (Stone and Thomason 2003). Our model improves on these approaches by providing means to deal with uncertain, incomplete, or possibly wrong information (Kruijff, Janíček, and Lison 2010), as is typical for spoken dialogue processing as such, and particularly for *situated* dialogue processing which has to content with uncertainty and incompleteness pervasive throughout the robot's understanding of the world, and of other actors. Another aspect is that we are currently extending the approach to include an explicit model of the social dynamics in human robot-teams.

Modeling team interaction

The situatedness in situated dialogue is about more than just "the world." There is the environment as it is described, there is the past, present, and the future, there is simply the fact that the actors themselves are in or connected to that environment; See also (Ginzburg 2011). Each actor has a personal perspective on that reality. And, as (Murphy and Burke 2010) argue for, this perspective is determined to an important degree by the role that actor plays in a team. For example, in a UAV team, the pilot closely watches that part of the situation in which the UAV is flying, whereas the mission specialist uses the UAV's on-board video camera to look at the situation from the UAV's viewpoint. These are typically different views. However, they need to be aligned in communication if the pilot and the mission specialist are to maintain a common ground in understanding the situation, to coordinate actions.

In (Kruijff and Janíček 2011) we describe a model of human-robot team roles, and the social network between these roles. The model follows up on the communicative analysis of (Burke et al. 2004) but expands it with the notion of level of (adaptive) autonomy from (Parasuraman, Sheridan, and Wickens 2000) to be able to explicitly model ranges of shared control between roles, and to provide a basis for reasoning about the dynamics of role shifting (delegation) within a team. Given an instantiation of roles to actors in a team, and grounding the information that gets communicated between the actors (as per their roles), the model yields a perspective on team situation awareness that is highly distributed – namely, tied to roles and interdependency between roles, not to "the team" as such; See also (Salmon et al. 2009). To investigate the dynamics of human-robot teaming in practice, we have conducted real-life exercises at a training center of the Fire Department of Dortmund, and are using a setup to explore specific aspects of this interaction in a controlled setting.

To support the latter investigation we have developed a collaborative tagging environment and tailored it to the HRI domain. The system is called Trex: Tagging-based Realtime Exhibitor. It contains views with basic functionalities such as indicating dangerous areas on a map, leaving messages and camera-images on it, and seeing where your colleagues are. With respect to organizational issues, we have developed a so-called organisation awareness view, illustrated in Fig. 11. The user interface is based on an information structure that represents three main aspects: organization, missions and resources. Organization covers all aspects of the organizational structure, such as roles with corresponding authorizations and responsibilities and hierarchical relations between roles. The mission includes all information related to the mission, such as the main mission goal, a division in sub-goals, and a mission plan of how to achieve those goals. Finally, resources include human actors (e.g. tele-operators) or system resources (robots, UAV's) with their capabilities and states (position, availability, etc.).

Figure 11: Exploring human-robot teaming using Trex

In an operational organization, a tight interconnection exists between all the aspects above. For example: a human actor enacts an organizational role, and therewith he takes up the responsibility for the mission objectives that were assigned to the role. In the organization ontology we specify the semantic relations between knowledge elements in such way that these derivations can be automatically derived by an OWL reasoning system.

The interconnection between the different aspects also implies that changing one aspect has impact on the others. For example, if an actor is no longer capable to fulfill his task, the consequence could be that the mission plan is no longer executable, and needs a change. Several cross sections of the information can be shown depending on the needs of a user. For example, a hierarchical tree is convenient to show the organisational structure in terms of superior relations between the roles. Another possibility is to combine the organisation awareness aspects with for example position information. Fig. 11 shows such geographical interface. Each actor is plotted as a symbicon with their current task written below. The green and red bars on the sides indicate whether they are capable of and authorized for their tasks. In this particular shot, the UGV (indicated with a G) is not able to Explore. By sending a query for the Explore capability to this interface, the known resources with that capability will pop up (e.g. the UAV).

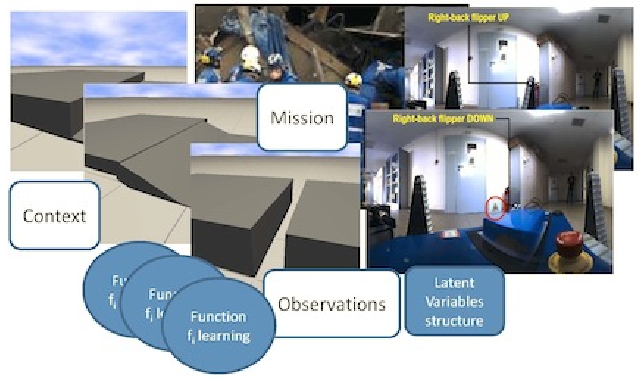


Figure 12: Hierarchical modeling of morphological adaptation, for two kinds of contexts: step actions, task switching.

Intelligence in Team Cooperation

A human-robot team consists of at least one rover (UGV) and a microcopter (UAV), several humans located at a remote control post, and possibly one or more human operators in-field. The team is thus geographically dispersed. For team cooperation this requires the approach to be able to integrate different perspectives on the environment, (e.g. UAV, UGV, and descriptions from an in-field operator), and to facilitate different perspectives and needs, see (Salmon et al. 2009). Two aspects are here stressed: how the robot can adapt to the asperities of the situation, in terms of terrain piles of debris and clutter, still keeping all its parts in full functioning, and how it can coordinate the control of its components with the state of the other team members. These issues address question 4 from the introduction. In the next two sections we first introduce a brief description of the morphological adaptation problem and further introduce the basics elements of the planning model, supporting a coordinate execution.

Morphology-adaptive planning for operation

Morphology adaptation here is intended as the ability of the robot to face territory harshness subject to the requirements of the mission. We have designed a new robot platform that is capable of both passive and active forms of morphological adaption, see Fig. 13.

Given the available information from the sensors on the surrounding region where to operate the task, and given the support of the UAV and the team operator, *planning* morphological adaptation ought to: (1) choose the best robot configuration at each time step t, to consistently face the terrain conditions under the kinematic constraints; and, (2) choose the best next sequence of actions, consistently with the robot

Figure 13: Robot with active and passive morphology

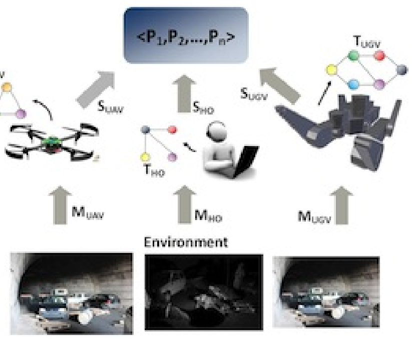


Figure 14: Properties $P_1, ..., P_n$ are defined on top of the inner states S of each team unit, to give a uniform representation of the multi-agent system. M denotes perceptual model of rescue unit, T temporal model of unit activities.

configuration constraints, to reach the next goal state. Early modeling of these two crucial aspects of robot motion and action execution, requires supervised learning of primitive actions to accommodate any-terrain path planning and also dynamic adjustments with respect to task prescriptions. In other words, morphological adaptation is a two way model, on one side towards the correct kinematic configuration of the robot, on the other side towards the strict tasks requirements. A multilevel regression model (see (Gelman and Hill 2006)) can account for different predictors levels: for primitive actions, for sequences of actions, for classes of actions (namely bag of situations) when several contexts, including team interaction have to be taken into account. Observations can be drawn both online, while the robot is teleoperated or via its simulation designed in Gazebo (Koenig and Howard 2004), see Figure 12. These observations form a random vector $x \in R^n$ drawn independently from different unknown distributions. Therefore observations are coupled with a latent variable structure accounting for context switching. The problem we face is that of determining, according to the context, the function and the parameters that best approximate the supervisor response, in each context. Variation is therefore accounted both with respect to primitive actions (those determined by angle, acceleration, velocity, and similar parameters) and with respect to high level actions (those determined by state preconditions parameters). Predictors are, therefore modulated by the different contexts.

Flexible temporal planning for co-operation

The dynamics of the UGV and UAV can be modeled separately by defining two different temporal declarative models in the Temporal Flexible Situation Calculus (TFSC) (Finzi and Pirri 2005). The UAV can act in strict cooperation with the UGV, to this end the TFSC model ought to know the states of both the system components, via a common language and model. This is briefly specified in the sequel. The hybrid framework combines temporal constraint reasoning and reasoning about actions. The flexible behaviors of the UAV and UGV are specified in a compact representation by temporal constraint networks T_{UAV} and T_{UGV}, while any team operator can be represented, in turn, with a network T_{HO}. These causal and temporal relations, with their constraints are learned by continuous interaction with humans, via demonstration and by the collected observations of successful processes achieved in classified contexts (Pirri 2011; Khambhaita et al. 2011). The networks are mapped into a structure managing time, resources and actions, namely, the model-based control. The model accounts for timelines with time flexibly assigned to each component so as to satisfy priorities for both resources and tasks, and that rely on online acquisition of sensor data (Gianni et al. 2011). The whole set is managed by the execution monitor that loops over the updating of the environment model $\{M_{HO}, M_{UAV}, M_{UGV}\}$ and of the inner states $\{S_{HO}, S_{UAV}, S_{UGV}\}$. The execution loop will ensure that the network is kept satisfiable, and it is extended accordingly. Indeed, it has been proved that this implies satisfiability of the processes modeled. The inner states S_{UAV} and S_{UGV} represent the internal loop that check on all of the machine components, namely both of the UAV and UGV. The human-robot(s) team shares the information about the environment and the mission, combining together their models of the current percepts. In order to integrate the different abilities of the UAV and the UGV with human operators intervention, a set of properties $P_1, ..., P_n$ are defined on top of the inner states of the team units bridging the different dynamic models (see Fig. 14).

These properties are modeled in a common language and constitute the substrate of knowledge communication among agents. A priority queue can be established on the set of properties to ensure that a process is started by the first team member that can subscribe to the requested execution. When a process is initiated by a team member, unless otherwise specified, the process is attributed to it. The underlying properties, which the task execution satisfies, are entailed in the model of the ascribed team member.

Conclusions

The paper presents a (dense) overview of how we combine different techniques from Artificial Intelligence and Robotics to build intelligent robots which can act as team members in an USAR human-robot team. Techniques from AI and robotics *enhance* each other. There is no “AI layer” separate from a “robotics layer.” They are used across the board, combining probabilistic and logical forms of inference, to deal with uncertainty and incompleteness in observing, acting, interacting, and understanding while humans and robots jointly explore a complex environment.

References

Birk, A., and Carpin, S. 2006. Rescue robotics - a crucial milestone on the road to autonomous systems. *Advanced Robotics* 20(5):595–605.

Burke, J.; Murphy, R.; Coovert, M.; and Riddle, D. 2004. Moonlight in Miami: An ethnographic study of human-robot interaction in USAR. *Human Computer Interaction* 19((1–2)):85–116.

Fellbaum, C., ed. 1998. *WordNet: an electronic lexical database*. MIT Press.

Finzi, A., and Pirri, F. 2005. Representing flexible temporal behaviors in the situation calculus. In *Proceedings of the 19th international joint conference on Artificial intelligence*, IJCAI'05, 436–441. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Gelman, A., and Hill, J. 2006. *Data Analysis Using Regressio and Multilevel/Hierarchical Models*. Cambridge University Press.

Gianni, M.; Papadakis, P.; Pirri, F.; Liu, M.; Pomerleau, F.; Colas, F.; Zimmerman, K.; Svoboda, T.; Petricek, T.; Kruijff, G. J. M.; Zender, H.; and Khambhaita. 2011. A unified framework for planning and execution-monitoring of mobile robots. In *Proceedings of the AAAI-11 Workshop on Automated Action Planning for Autonomous mobile Robots*.

Ginzburg, J. 2011. Situation semantics and the ontology of natural language. In Portner, P.; Maierborn, C.; and von Heusinger, K., eds., *The Handbook of Semantics*. de Gruyter. 830–851.

Goodrich, M. A., and Schultz, A. C. 2007. Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction* 1(3):203–275.

Grisetti, G.; Stachniss, C.; and Burgard, W. 2007. Improved techniques for grid mapping with Rao-Blackwellized particle filters. *IEEE Transactions on Robotics* 23(1):34–46.

Grosz, B. J., and Sidner, C. L. 1986. Attention, intention and the structure of discourse. *Computational Linguistics* 12(3):175–204.

Hawes, N., and Wyatt, J. 2010. Engineering intelligent information-processing systems with cast. *Advanced Engineering Informatics* 24(1):27 – 39.

Khambhaita, H.; Kruijff, G.; Mancas, M.; Gianni, M.; Papadakis, P.; Pirri, F.; and Pizzoli, M. 2011. Help me to help you: How to learn intentions, actions and plans. In *Proc. AAAI Spring Symposium Help Me Help You: Bridging the Gaps in Human-Agent Collaboration*.

Koenig, N., and Howard, A. 2004. Design and use paradigms for gazebo, an open-source multi-robot simulator. In *In IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2149–2154.

Krieger, H. 2011. A temporal extension of the Hayes and ter Horst entailment rules for RDFS and OWL. In *AAAI 2011 Spring Symposium "Logical Formalizations of Commonsense Reasoning"*.

Kruijff, G., and Janíček, M. 2011. Using doctrines for human-robot collaboration to guide ethical behavior. In *Proceedings of the AAAI 2011 Fall Symposium Robot-human team-work in dynamic adverse environments*. AAAI.

Kruijff, G.; Janíček, M.; and Lison, P. 2010. Continual processing of situated dialogue in human-robot collaborative activities. In *Proceedings of the 19th International Symposium on Robot and Human Interactive Communication (RO-MAN 2010)*.

Larochelle, B.; Kruijff, G.; Smets, N.; Mioch, T.; and Groenewegen, P. 2011. Establishing human situation awareness using a multi-modal operator control unit in an usar human-robot team. In *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE.

Lison, P.; Ehrler, C.; and Kruijff, G. 2010. Belief modelling for situation awareness in human-robot interaction. In *Proceedings of the 19th International Symposium on Robot and Human Interactive Communication (RO-MAN 2010)*.

Liu, M.; Colas, F.; and Siegwart, R. 2011. Regional topological segmentation based on mutual information graphs. In *Proc. of the IEEE International Conference on Robotics and Automation (ICRA)*.

Murphy, R., and Burke, J. 2010. The safe human-robot ratio. In Barnes, M., and Jentsch, F., eds., *Human-Robot Interactions in Future Military Operations*, Human Factors in Defence. Ashgate. 31–49.

Murphy, R.; Tadokoro, S.; Nardi, D.; Jacoff, A.; Fiorini, P.; Choset, H.; and Erkmen, A. 2008. Search and rescue robotics. In Siciliano, B., and Khatib, O., eds., *Springer Handbook of Robotics*. Springer Verlag. Part F, 1151–1173.

Murphy, R. 2004. Human-robot interaction in rescue robotics. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 34(2):138–153.

Parasuraman, R.; Sheridan, T. B.; and Wickens, C. D. 2000. A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics. Part A: Systems and Humans* 30:286–297.

Pirri, F. 2011. The well-designed logical robot: Learning and experience from observations to the situation calculus. *Artificial Intelligence* 175(1):378 – 415. John McCarthy's Legacy.

Pomerleau, F.; Magnenat, S.; Colas, F.; Liu, M.; and Siegwart, R. 2011. Tracking a depth camera: Parameter exploration for fast icp. In *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

Salmon, P.; Stanton, N.; Walker, G.; and Jenkins, D. 2009. *Distributed Situation Awareness: Theory, Measurement, and Application to Teamwork*. Human Factors in Defence. Ashgate.

Stone, M., and Thomason, R. 2003. Coordinating understanding and generation in an abductive approach to interpretation. In *Proceedings of DIABRUCK 2003: 7th workshop on the semantics and pragmatics of dialogue*.

Woods, D.; Tittle, J.; Feil, M.; and Roesler, A. 2004. Envisioning human-robot coordination for future operations. *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews* 34(2):210–218.

Wurm, K. M.; Hornung, A.; Bennewitz, M.; Stachniss, C.; and Burgard, W. 2010. OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems. In *Proc. of the ICRA 2010 Workshop on Best Practice in 3D Perception and Modeling for Mobile Manipulation*. Software available at `http://octomap.sf.net/`.

Zimmermann, K.; Hurych, D.; and Svoboda, T. 2011. Improving cascade of classifiers by sliding window alignment in between. In *The 5th International Conference on Automation, Robotics and Applications*. IEEE.