Towards Explainability for a Civilian UAV Fleet Management using an Agent-based Approach

Yazan Mualla
CIAD, Univ. Bourgogne
Franche-Comté, UTBM,
90010 Belfort, France
yazan.mualla@utbm.fr

Amro Najjar
AI-Robolab/ICR, Computer
Science and Communications,
University of Luxembourg,
4365 Esch-sur-Alzette,
Luxembourg

Timotheus Kampik
Department of Computing
Science, Umeå University,
90736 Umeå, Sweden

Igor Haman Tchappi
University of Ngaoundere
454 Ngaoundere, Cameroon

Stéphane Galland
CIAD, Univ. Bourgogne
Franche-Comté, UTBM,
 90010 Belfort, France

Christophe Nicolle
CIAD, Univ. Bourgogne
Franche-Comté, UB,
21000 Dijon, France

Abstract
This paper presents an initial design concept and specification of a civilian Unmanned Aerial Vehicle (UAV) management simulation system that focuses on explainability for the human-in-the-loop control of semi-autonomous UAVs. The goal of the system is to facilitate the operator intervention in critical scenarios (e.g. avoid safety issues or financial risks). Explainability is supported via user-friendly abstractions on Belief-Desire-Intention agents. To evaluate the effectiveness of the system, a human-computer interaction study is proposed.

Author Keywords
Intelligent Transport Systems, Explainability, Multiagent Systems

Introduction
With the rapid increase of the world's urban population, the infrastructure of the constantly expanding metropolitan areas is subject to immense pressure. To meet the growing demand for sustainable urban environments and improve the quality of life for citizens, municipalities will increasingly rely on novel transport solutions. In particular, Unmanned Aerial Vehicles (UAVs), commonly known as drones, are expected to have a crucial role in future smart cities thanks to relevant features such as autonomy, flexibility, mobility, adaptive altitude, and small dimensions [18]. Therefore, over the past

few years, an increasing number of public and private research laboratories have been working on civilian, small, and human-friendly drones.

Still, several concerns exist regarding the possible consequences of introducing UAVs in crowded urban areas, especially regarding people's safety, e.g. if a mechanical failure causes a crash. To guarantee it is safe for UAVs to fly close to human crowds and to reduce costs, different scenarios must be modeled and tested. Yet, regulations restrict the use of UAVs in cities. Additionally, to perform tests with real UAVs, one needs access to expensive hardware. Moreover, field tests usually consume a considerable amount of time and require trained and skilled people to pilot and maintain the UAVs. Furthermore, on the field, it may also be hard to reproduce the same scenario several times [14]. In this context, the development of simulation frameworks that allow transferring real world scenarios into executable models using computer simulation frameworks, i.e. simulating UAVs activities in a digital environment are highly relevant [17], [15]. However, simulation frameworks have their own drawbacks, i.e. it is impossible to fully reproduce the real environment.

Intelligent software agents have been established as a suitable technique for implementing autonomous control and decision making in computer systems [22]. It is also used to for semi-autonomous systems in different domains [16]. The use of Agent-Based Simulation (ABS) frameworks in UAVs is gaining more interest in complex civilian application scenarios where coordination and cooperation are necessary [15]. ABS models a set of interacting intelligent entities that reflect, within an artificial environment, the relationships in the real world [22]. Due to operational costs, safety concerns, and legal regulations, ABS is commonly used to implement models and conduct tests. This has resulted in a range of research works addressing ABS in UAVs [17]. The results

make ABS a natural step forward towards better understanding and managing the complexity of today's business and social systems.

However, as UAVs are like any other robot, communication with humans is a challenge, since the human user is not by default capable of understanding the robot's State of Mind (SoM) [11].This problem is even more accentuate in the case of UAV since–as has been confirmed by recent studies in the literature [4, 10]–remote robots tend to instill less trust than those co-located. For this reason, working with remote robots is more challenging task specially in high-stakes scenarios such as flying UAVs in the urban environment. To overcome this challenge, this paper relies on the recent advances of the domain of eXplainable Artificial Intelligence (XAI) [8], [2], [6] in order to trace the decisions of the agents, and enable the validation of their behaviors when they are applied in a fleet of civilian UAVs that are interacting with other objects in the air or in the smart city. In this paper, we present a conceptual design of a MultiAgent System (MAS) that simulate the civilian UAVs' fleet, and provides tools for building the explainability of the system.

The rest of this paper is structured as follows. Firstly an overview of related works is given. Then, the explainable MAS for aerial transportation is highlighted and a use case in a smart city is presented. Finally, a conclusion and future research directions are presented.

Related Work

The rise of XAI

In the last couple of years, work on XAI is gaining momentum both in research and industry. Primarily, this surge can be explained by the success of black-box machine learning mechanisms whose inner workings are incomprehensible by human users [9]. Therefore, XAI aims to "open" the black-

box and explain the sometimes intriguing results of its mechanisms e.g. a Deep Neural Network (DNN) mistakenly classifying a tomato as a dog [21]. In contrast to this data-driven explainability [2], more recently, this tendency has been extended to explain the complex behavior of goal-driven systems such as robots/agents [2][11] since: *(i)* as has been shown in the literature, humans tend to assume that these agents & robots have their own SoM [11] and that with the absence of a proper explanation, the user will come up with an explanation that might be flawed or erroneous, *(ii)* these agents/robots are expected to be omnipresent in the daily lives of their users (e.g. social assistive robots and virtual assistants).

Explainability of UAVs

Recently, both data-driven and goal-driven explainability have been introduced in UAVs. In the former case, XAI aims to *interpret* the opaque machine learning mechanisms used by those UAV to analyze the input originating from their rich data streams. One ongoing work investigates how to interpret the decisions of convolution neural networks analyzing aviation related images as inputs [7]. Explainability is explored by reviewing three feature visualization methods in a layer-by-layer approach. In the latter case, explainability aims to make the autonomous behavior of the UAV understandable. One work is based on fuzzy logic: the explainable model is presented on a visual platform in the format of if-then rules derived from the fuzzy inference model [13]. Our approach belongs to the goal-driven case and is different than other related work as it relies on a decentralized solution using MAS. This choice is supported by the fact that the management of a UAV fleet must consider the physical distance between UAVs and other actors in the system. Additionally, autonomous agents represent in our opinion an adequate implementation of the autonomy of UAVs. The choice of Belief-Desire-Intention (BDI) [5] model is to support the explainability in the UAV fleet as detailed in the next section.

Explainable MAS for Aerial Transportation

Agent architectures, like BDI model, are frequently applied to equip UAVs with greater autonomy. By designing proactive agents that control UAVs, the latter become capable of autonomously managing their actions and behavior to reach their goals [3, 20].

As shown by Padgham et al. [19], BDI models can facilitate explainable agency, as they provide a well-structured snapshot of agent internals at any point of time. The BDI paradigm is frequently applied in ABS. Adam and Gaudou [1] present an extensive analysis and evaluation of approaches for integrating BDI models in ABS and highlight the previously mentioned benefits of BDI models as a way to implement descriptive agents that use richer and–from a human perspective– better interpretable abstractions than purely reactive agents.

Figure 1 outlines the contextual model of our work; in the bottom left, we see the technical requirements and infrastructure in a smart city, into which the aerial transportation management system (middle left) is embedded. In the top, the reactive agents guarantee autonomy when performing the normal tasks of an intelligent transport system, while there is a need for explainability that is associated with BDI agents and allows for human-in-the-loop control (top right). The UAVs in the fleet explain to the human their autonomous behavior and decisions along with any deviation from the planned mission. Table 1 shows a potential example of a high-level perspective of an agent's BDI snapshot.
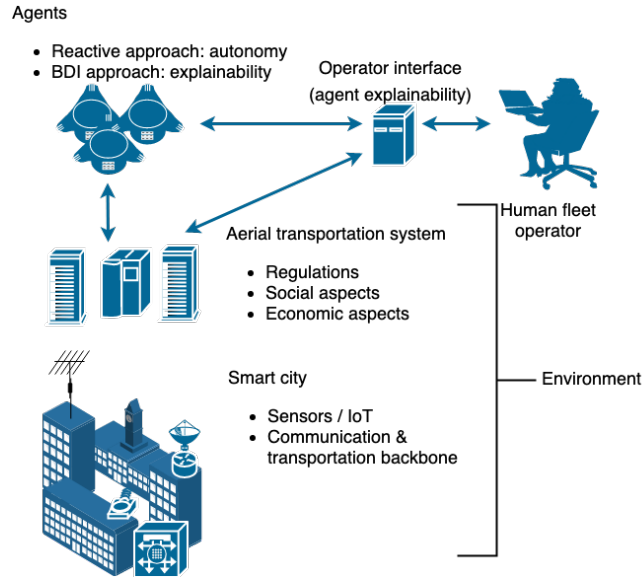
Agents

- Reactive approach: autonomy
- BDI approach: explainability

Operator interface
(agent explainability)

Human fleet
operator

Aerial transportation system

- Regulations
- Social aspects
- Economic aspects

Smart city

- Sensors / IoT
- Communication &
  transportation backbone

Environment

Figure 1: Contextual Model

Table 1: A high-level perspective of an agent's BDI snapshot.

Beliefs	`package_1: { to: Townhouse_27a}`
	
	`package_2: { to: School_1 }`
	`Townhouse_27a:`
	`{ time: 20min, charger: no }`
	`School_1:`
	`{ time: 20min, charger: yes }`
	`batteryTime: 22min`

Desires	`deliver: package_1`
	
	`deliver: package_2`

Intentions	`deliver: package_2`
Plans	`moveTo(School_1)`

Figure 2 shows the interaction between the actors in the proposed use case. In the following, the steps of the use case are detailed:

1. When a client puts a request for transporting a package/passenger, a notification is sent to the OAA.

2. The OAA will send it to all drones, so all drones are connected with each other and with the OAA using an assumed reliable network.

3. Drones that are near, with a specific radius, to the package/passenger will coordinate to complete the transportation mission. The decentralized coordination (without the approval of the operator) can be for several reasons:

   - Who will deliver the package/passenger according to constraints: actual distance to the package/passenger, battery size, having other packages/passengers in hand, having a mission with a near destination, etc.

DroneAgent-Delivery: a use case of human or package transportation in a smart city

The use case is about investigating the role of XAI in the communication between drones and humans in the context of human or package transportation in a smart city.
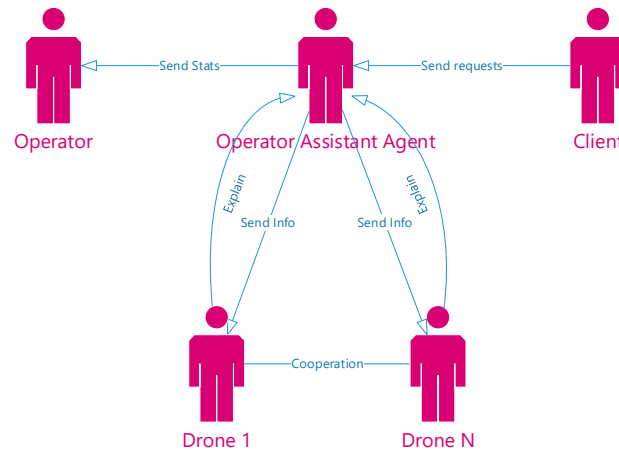
In the scenario, one operator is in charge of several drones that will provide transportation services to clients. These drones will autonomously conduct tasks and take decisions when needed. Additionally they need to communicate and discuss with each other and may cooperate to complete a specific task. The drones will explain to the Operator Assistant Agent (OAA) the progress of the mission including the unexpected events along with the decisions made by them.

Figure 2: Interaction of Actors in the System

- The load is heavy, and it needs several drones working as a swarm to lift it.

- There is a need to cooperate to deliver the package/passenger between several drones, where each drone delivers the load part of the way and then hands it to another drone.

4. Every drone will explain its own arguments to the OAA. At the end, the result of the coordination discussion will be sent to the OAA that will show it with/without filtration to the operator.

5. If the package/passenger is picked up by a drone from a competitor (external events), we have two situations:

   - The client sends to the OAA that the package or the passenger is picked up, and the assistant agent will inform the assigned drone to stop the mission;

   - The client does not send a notification that the package/passenger is picked up (because of self-ishness or laziness). In this situation, the drone will go to the place and observe the absence of the package/passenger, and it needs to explain this to the OAA.

6. The explanation needed from the drone is generally about the mission progress, its decisions and its status, e.g. the drone needs charging and that is why it ignores a nearby package/passenger, etc. Other kinds of explanation needed from the drones are the unexpected events, e.g. the drone arrives at the package location and see that is it damaged, or not according to the description (maybe heavier).

7. the OAA may filter the explanations received from the drones to give a summary of the most important explanations to avoid overwhelming the operator with a lot of details.

8. The operator at any time can either look at full explanations from all drones or only the results filtered by the OAA.

Evaluation of the autonomy and explainability
For evaluation, there is a need for an ABS to simulate an application of drones' autonomy and explainability. The evaluation is performed as a human-computer interaction study. The participants will try the simulation and fill out a questionnaire to discover if a human user (as the operator) can understand the explanations provided by the UAVs. The questionnaire will be built specifically according to the XAI metrics provided in the literature [12].

The evaluation will have three levels:

- First level: we test the explainability aspect. The experiment requires dividing the participants into three groups. The first, second and third groups will try the simulation without any explanation, with full explanation, and with filtered explanation, respectively.

- Second level: we test the shared autonomy aspect. We ask the participants to perform a tedious task like filling an excel sheet while the simulation is running in two modes (First mode: The operator only looks at the full explanations by all drones. Second mode: The operator is looking at only the filtered explanations from the OAA). Then we measure in what mode the participants completed more percentage of the tedious task assigned to them.

- Third level: we get the results and statistics from the simulation, e.g. how many packages delivered, if a drone did not deliver any package, the distance moved by all drones, etc. The operator benefits from these results to change the initial parameters of the simulation for the second run and check if the results are better.

Conclusion and future work

This paper presented a concept and a specification for an agent-based civilian UAV fleet management approach with a focus on explainability. The presented work is an initial step towards the goal of providing agent-based tools that allow for a human-in-the-loop approach supporting semi-autonomous UAV fleet management. In particular, the following work is of importance:

Engineering research should be conducted to design an explainable multiagent management system whose architecture is sufficiently generic to be applied for use cases beyond UAV fleet management.

An empirical assessment should evaluate the effectiveness of the approach in human-computer interaction studies.

Acknowledgements

REFERENCES

1. Carole Adam and Benoit Gaudou. 2016. BDI agents in social simulations: a survey. *The Knowledge Engineering Review* 31, 3 (2016), 207–238.

2. Sule Anjomshoae, Amro Najjar, Davide Calvaresi, and Kary Främling. 2019. Explainable agents and robots: Results from a systematic literature review. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, 1078–1088.

3. Willson Amalraj Arokiasami, Prahlad Vadakkepat, Kay Chen Tan, and Dipti Srinivasan. 2016. Interoperable multi-agent framework for unmanned aerial/ground vehicles: towards robot autonomy. *Complex & Intelligent Systems* 2, 1 (2016), 45–59.

4. Wilma A Bainbridge, Justin Hart, Elizabeth S Kim, and Brian Scassellati. 2008. The effect of presence on human-robot interaction. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 701–706.

5. Michael Bratman. 1987. *Intention, Plans, and Practical Reason*. Center for the Study of Language and Information.

6. Davide Calvaresi, Yazan Mualla, Amro Najjar, Stéphane Galland, and Michael Schumacher. 2019. Explainable Multi-Agent Systems through Blockchain Technology. In *Proceedings of the 1st International Workshop on eXplanable*

TRansparent Autonomous Agents and Multi-Agent Systems (EXTRAAMAS 2019).

7. Chester V Dolph, Loc Tran, and Bonnie D Allen. 2018. Towards Explainability of UAV-Based Convolutional Neural Networks for Object Classification. In *2018 Aviation Technology, Integration, and Operations Conference.* 4011.

8. Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, Franco Turini, Fosca Giannotti, and Dino Pedreschi. 2019. A survey of methods for explaining black box models. *ACM computing surveys (CSUR)* 51, 5 (2019), 93.

9. David Gunning. 2017. Explainable artificial intelligence (xai). *Defense Advanced Research Projects Agency (DARPA), nd Web* (2017).

10. Helen Hastie, Xingkun Liu, and Pedro Patron. 2017. Trust triggers for multimodal command and control interfaces. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction.* ACM, 261–268.

11. Thomas Hellström and Suna Bensch. 2018. Understandable robots-what, why, and how. *Paladyn, Journal of Behavioral Robotics* 9, 1 (2018), 110–123.

12. Robert R Hoffman, Shane T Mueller, Gary Klein, and Jordan Litman. 2018. Metrics for explainable AI: Challenges and prospects. *arXiv preprint arXiv:1812.04608* (2018).

13. Blen M Keneni, Devinder Kaur, Ali Al Bataineh, Vijaya K Devabhaktuni, Ahmad Y Javaid, Jack D Zaientz, and Robert P Marinier. 2019. Evolving Rule-Based Explainable Artificial Intelligence for Unmanned Aerial Vehicles. *IEEE Access* 7 (2019), 17001–17016.

14. Fabian Lorig, Nils Dammenhayn, David-Johannes Müller, and Ingo J Timm. 2015. Measuring and Comparing Scalability of Agent-Based Simulation Frameworks. In *German Conference on Multiagent System Technologies.* Springer, 42–60.

15. Yazan Mualla, Wenshuai Bai, Stéphane Galland, and Christophe Nicolle. 2018. Comparison of Agent-based Simulation Frameworks for Unmanned Aerial Transportation Applications. *Procedia computer science* 130, C (2018), 791–796.

16. Yazan Mualla, Amro Najjar, Olivier Boissier, Stéphane Galland, Igor Tchappi Haman, and Robin Vanet. 2019a. A Cyber-Physical System for Semi-Autonomous Oil&Gas Drilling Operations. In *2019 Third IEEE International Conference on Robotic Computing (IRC).* IEEE, 514–519.

17. Yazan Mualla, Amro Najjar, Alaa Daoud, Stéphane Galland, Christophe Nicolle, Elhadi Shakshuki, and others. 2019b. Agent-based simulation of unmanned aerial vehicles in civilian applications: A systematic literature review and research directions. *Future Generation Computer Systems* 100 (2019), 344–364.

18. Yazan Mualla, Amro Najjar, Stéphane Galland, Christophe Nicolle, Igor Haman Tchappi, Ansar-Ul-Haque Yasar, and Kary Främling. 2019c. Between the megalopolis and the deep blue sky: Challenges of transport with uavs in future smart cities. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems.* International Foundation for Autonomous Agents and Multiagent Systems, 1649–1653.

19. Lin Padgham, David Scerri, Gaya Buddhinath Jayatilleke, and Sarah L. Hickmott. 2011. Integrating BDI reasoning into agent based modeling and simulation. *Proceedings of the 2011 Winter Simulation Conference (WSC)* (2011), 345–356.

20. Domenico Pascarella, Salvatore Venticinque, and Rocco Aversa. 2013. Agent-based design for UAV mission planning. In *8th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC).* IEEE, 76–83.

21. Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, and Rob Fergus. 2013. Intriguing properties of neural networks. *arXiv preprint arXiv:1312.6199* (2013).

22. Michael Wooldridge and Nicholas R Jennings. 1995. Intelligent agents: Theory and practice. *The knowledge engineering review* 10, 2 (1995), 115–152.