Delegation in Human-machine Teaming: Progress, Challenges and Prospects

Jurriaan van Diggelen, Jonathan Barnhoorn, Ruben Post, Joris Sijs, Nanda van der Stap, Jasper van der Waa

TNO,
the Netherlands
{jurriaan.vandiggelen, jonathan.barnhoorn, ruben.post, joris.sijs, nanda.vanderstap, jasper.vanderwaa}@tno.nl

Abstract. This paper describes DASH: Delegation to Autonomous Systems within Human-machine teams. The purpose of DASH is to combine ideas from recent progress in human machine interaction to confront current and prospected challenges of distributed delegation and meaningful human control. We describe the design considerations of DASH and illustrate how delegation using plays, a web of interaction modules, and interdependency graphs are used in the context of a military human-robot team.

Keywords: Human Factors · Human-systems Integration · Systems engineering

1 Introduction

For over two decades, delegation has been recognized as a crucial notion within the context of autonomy, cooperation, and collaboration [1]. It allows a delegator to set an agenda either broadly or specifically, but leaves some authority to the subordinate to decide exactly how to achieve the commands supplied by the delegator [2]. Research has made significant progress on a wide range of issues that are essential for enabling humans to delegate tasks to machines, such as developing appropriate levels of trust [3], allowing the human to have situation awareness of a diverse range of aspects of the work domain [4], and characterizing different levels of automation between human delegator and the machine [5].

Recent developments in artificial intelligence and robotics have significantly increased the application possibilities of delegation [6], yet raised new challenges as well:

Meaningful human control. As AI becomes ever more capable, it is applied to an increasingly wide range of problems, including ethically sensitive ones such as medical diagnosis, balancing driving speed and safety, and warfighting. While the topic is heavily discussed in many respects [7], there is general consensus that the delegation of task execution, responsibility, and accountability from humans to machines is not without limits. The human must always remain in control of ethically sensitive decisions. Designing delegation that safeguards this is one of present time's profound challenges.

Distributed delegation. AI is hardly used by one user [8, 9], nor does it take place at one particular moment. For example, an organization might set behavioral constraints to a system during deployment, while an operator commands the AI system to pursue a certain goal during operation. Both of these instructions determine the AI system's behavior. Furthermore, the machine may also act as a delegator, e.g. for delegating subtasks to other machines or even back to humans (as is the case when humans and machines work together as equal teammates). Most current delegation systems assume a one to one relation between delegator and subordinate. To be future proof, these mechanisms should be extended to enable various forms of delegation that are distributed in time and source of the command (see Figure 1). As shown in Figure 1d, the most advanced form of delegation in a HAT also allows machines to delegate tasks to humans, leading to complex delegation chains.

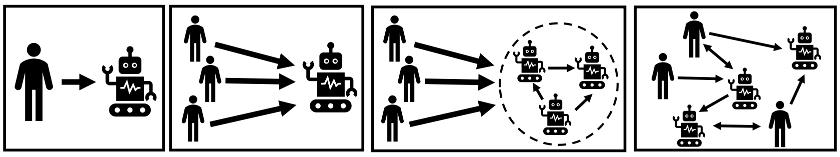

Figure 1: Various forms of distributed delegation

The challenges of distributed delegation and meaningful human control are closely related. A useful analysis of this dynamic was written by Chen et al. [12], describing various complexities such as development of trust over time, and mutual understanding of each other's capabilities.

This paper presents our experiences confronting these challenges while designing a uniform delegation system (called DASH) for controlling multiple intelligent assets in an ISR (Intelligence, Surveillance and Reconnaissance) task. DASH is based upon insights from ecological interface design [4], and aims at increasing the human's awareness of the sociotechnical system and environment by providing customizable views at different levels of abstraction, tailored to the desired level of involvement. Furthermore, DASH supports various ways of distributed delegation by connecting multiple humans and machines to the interface. This paper lays out the design principles of DASH, our initial experience with them, and describes their relation to the current challenges of delegation systems.

2 DASH system design

The challenges described above require us to rethink delegation on a functional, knowledge, system, and interaction level. The main design considerations at each of these levels is specified below.

2.1 Functional design

The table below describes the most important requirements for Observability, Predictability and Directability (resp. OR, PR, and DR).

OR1	DASH provides insight in the work system from the perspectives of tasks, goals, constraints, and resources
OR2	DASH provides insight in all directives that were issued, by whom, and how.
PR3	DASH provides insights in which behaviors are expected in the future.

DR4	DASH allows delegation by multiple actors at multiple moments.
DR5	DASH allows delegation at various levels of abstraction.

DASH should allow the user to issue a directive at an abstract level (DR5), but to drill down if needed to set the details (depending on how much freedom the delegator wants to grant to the subordinate). To maintain *meaningful human control*, humans must be aware of the current situation (OR1) and expected situations in the future (PR3, OR2). Distributed delegation is formulated by DR4, OR2.

2.2 Interaction design

To translate the functional requirements described above into design solutions, we propose three components:

- **Web of DIMs** Understanding the workspace (OR1) completely and from all points of view would be too complicated and would lead to overly complex interfaces that require fulltime attention of a supervisory control operator (e.g. [4]). Dependent on specific user-needs, the web of DIMs (Dash Interaction Modules) provides insight in one particular aspect from any particular point of view by adjusting the workspace's perspective (task, goal, resources, constraints), visualization (e.g., map or temporal) and level of detail (OR1, OR2).
- **Plays:** Whereas DIMs serve to satisfy the observability requirements, plays are used to sets things in motion (directed at DR4, DR5). A play (as originally proposed in [2]), allows for rapidly communicating a directive. This can be done with a *low level of human control*, e.g. by specifying which goal must be achieved to a highly capable machine, or with a *high level of human control,* e.g. by specifying exactly which resources must perform which tasks (OR1). Because all types of plays are known by the different actors in advance, this allows them to anticipate on each other's actions (PR3).
- **Interdependency graphs:** Large sociotechnical systems comprise a complicated network of interdependencies between its members (human or machine), which can be made transparent by so-called *interdependency graphs.* For example, they facilitate coordinating the use of a shared asset for multiple purposes (OR2, DR4).

2.3 Knowledge design

DASH follows a knowledge-based systems engineering approach [10], meaning that the interactions between humans and machines are for the most part defined by a shared knowledge model or ontology. The ontology must be aligned with both the human mental model of the work domain, as well as with the machine internal workings. This led us to adopt the following main concepts (adapted from the well-established theory CWA [11], and ecological interface design [4]):

- **resources** are agents, tools and consumables, that are required to perform tasks in an environment. They exist at various levels of aggregation (e.g. *Human Machine Teams, humans, robots, cameras, locomotion systems*, etc.).
- **goals** are the objectives that agents aim to achieve. An example of a high-level goals is *provide safety.* Examples of low-level goals include *obtain-optical-image,* or *located-at-waypoint-X.*
- **tasks** describe activities that agents perform to achieve their goals. Again, these exist at various levels of detail, e.g. *carry-out-air-support*, or *move-forward*.

- **constraints** are limitations on how goals should be achieved and tasks may be performed. These may be legal constraints (e.g. *rules of engagement*) that apply to the whole human machine team, or more specific user-defined constraints (e.g. on the usage of certain sensors), and can be aggregated on different levels.

The concepts above allows a *passive* understanding of the work system but are not yet sufficient to initiate actual activity. Action occurs when goals are required to perform above a certain ***threshold*** (according to a well-defined ***performance measure***); or when **tasks** are scheduled to be executed within a ***plan***. This is done by initiating **Plays** which can be understood as parameterizable templates for commonly occurring forms of delegation. For example, a play might state a sequence of **tasks** that a UAV must visit a specified location and record camera-imagery. A play may also be more abstract in terms of a high level **goal** and **constraints**, e.g. by stating that threat-level must remain minimal and use of force must be avoided at all times. Specifying multiple plays simultaneously (by multiple actors), or specifying static constraints in the work system and calling plays afterwards, can be regarded as ways to achieve distributed delegation.

3. Scenario

The following scenario illustrates some of the requirements and proposed functionality that we implemented in an early DASH prototype.

Figure 2: Scenario

An embassy has been ambushed. A VIP is held hostage and a rescue operation is initiated and coordinated from a Forward Operating Base (FOB) on sea. One of two teams on the shore performs perimeter control. The other team rescues the hostages from the embassy terrain. One of the latter team's members is "John", whom we will focus on. Three UAVs support the operation: a "Logistics" UAV supporting both teams when requested, a "Spotter1" UAV providing SA to the perimeter team, and a "Spotter2" UAV providing, via John, SA to the team raiding the embassy. Personnel on the FOB supports the UAVs for activities that the machines cannot perform independently, among which task selection (in case of conflicting requests) and verifying identification of contacts. Since the teams are working in a contested electromagnetic environment, dropping radio connections are a considerable risk.

Event 1: Connections between the FOB and the shore become unstable.

FOB personnel uses interdependency graphs to establish which UAV activities currently depend on interaction with the FOB. Personnel from the perimeter team is requested to temporarily take-over supporting the UAVs in these tasks. The task of verifying identification of contacts cannot be performed from land and thus John is warned that this functionality will be disabled in case of lost connections.

- In this event, the interdependency graphs provide insight into the current composition of the network based on an operational need. Gaining understanding of the

- problems that would result of a connection-drop allows humans to respond pro-actively (see OR1).
- Losing Spotter1's ability to identify contacts is ethically sensitive and by making this known to the relevant actors, ***meaningful human control*** is supported.

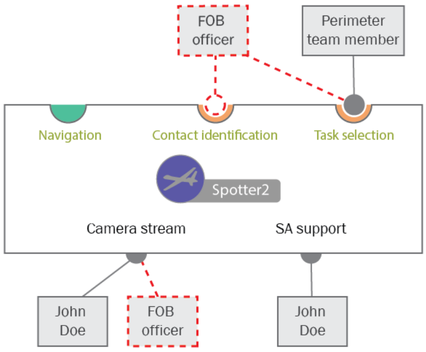

Figure 3: Interdependency graph illustrating the consequences of connection loss between shore and FOB.

Event 2: John finds a hostage.
Using the plays-interface, he quickly re-tasks Spotter1 to first scan the route to the drop zone at the embassy entrance, return and guide John to the entrance, and then return to the next building to be searched.

- Using the DASH interface (see Figure 4 allows John to delegate a set of tasks immediately using plays, increasing efficiency and reducing attentional demands (DR5).
- The initial play 'follow and look ahead' comprises a large set of tasks, called for using a single abstract play with many details filled in by the UAV. The 'surveil route' task is an example of a much more specific play (DR5).

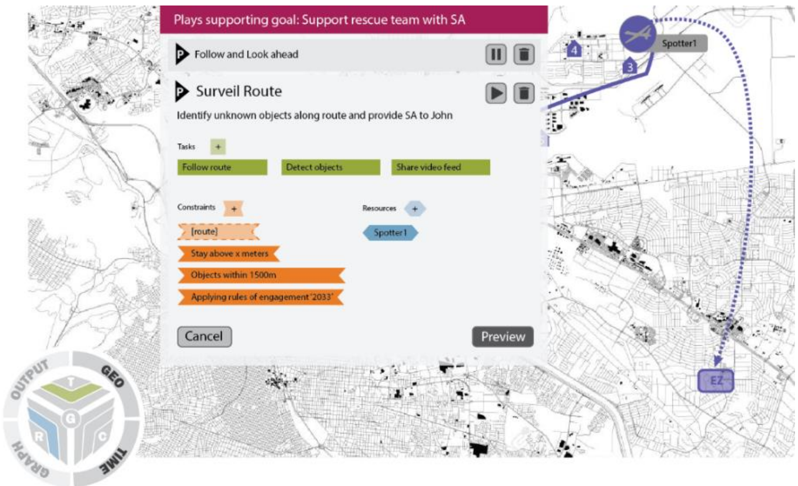

Figure 5: An overview of the play interface, which is overlayed on a DIM the user is currently viewing.

Event 3 Connections between FOB and the UAVs are re-established
John is informed about this. He wants to know whether Spotter1 is receiving support in identifying contacts again and checks the interdependency graph for Spotter1.

- Each actor in the network has different needs in terms of understanding the work system, here, John finds information relevant for him easily by selecting the right combination of Resource (Spotter1), Constraints (specifically: tasks constrained by a requirement for human-support), and Level of detail (see OR1).

4. Conclusion

This paper discusses progress, challenges and prospects in delegation architectures. In the light of rapidly advancing AI technology, distributed delegation and meaningful human control can be identified as two main challenges. We believe that the following areas of progress in intelligent human systems integration will play a vital role: play-based technology that allows delegation at multiple levels of abstraction (thereby varying the level of human control) by different users, and observability technology which allows humans to observe a work system from different perspectives at different levels of detail. We described DASH as an example that combines the two technologies and illustrated its merits in a multi-robot military mission.

Future work on DASH will focus on evaluating the system against the observability, predictability, and directability requirements, and designing an appropriate set of plays and testing them against requirements of meaningful human control.

References

1. Castelfranchi, C., & Falcone, R. (1998). Towards a theory of delegation for agent-based systems. Robotics and Autonomous systems, 24(3-4), 141-157.
2. Miller, C. A. (2005). Delegation Architectures: Playbooks and Policy for Keeping Operators in Charge. WS3, 28.
3. J. Y. C. Chen, M. J. Barnes and M. Harper-Sciarini, "Supervisory Control of Multiple Robots: Human-Performance Issues and User-Interface Design," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 41, no. 4, pp. 435-454.
4. Bennett, K. B. (2017). Ecological interface design and system safety: One facet of Rasmussen's legacy. Applied ergonomics, 59, 625-636.
5. Parasuraman, R., Sheridan, T. B., & 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(3), 286-297.
6. Peeters, M. M., van Diggelen, J., Van Den Bosch, K., Bronkhorst, A., Neerincx, M. A., Schraagen, J. M., & Raaijmakers, S. (2020). Hybrid collective intelligence in a human–AI society. AI & SOCIETY, 1-22.
7. Horowitz, M., & Scharre, P. (2015). Meaningful human control in weapon systems: A primer. Washington: Center for a New American Security.
8. Roth, G., & Schulte, A. (2020, July). A Concept on the Shared Use of Unmanned Assets by Multiple Users in a Manned-Unmanned-Teaming Application. In International Conference on Human-Computer Interaction (pp. 189-202). Springer, Cham.
9. Voshell, M., Tittle, J., & Roth, E. (2016). Multi-Level Human-Autonomy Teams for Distributed Mission Management. AAAI Spring Symposia.
10. Beckers et al (2019). Intelligent autonomous vehicles with an extendable knowledge base and meaningful human control. In Counterterrorism, Crime Fighting, Forensics, and Surveillance Technologies III (Vol. 11166, p. 111660C). Int. Soc. for Optics and Photonics.
11. Naikar, N., Hopcroft, R., & Moylan, A. (2005). Work domain analysis: Theoretical concepts and methodology (No. DSTO-TR-1665). Defence Science and Technology Organisation Victoria (Australia) Air Operations Div.