

HUMAN-SYSTEM INTERFACE WITH EXPLANATION OF ACTIONS FOR AUTONOMOUS ANTI-UAV SYSTEMS

John Kontos

Department of Cognitive Science,
National and Kapodistrian University of Athens, Athens, Greece

ABSTRACT

Research on explanation is currently of intense interest as documented in the DARPA 2021 investments reported by the USA Department of Defense. An emerging theme for explanation techniques research is their application to the improvement of human-system interfaces for autonomous anti-drone or C-UAV defense systems. In the present paper a novel proposal based on natural language processing technology concerning explanatory discourse using relations is briefly described. The proposal is based on the use of relations pertaining to the possible malicious actions of an intruding alien drone swarm and the defense decisions proposed by an autonomous anti-drone system. The aim of such an interface is to facilitate the supervision that a user must exercise on an autonomous defense system in order to minimize the risk of wrong mitigation actions and unnecessary spending of ammunition.

KEYWORDS

UAV, anti-drone, explanation, human-system interface

1. INTRODUCTION

As stated in [1] “UAS operations are complex symphonies of activities of many operators, both automated and humans. Understanding and analyzing interfaces can provide the C-UAS mission planner with many opportunities for vulnerability exploitation”.

In the present paper a novel proposal for a natural language based interface concerning explanatory discourse aided by relations defined below is briefly described. The aim of such an interface is to facilitate the supervision that a user must exercise on an autonomous defense system in order to minimize the risk of wrong mitigation actions and unnecessary spending of ammunition. Such an interface is an augmentation of the usual image-only man-system interfaces that in the case of a drone swarm intrusion may be so much cluttered that render the prompt reaction of the human supervisor very difficult.

The proposal concerns autonomous but accountable anti drone systems that have the task of dealing with intruding alien swarms of drones. The security of the defense operation of autonomous systems is of major concern because lethal decisions must ideally be supervised by a human.

This human will take the responsibility for the decisions to act against alien drones. The XAI part of the user-C-UAS (or anti-drone system) interface must be able to participate in a discourse concerning its operation and its decisions in order to support any visual display presented to the user of the C-UAS.

This discourse is mainly based on pictorial data depicting the motion of objects in space. By “discourse” we mean the exchange of questions and answers and proposals for mitigating actions between a supervising user and an autonomous defense system.

As reported in [2] DARPA is scheduling investments for the next years on several projects that include explanation among their goals. A recent review of existing anti-UAV technology is given in [3]. In the following the proposal mentioned above is briefly described.

2. RHETORIC RELATIONS FOR EXPLANATION

Rhetoric relation methods have been used for improving argument mining [4]. Similar methods for the generation of user friendly natural language explanations could be used as proposed in [5].

In the case of explanations of the action either taken by a real autonomous system or a simulated one it is proposed here to use an enhanced set of relations in the lines of the method described in [5]. This is necessary because the discourse concerning the actions of a moving autonomous system to be intercepted involves the reference to dynamic pictorial information.

This information depicts the state of affairs of the system and its environment that must be translated to natural language sentences that a human user may easily and quickly understand. These sentences must inform the user for the action proposal of the anti-drone system and the justification or explanation of its decision.

In this way the user is supported in his task of deciding how to deal with the battle situation while she is observing a screen cluttered with plethora of data and traces of many tens of intruding drones and their possible targets that are difficult to manage.

The relations defined here can serve as a formal representation of the qualitative aspects of the behavior of the systems involved. They are deduced from the data collected by the sensor subsystems of the anti-drone system. Using these relations formal reasoning may be executed prior to generating the natural language text presented to the user asking for her approval of actions to be performed.

3. TEXT AND IMAGE RELATIONS

In [5] a set of relations between text and still images analogous to but not included in the classical set of intra-text rhetorical relations was proposed. The classical set of rhetorical relations refer to relations between sentences or sets of sentences of a natural language text.

As stated in [4]: “Many forms of argument were identified and discussed by Aristotle. Argument schemes are patterns of human reasoning used for both inventing and evaluating arguments and can be seen as historical descendants of Aristotle’s Topics. ... argument schemes are forms of inference from premises to a conclusion of the kind used in everyday conversational exchanges in which one party is trying to get another to accept a conclusion at issue. ... Schemas in Rhetorical Structure Theory (RST), on the other hand, define the structural constituency arrangements of text. As a descriptive framework for text, RST provides features that can be useful in discourse studies. Representation of communicative roles of a text structure via schemas certainly aids in discourse representation for argument mining, ... RST defines twenty-three relations that can hold between spans of a text...From a theoretical point of view Azar ... investigates five RST relations Evidence, Motivation, Justify, Antithesis and Concession”.

The relations proposed in [5] generalize the intra-text relations of the RST mentioned above to relations between text spans and images. Examples of text-image relations may be defined for instance in cases such as

1. A sentence describes the whole or part of an image
2. The content of an image supports the truth of a sentence
3. A sentence refers to the sequence of parts of a waveform depicted in an image.

A kind of relations of case 3. above (symbolized as mbr) was proposed in [5] and they formalize the text description of the time behavior of the variables of a dynamic system.

The sequence of number tuples of such system variables are treated as “numerical narratives” with a “rhetoric” structure of events corresponding to the “pattern structure” of graphs where their notable variations play the role of events.

Graphs of the time evolution of variables may show peaks and valleys that are the main features analogous to the events of a narrative.

4. SYSTEM NATURE RELATIONS (SNR)

In view of the generalization proposed in [5] the relations involved will be hence symbolized as “snr” standing for “system nature relations”. The formal representation of the relations proposed has the form:

$r(\text{PAR}_1, \dots, \text{PAR}_n)$ where the arguments PAR_1 to PAR_n stand for n parameters that define the relation between two “objects”. These “objects” may be of different kinds such as physical objects, commands, questions, explanations, equation variables and data collected by sensors. The particular formal representation used in the present paper being a special case of the above general form is:

$r(\text{relation_name}, \text{relation_kind}, \text{first_object_kind}, \text{first_object_identifier}, \text{second_object_kind}, \text{second_object_identifier})$

and is used for the formalization of relations useful for generating the discourse between a user and an autonomous anti-drone defense system. In some cases a single object may be sufficient.

The variable “Behrel” below stands for relations between some entity “C” of the system and the time behavior of a set of variables “V1, ... Vn” of the object involved that are expressed as verbs in the explanation generated.

Formally: “ $r(\text{Behrel}, \text{snr}, \text{drone}, \text{A}, \text{variables}, \text{V1}, \dots \text{Vn})$ ”.

In the above formula or predicate, as called in Prolog, the variable Behrel stands for a relation belonging to the relation category “snr”.

We give below some simple examples in order to demonstrate the use of such relations in the generation of explanations by a user-C-UAV interface.

The relational representation facilitates the automatic reasoning that may be needed in order to generate complex explanations of the actions proposed by the autonomous anti-drone system to its supervising user.

Such relational reasoning is easily programmed with AI programming languages like PROLOG as shown by the author and his team in [5], [6], [7] and [8].

5. SIMPLE EXAMPLES

The first simple example is based on the following scenario:

An intruding swarm of possibly armed drones has been identified by an autonomous anti-drone or C-UAV system S and all drones of the swarm are being tracked. S decides to attack a drone “a” of the intruding swarm that its defense algorithm has concluded that “a” threatens target “t” because it approached it nearer than a limit distance “d”.

S before attacking “a” asks its supervisor user the following question:

“Do you allow attack of “a” **because it approached “t” nearer than d**” ?

The final decision will be taken by the user by selecting Yes or No.

The bold part of the above question to the user is an “explanation” that must be generated by the XAI subsystem of the user-system interface of S.

This explanation uses the motion verb “approach” that S must have some knowledge of its semantics in order to use it appropriately.

The semantic grounding of motion verbs is defined in [6] but other categories of verbs such as causal or ordering verbs (e.g. “causes” and “follows”) may need to be generated using appropriate natural language processing techniques.

A brief account of the processing steps, using Prolog as example language, for handling the relation denoted by the verb “approach” is given below.

The coordinates of each drone recognized as member of the intruding alien swarm are dynamically recorded as a function of time. These records may be transformed into Prolog facts or entries in a Prolog-processable file in order to be used when running the Prolog program generating the discourse with the user that contains an explanation. These steps are:

1. The distance of each intruding alien drone from all targets defended by the anti-drone system are calculated in real time.
2. Whenever the distance of a possibly armed alien drone becomes equal or smaller than the distance defined as “safe” a relation of the form defined above is instantiated corresponding to a drone-target pair.
3. The instantiated relation after possibly having been logically related with other relevant relational knowledge is used for the generation of an alarm text easily understandable by the supervising human asking for her approval of the action decided by the autonomous defense system algorithm which contains an explanation.
4. The supervising human selects an answer that determines the action of the defense system.

A possible implementation of the basic steps only from the ones given above may be described in a Prolog-like style as follows:

r (approach,snr,entity,A,variables,X,Y,Z):-

distance (X,Y,Z,T, unsafe, D), drone(A),target(T),

print ("because it approached" T "nearer than" D).

The predicate "distance" is defined on the basis of the calculations that determine whether "a" is threatening "t".

The "drone" and "target" predicates will be facts dynamically inserted into the program on the basis of the sensor signals that the anti-drone defense system collects and combines in fused form.

The second simple example is based on the following scenario:

An intruding swarm of possibly armed drones has been identified by an autonomous C-UAV system S while all drones are being tracked. S decides to attack a drone "d1" of the intruding swarm that its defense algorithm has concluded that "d1" threatens target "t" because it is nearer than d2 to t.

S before attacking "d1" asks its supervisor user e.g. the following question:

"Do you allow attack of "d1" because it threatens "t" as it is nearer to t than d2 which also approaches t" ?

The final decision will be taken by the user by selecting Yes or No. The bold part of the above question to the user is an "explanation" that must be generated by the XAI subsystem of the user-system interface of S.

A brief account of some processing steps proposed using PROLOG as example programming language is given below.

It is presupposed that the coordinates of each recognized drone of the intruding alien swarm are dynamically recorded as parameters of PROLOG predicates so that they can be used by the reasoning process of an intelligent interface.

These steps are:

1. The distance of each intruding alien drone from all targets defended by the anti-drone system are calculated in real time.
2. Whenever the distance of a possibly armed alien drone becomes equal or smaller than the distance defined as "safe" a relation of the form defined above is instantiated corresponding to a drone-target pair.
3. Drones d1 and d2 both approach t but d1 is detected as larger than d2.
4. The instantiated relation after possibly having been logically related with other relevant relational knowledge is used for the generation of an alarm text easily understandable by the supervising human asking for her approval of the action decided by the autonomous defense system algorithm which contains an explanation.
5. The supervising human selects an answer that determines the action of the defense system.
- 6.

A possible implementation of the basic steps from the ones given above may be described in a Prolog-like style, using predicates already defined above, as follows:

r (threatens,snr,entity,D1,variables,X1,Y1,Z1):-

```
distance(X1,Y1,Z1,T,unsafe,D), distance(X2,Y2,Z2,T,unsafe,D),
drone (D1), drone(D2),
nearer (X1,Y1,Z1,X2,Y2,Z2,T), target(T),
print ("because it threatens", T, " as it is nearer to". T, "than", D2, "which also approaches",
T).
```

The third simple example is based on the following scenario:

An intruding swarm of possibly armed drones has been identified by an autonomous C-UAV system S.

All drones will be tracked automatically and the tuples of their position in space recorded in appropriate databases.

S decides to attack a drone "d1" of the intruding swarm that its defense algorithm has concluded that "d1" threatens target "t" because it is larger than d2 although both d1 and d2 approach t.

S before attacking "d1" asks its supervisor user e.g. the following question:

"Do you allow attack of "d1" because it threatens "t" as larger than d2 which also approaches t" ?

The final decision will be taken by the user by selecting Yes or No. The bold part of the above question to the user is an "explanation" that must be generated by the XAI subsystem of the user-system interface of S.

A brief account of the processing steps proposed are given below assuming that the coordinates of each recognized drone of the intruding alien swarm are dynamically recorded as above. These steps are:

1. The distance of each intruding alien drone from all targets defended by the anti-drone system are calculated in real time.
2. Whenever the distance of a possibly armed alien drone becomes equal or smaller than the distance defined as "safe" a relation of the form defined above is instantiated corresponding to a drone-target pair.
3. Drones d1 and d2 both approach t but d1 is detected as larger than d2.
4. The instantiated relation after possibly having been logically related with other relevant relational knowledge is used for the generation of an alarm text. This text should be easily understandable by the supervising human.
5. This text consists of a request for her approval of the action decided by the autonomous defense system algorithm and an explanation of the decision taken by the system.
6. The supervising human selects an answer that determines the action of the defense system.

A possible implementation of the basic steps from the ones given above may be described in a Prolog-like style, using predicates already defined above, as follows:

```
r (threatens,snr,entity,D1,variables,X1,Y1,Z1):-
distance (X1,Y1,Z1,T,unsafe,D), distance(X2,Y2,Z2,T,unsafe,D),
drone (D1), drone(D2), larger(D1,D2),target(T),
print ("because it threatens", T, "as larger than", D2 which also approaches T "
```

The presentation of a full program for the tasks executed by a real C-AUA in the situations described above is beyond the scope of the present paper.

6. CONCLUSIONS

In the present paper a novel proposal for the design of a friendly human-system interface of an autonomous anti-drone system was described. The proposal is based on the use of relations pertaining to the motion actions of a drone swarm system.

The aim of such an interface is to facilitate the supervision that a user must exercise on an autonomous defense system in order to minimize the risk of wrong mitigation actions and unnecessary spending of ammunition.

The relations used are convenient representations of information necessary for the justification of defense decisions of an autonomous anti-drone system.

The interface designed facilitates the supervision that a user must exercise on an autonomous defense system.

The usefulness of the interface is to maximize security of actions taken by the system and minimize the unnecessary spending of ammunition.

In the absence of such an intelligent interface the user of an anti-drone system must cope with a cluttered computer display depicting the dynamic behavior of tens or even hundreds of moving objects in the battle field and take decisions that require reasoning beyond the processing speed possible by a human brain.

The development of the proposal of the present paper can be based on methods of Artificial Consciousness and Natural language Generation. Such methods were presented in [5], [6], [7] and [8] by the author and his team offering enhancement of the friendliness of user human-system interfaces. Recent systems such as those described in [9],[10],[11], [12], [13] and [14] lack such functionalities and may benefit from the leveraging of the techniques described in the present paper making the task of the supervising user much safer and easier.

These functionalities are necessary when the invading swarm invading an area to be defended consists of a large number of drones. In such a case the display on a screen of the conflict situation may be so cluttered that constitutes the job of the supervising human quite difficult. The implementation of an interface of the kind proposed here may be improved using methods of XAI technology that keeps developing by recent research.

It is mandatory that the research and development in the area is supported with funds from funding organizations of countries that seek advances in technological development of anti-drone systems that enhance the performance of their autonomous anti-drone systems. This means that the engineering development of anti-drone systems must be complemented by activities in the areas of natural language generation and XAI. The present paper deals with the XAI aspects of a real-time system different than the usual off-line systems. [15]

The user-friendliness of interfaces of anti-drone systems is very important for facing invasions of large swarms of drones that unaided humans are unable to cope with due the clutter of traditional screen interfaces. The proposal of the present paper shows a direction that such research must follow for improving the human-system interface used.

REFERENCES

- [1] Mueller S. T. et al, "Explanation in Human-AI Systems: A Literature Meta-Review. Synopsis of Key Ideas and Publications and Bibliography for Explainable AI", DARPA XAI Literature Review, 2019.
- [2] Department of Defense, "Fiscal Year (FY) 2021 Budget Estimates", 2020.
- [3] Hao Q. et al, "Research on anti UAV swarm system in prevention of the important place", Journal of Physics: Conference Series, 1507, 2020.
- [4] Mitrovic J., "Ontological representations of rhetorical figures for argument mining", Argument & Computation, 8, 267–287, 2017.
- [5] Kontos J. et al, "Question Answering and Rhetoric Analysis of Biomedical Texts in the AROMA System". HERCMA 2005, Athens, 2005.
- [6] Kontos, J. and I. Malagardi, "Motion Verbs and Vision" HERCMA 2007, Athens, 2007.
- [7] Kontos J. et al, "Metagnostic Deductive Question Answering with Explanation from Texts", Universal Access in HCI, Part IV, HCI 2011, , pp. 72–80, Springer-Verlag Berlin Heidelberg, 2011.
- [8] Kontos J., "Machine Consciousness and Question Answering". American Journal of Data Mining and Knowledge Discovery, 1(1): 7-15, 2016.
- [9] SouliN. et al, "Horizon Block, Implementation of an Autonomous Counter-Drone System", International Conference on Unmanned Aircraft Systems (ICUAS), Athens, Greece. September 1-4, 2020.
- [10] Martín-Barrio A. et al, "Application of immersive technologies and natural language to hyper-redundant robot teleoperation" Springer Nature 2019.
- [11] Mendonça N. M. A., "Control of Robot Swarms through Natural Language Dialogue", M. Sc, 2019.
- [12] Kang H. et al, "Protect Your Sky: A Survey of Counter Unmanned Aerial Vehicle Systems", IEEE Access, September 2020.
- [13] Park S. et al, "Survey on Anti-Drone Systems: Components, Designs, and Challenges", IEEE Access, 2021, (in press).
- [14] Roldán-Gómez J. J, et al, "A Survey on Robotic Technologies for Forest Firefighting: Applying Drone Swarms to Improve Firefighters' Efficiency and Safety", Appl. Sci. 11, 363, 2021.
- [15] Demajo L. M. et al , "An Explanation Framework for Interpretable Credit Scoring" , International Journal of Artificial Intelligence and Applications (IJAIA), Vol.12, No.1, January 2021.