

Human-Autonomy Teaming: Supporting Dynamically Adjustable Collaboration

(STO-TR-HFM-247)

Executive Summary

Ongoing evolution of uninhabited systems and the interconnecting agent and cloud technologies reveal a rapid blurring of boundaries between the roles of humans and machines as well as the difference between the notions of uninhabited systems and other automated processes. In the future we need to focus less on supervising uninhabited systems and more on the multi-faceted, collaborative and reciprocal interaction between human and automated processes. Effective human-automation teaming technologies will enable autonomous systems to operate robustly in hazardous environments, support the tempo of rapid operations, or operate around and interact with humans, manned platforms, and other autonomous systems in highly constrained and chaotic environments.

Given the evolving roles of human and autonomous agents, NATO should strive toward flexible, adjustable and trustworthy automation that is being operated under user sovereignty. Research Task Group (RTG) HFM-247 explored the rapidly developing area of human-autonomy teaming. This RTG consisted of 20 dedicated scientists representing seven NATO nations. HFM-247 identified and demonstrated successful teaming methodologies and interface design practices that allow for shared situation awareness of task and environment, bi-directional understanding of intent, dynamic work distribution, and effective human-autonomy mission collaboration. The ultimate goal of this research activity is to increase the effectiveness of NATO operations through flexible and robust human-autonomy teams.

HFM-247 took a structured approach to studying Human-Autonomy Teaming (HAT) that allowed for creative explorations in select areas. The group began by establishing an initial list of human-autonomy teaming research challenges which was revisited and refined at the end of the effort, based upon knowledge gained over the lifespan of the group. The group first reviewed the literature of Human-Human Teaming (HHT) to identify lessons learned, guidelines, and metrics that might be applicable to HAT. This effort proved very informative for understanding the factors that most influence team dynamics. This Human-Human Teaming investigation also allowed the group to identify how HAT is different with unique challenges that might arise.

In order to stimulate enhanced interactions and collaboration within the group, each participating nation identified at least one HAT-related Technical Activity (TA) being conducted within that nation which served as that team member's 'contribution' to the group. These TAs represented significant ongoing HAT-related research or analyses that were tracked by the TG over the group's lifespan. A total of 10 TAs were identified across the seven nations and mapped to the HAT research challenges in order to identify similar research objectives and to inspire potential collaborations. This report summarizes these activities with the key accomplishments and lessons learned.

This RTG also addressed the very real difficulties associated with generalizing HAT solutions from one application to another by conducting a novel exploration of HAT design patterns. The application of design patterns has proven to be successful in the human-computer interaction and software engineering domains, so the group examined the possibility of leveraging this concept for use in HAT applications, i.e., to provide a common structure and vocabulary for relating each other's HAT artifacts. This report documents the

standardization of nomenclature describing highly automated human-machine systems along with top-level reoccurring problems faced in HAT applications, and then proposes a few candidate HAT solutions that might serve as design pattern instantiations. This promising effort requires more research before a final judgement can be made as to the utility of this approach. However, this report describes the initial motivation and approach taken as well as some candidate patterns for continued and more rigorous assessment.

HFM-247 successfully addressed a variety of important issues such as agreement on a common system-level framework pattern approach, HAT performance metrics, and possible solutions focused on dialog management, working agreements, transparency, and explicit intent communication. Through the TAs, the RTG also concentrated on the identification and demonstration of successful HAT applications, methodologies and design practices for HAT. However, this area is complex, and there are still significant challenges to overcome before true HAT can be achieved across NATO applications. As discussed further in the report, the following areas of research are suggested for further focus and development:

- 1) Meaningful human control: how to establish and maintain across all AI systems;
- 2) Team design patterns for dynamic evolving behaviors;
- 3) Continuous trust-calibration for proper reliance on automation;
- 4) Scope enlargement to cover all relevant teaming structures and characteristics;
- 5) Explainable AI in human-agent teamwork; and
- 6) Evolving hybrid intelligence by co-learning.

Équipe humain-autonomie : soutien d'une collaboration ajustable (STO-TR-HFM-247)

Synthèse

L'évolution actuelle des systèmes sans pilote et les technologies de cloud et d'agent d'interconnexion brouillent rapidement, d'une part, la frontière entre le rôle de l'homme et celui de la machine et d'autre part, la différence entre la notion de systèmes sans pilote et celle d'autres processus automatisés. À l'avenir, nous devons nous concentrer moins sur la supervision des systèmes sans pilote et plus sur les interactions multiformes, collaboratives et réciproques entre l'humain et les processus automatisés. Les technologies constituant des équipes réelles entre humains et automates permettront aux systèmes autonomes de fonctionner de manière robuste dans les environnements dangereux, de soutenir le rythme des opérations rapides ou de fonctionner autour et d'interagir avec les humains, les plateformes avec pilote et les autres systèmes autonomes dans des environnements chaotiques à fortes contraintes.

Étant donné l'évolution du rôle des humains et des agents autonomes, l'OTAN devrait s'efforcer de parvenir à une automatisation souple, ajustable et digne de confiance, exploitée souverainement par l'utilisateur. Le groupe de recherche (RTG) HFM-247 a étudié le domaine en plein essor des équipes humain-autonomie. Ce RTG se composait de 20 scientifiques spécialisés représentant sept pays de l'OTAN. Le HFM-247 a identifié et fait la démonstration de méthodologies d'équipe et de pratiques de conception d'interface couronnées de succès, qui permettent une connaissance partagée de la tâche et de l'environnement, une compréhension bidirectionnelle de l'intention, une répartition dynamique du travail et une collaboration humain-autonomie efficace en mission. Le but ultime de cette activité de recherche est d'accroître l'efficacité des opérations de l'OTAN grâce à des équipes humain-autonomie souples et robustes.

Le HFM-247 a adopté une démarche structurée de l'étude des équipes humain-autonomie (HAT, Human-Autonomy Teaming), qui a permis des explorations créatives dans certains domaines. Le groupe a commencé par établir une liste initiale de défis de recherche sur les équipes humain-autonomie, qui a été révisée et affinée à la fin, à partir des connaissances acquises pendant les travaux du groupe. Le groupe a tout d'abord passé en revue la littérature sur les équipes humain-humain (HHT, Human-Human Teaming) pour en identifier les enseignements, les principes directeurs et les indicateurs applicables aux HAT. Ce travail s'est révélé très utile pour comprendre les facteurs qui influencent le plus la dynamique d'équipe. Cette enquête sur les équipes humain-humain a également permis au groupe de déterminer la spécificité des HAT et les questions uniques susceptibles de se poser.

Afin de stimuler les interactions et la collaboration au sein du groupe, chaque pays participant a identifié au moins une activité technique (TA) liée aux HAT menée en son sein, qui a constitué la « contribution » de ce pays au groupe. Ces TA étaient des recherches en cours ou des analyses liées aux HAT et le RTG les a suivies tout au long de sa durée de vie. Au total, dix TA ont été identifiées parmi les sept pays et mises en correspondance avec les défis de la recherche sur les HAT, afin d'identifier les objectifs de recherche similaires et d'inspirer de potentielles collaborations. Le présent rapport résume ces activités, avec les résultats essentiels et les enseignements.

Ce RTG a également traité des difficultés bien réelles associées à la généralisation des solutions HAT d'une application à l'autre, en étudiant de manière inédite les schémas de conception des HAT. L'application

de schémas de conception s'étant révélée efficace dans les domaines de l'interaction humain-ordinateur et de l'ingénierie des logiciels, le groupe a examiné la possibilité d'exploiter ce concept dans les applications HAT, autrement dit, pour fournir une structure et un vocabulaire communs servant à relier entre eux les artefacts HAT des partenaires. Le présent rapport documente la normalisation de la nomenclature décrivant des systèmes humain-machine extrêmement automatisés, ainsi que les problèmes récurrents de haut niveau rencontrés dans les applications HAT, puis propose quelques solutions HAT candidates pouvant servir d'instanciations de schéma de conception. Ces travaux prometteurs nécessitent plus de recherches avant que leur utilité puisse être établie. Cependant, le présent rapport décrit la motivation initiale et l'approche adoptée, ainsi que quelques schémas candidats pour une évaluation continue et plus rigoureuse.

Le HFM-247 a traité avec succès diverses questions importantes, telles que l'accord sur une approche commune du schéma-cadre au niveau du système, les indicateurs de performance HAT et les solutions possibles axées sur la gestion du dialogue, les accords de travail, la transparence et la communication explicite de l'intention. À travers les TA, le RTG s'est également concentré sur l'identification et la démonstration d'applications HAT, de méthodologies et de pratiques de conception pour les HAT. Toutefois, ce domaine est complexe et il reste de grands obstacles à surmonter avant de parvenir à une véritable HAT dans les applications de l'OTAN. Comme en discute le rapport, nous suggérons les domaines de recherche suivants :

- 1) Contrôle humain sensé : comment l'établir et le maintenir dans tous les systèmes d'intelligence artificielle ;
- 2) Schémas de conception des équipes pour l'évolution dynamique des comportements ;
- 3) Étalonnage continu de la confiance pour une confiance adéquate dans l'automatisation ;
- 4) Élargissement du champ d'application à toutes les structures et caractéristiques pertinentes des équipes ;
- 5) IA explicable dans le travail d'équipe humain-agent ;
- 6) Évolution de l'intelligence hybride par le co-apprentissage.