A learning model for RPAS sensor operators and its implications for training

Jan Joris Roessingh[1], Olaf Brouwer[1], Joost van Oijen[1], Gerald Poppinga[1]
[1]Netherlands Aerospace Centre NLR, Amsterdam, the Netherlands

Abstract — We developed a competency profile for the Sensor Operator (SO) of Remotely Piloted Aircraft Systems (RPAS) and subsequently developed a Reinforcement Learning (RL) agent that is capable to resolve tasks that are illustrative for an SO's job. We compare performance results of humans and Reinforcement Learning (RL) agents in performing these tasks. We analysed that there are similarities in the learning process of humans and RL agents. Experimental results potentially allow us to identify human task requirements, training needs, selection criteria and cut-off benchmarks from data generated by RL agents. We present tasks that cover different cognitive abilities required for an SO, using games as a method for learning. An RL agent provides insight in SO task performance, enables the identification of learning transfer between tasks and enables the development of effective training for SO's.

1 RPAS sensor operator tasks & competences

Large Remotely Piloted Aerial Systems (RPAS) provide the armed forces with advanced Intelligence, Surveillance and Reconnaissance (ISR) capabilities, to cope with complex operations. The flight crew, often consisting of a pilot and a sensor operator (SO), faces challenges that are typical for large RPAS operations. These include continuous mission durations without interruption, massive data gathering and operating in an only partially observable mission environment. This kind of environment requires new abilities and competencies.
Reported Human Factor Problems [1] with military long-endurance RPAS operations include non-adherence to procedures, suboptimal display design, decision errors, lack of alertness, perceptual errors and lack of teamwork. Our analysis [2] yields that the most critical SO tasks during Intelligence, Surveillance, and Reconnaissance (ISR) missions are:

- maintaining the operational picture,
- information management,
- employment of sensors and systems,
- in-flight duties and standard operating procedures,
- tasks related to operational safety, and
- mission coordination.

A longer list of human attributes needed to perform these SO tasks was assembled on the basis of literature analysis, of which the following items are the most important:

- Situational Awareness
- Cognitive task prioritization
- Adaptability / flexibility
- Cognitive proficiency
- Visual perception
- Attention
- Spatial processing
- Short- and long-term memory
- Reasoning
- Interpersonal skills

Finally, an analysis was conducted to establish training priorities. The following sub-tasks were found to have the highest training priority:

- Detection, identification and monitoring air, ground and sea units
- Gathering, storing and distributing information

2 Sensor Operator mini-tasks

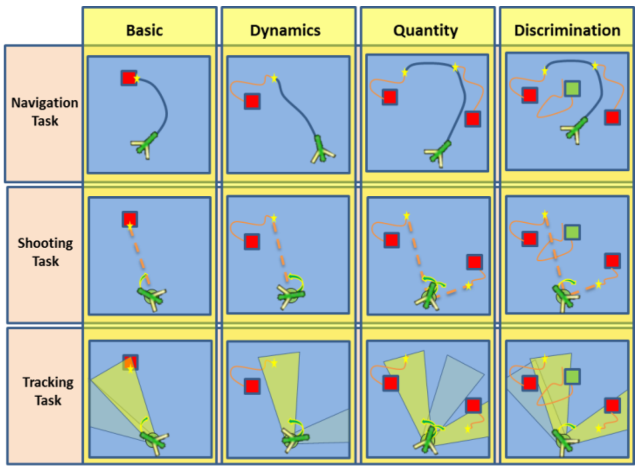

Fig. 1. Overview of the Sensor Operator mini-games

In order to devise an agent that could master SO-like tasks, we started with relatively simple machine learning algorithms and followed a bottom-up approach [3]. This involves breaking down the problem domain into so-called mini-games to investigate the performance of the Reinforcement Learning (RL) agent. The mini-games were designed to address one or more cognitive abilities (see Section 1) that would be required to master such games. Further game elements were incorporated that were illustrative for sub-tasks of an SO. At a later stage, separate tasks can be combined to represent more complex cognitive tasks.

This cumulative approach towards learning tasks is also seen in humans. For instance, in [4], the Space Fortress (SF) game was decomposed into separate sub-tasks, trained human subjects on the tasks and then verified the performance of the subjects on the overall game. In our ITEC presentation we will also address results collected from both human learning and RL agent learning on the SF game.

Concerning the experimental results on the SO-mini-games, which we earlier reported in [3], these indicate that training on the sub-games made the subjects perform better on the overall game. Fig. 1 provides an overview of the mini-games that have been proposed in order to investigate the performance of an RL agent on illustrative SO tasks and associated cognitive abilities. They can be categorized across two dimensions, namely *task-oriented* (navigation, shooting, tracking) and *worker-oriented* (basic, dynamic, quantity and discrimination).

Table 1. SO mini-games results

Navigation Task				
	Basic	*Dynamics*	*Quantity*	*Discrimination*
Agent score	**109%**	**91%**	**66%**	**22%**
Random score	3%	18%	30%	13%
Shooting Task				
	Basic	*Dynamics*	*Quantity*	*Discrimination*
Agent score	**120%**	**101%**	**76%**	**71%**
Random score	6%	21%	25%	16%
Tracking Task				
	Basic	*Dynamics*	*Quantity*	*Discrimination*
Agent score	**102%**	**102%**	**96%**	**92%**
Random score	18%	15%	21%	13%

When considering the results in Table 1, in the *basic* and *dynamics* variations the agent reaches super-human or (near) par-human performance in all tasks. The only significant super-human performance is seen in the basic shooting task in which the agent excels in reaction time and accuracy. In the *quantity* and *discrimination* variations, the agent performs sub-human and performance starts to degrade in relation to the complexity of the task. However, agent performance still increases at the end of the training time which suggests opportunities for an agent to improve when trained longer.

A difference in difficulty in learning a task is also clearly observed from the agent's performance for each task. In the navigation task, the lowest performance is seen, followed by the shooting task, followed by the tracking task. We believe this difference is due to the difficulty for an agent to obtain a reward for a task. In the navigation task, random actions in the early training phases rarely lead to rewards. In the tracking task, rewards are immediate and much easier to obtain.

In an attempt to decrease the training time for complex mini-games an additional experiment was performed. The goal of this experiment was to verify if training time of a complex variation could be reduced through progressive part-task training. Training results from a less complex variation were used as input for a more complex variation. Research has shown that such a strategy is beneficial for human learning, cf. [4]. We considered the navigation/discrimination task. Initial (non-part-task) training on this task reached 54% human performance after 75 million frames. Alternatively, this task was trained incrementally by first training the *dynamics* variation, followed by the *quantity* variation and concluding with the *discrimination* variation. Training time of part-tasks was divided equally (hereby reaching the same amount of training time). Results showed an increase of 20% of the score, resulting in 65% human performance. This suggests positive transfer of training, such that less overall training time is required when tasks are learned incrementally. Fig. 3. shows the learning progress of the *quantity* and *discrimination* variations with and without pre-training. The influence the specific ordering of variations has with respect to incremental part-task training has not been investigated and is left for future work.

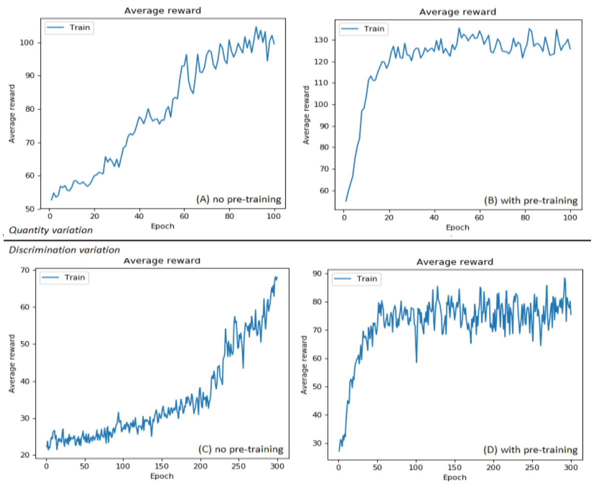

Fig. 2. Progressive part-task training results for the navigation task: comparing no pre-training (A,C) versus pre-training (B,D)

3 Conclusions and future work

We present RL agents in performing tasks which are illustrative for RPAS sensor operators. Gaming was used as a connection: on the one hand, games can be used by humans to enhance cognitive abilities that can be transferred to real-life tasks. On the other hand, recent developments in RL algorithms have shown promising results in their application to games. We (1) constructed tasks (games) that require some of the abilities of the SO, (2) devised RL agents that have to learn these tasks and (3) performed initial learning tests with RL agents on these tasks

References

[1] Brouwer, O., Roos, C., Roessingh, J.J.M. (2015). Flying 2020 - The Integrated Approach. NLR Technical Report NLR-TR-2015-292. Amsterdam, The Netherlands

[2] Brouwer, O. (2018). Flying 2020 – MALE RPAS Sensor Operator Tasks and Competencies. NLR-TR-2018-214. Amsterdam, The Netherlands

[3] Oijen, J. van, Poppinga, G., Brouwer, O., Aliko, A., Roessingh, J.J.M. (2017): Towards Modeling the Learning Process of Aviators using Deep Reinforcement Learning. 2017 IEEE International Conference on Systems, Man, and Cybernetics

(SMC), Banff Center, Banff, Canada, October 5-8, 2017; 10/2017
[4] Frederiksen, J. R., White, B. Y., "An approach to training based upon principled task decomposition," Acta Psychologica, vol. 71, pp. 89-146, 1989.

Author/Speaker Biographies

Jan Joris Roessingh holds a Ph.D. in Physics from Utrecht University in the Netherlands and is a simulation and training expert at the Netherlands Aerospace Centre NLR.

Olaf Brouwer holds an M.Sc. in Technology Management from Eindhoven Technical University in the Netherlands and is a Human Factors consultant at the Netherlands Aerospace Centre NLR.

Joost van Oijen holds a Ph.D. in Computer Science from Utrecht University in the Netherlands and is a Simulation and AI Expert expert at the Netherlands Aerospace Centre NLR.

Gerald Poppinga holds an M.Sc. from Groningen University in the Netherlands and is a Military Operations and AI Expert at the Netherlands Aerospace Centre NLR.