Autonomy Degradation Workspace - A Design Concept for Human-Automation Cooperation

Magnus Nylin, Jonas Lundberg, and Jimmy Johansson Westberg

Abstract—Lack of support for handling a degradation in autonomy in a highly autonomous automation may lead to a stressful situation for a human when being forced to take-over when a degradation occurs. In this article we present a design concept, the *Autonomy Degradation Workspace*, to address this issue. The starting point is that the human and the automation work together in parallel control processes, but at different levels of cognitive control, for example plans versus goals. When autonomy is degraded, the automation shall consult the human by providing information transformed to suit the level of which the human is working, and the timing of it shall be adapted to suit the human's working situation. This is made possible by letting the automation monitor the human in a third, separate process. This process together with the parallel control processes, the information level transformation, and timing of the presentation are the key characteristics of the *Autonomy Degradation Workspace*, which can be described in four steps: Identification of the need; evaluation of if, when, and how to present information; perception and response by the human; implementation of a solution by the automation. The timing of the information presentation shall be adapted in real-time to provide flexibility, while the level of the information provided shall be tuned off-line and kept constant to provide predictability. By following the *Autonomy Degradation Workspace* concept, the risk for surprising, stressful hand-over situations and the need to monitor the automation to avoid them should be reduced.

Index Terms—Autonomy degradation workspace, human automation cooperation, joint control framework, air traffic management, levels of automation, automation ironies

1 Introduction

When an autonomous automation working side by side with a human reaches a limit for its autonomy, there must be a way for the human to cope with the autonomy degradation. Taking over in a stressful situation requires an understanding of both the situation and how it was handled by the automation before the autonomy degradation. This may take more time than handling it without the automation. Furthermore, the human must maintain knowledge about the fallback procedures in case it happens, which may be both costly and time consuming. In the worst case, the human has not performed the tasks without the automation for a long time with a resulting loss of skills, and the situation could quickly get critical. This would mean that the automation actually leads to more work, demands on higher skills, and potentially in new dangers - contrary to why the automation was most likely introduced in the first place. All of these issues were described by Bainbridge [1] as automation ironies. Despite a lot of effort to mitigate these problems, they are still valid decades later [2], [3].

In this article, the design concept *Autonomy Degradation Workspace* (ADW), Fig.1 is proposed to mitigate some of these issues, specifically aiming at highly autonomous systems. When autonomy decreases and cooperation between automation and human is needed, the ADW should enable the automation to initiate the communication around the situation by consulting the human in a well-timed way that adds as little cognitive load as possible on the human. This puts the human in control to decide on if, how, and when to respond. The key aspects are to let the automation initiate communication using information transformed to a format, or level, that matches the role and responsibilities of the human. Likewise, the timing of the communication should be adapted to the humans situation (Fig.1). This is what we refer to as *the right level of information at the right time*. The ADW was developed using the Joint Control Framework (JCF) [4] and a case-driven design approach using a case from the domain of Air Traffic Management (ATM). The JCF was chosen in favour of other frameworks first and foremost for its ability to describe temporal aspects of control processes and the interaction between the processes.

The article first presents related work, then the ATM domain and a concept design case based on a future ATM system with a highly autonomous automation and a human working together, and the case is modelled using the ADW principles. Finally, different aspects of the ADW are discussed and how ADW can be implemented, some directions for future research are given, and what conclusions can be drawn from the case modelled.

- *M. Shell was with the Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, 30332. E-mail: see http://www.michaelshell.org/contact.html*
- *J. Doe and J. Doe are with Anonymous University.*

Manuscript received April 19, 2005; revised August 26, 2015.

2 Related Work

Humans working with automated systems is not a new and unknown phenomenon, yet there are still issues with respect to human-automation cooperation that to a large extent remain unsolved.

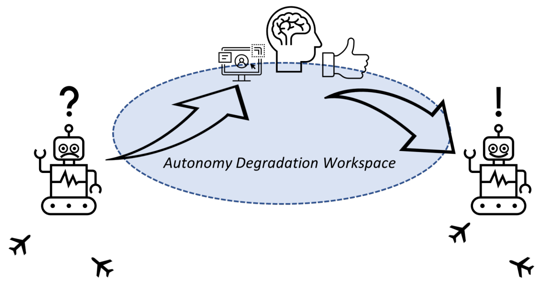

Fig. 1. The ADW is a concept for how to design an autonomous automation that can handle a degradation in autonomy by consulting the human as a colleague. It is achieved by presenting information at the right level and at the right time, thus making it possible for the human to understand what is asked for and time to provide an appropriate response to the automation.

2.1 Still Struggling to Team Up with the Automation

In 2012, Baxter et al. [2] made a review on what had happened during the 30 years since Bainbridge [1] formulated the now well recognized ironies (e.g. that the human's task shifts from working with a process to monitoring the automation now working with the process; a shift of tasks rather than a reduction in workload and to a task, monitoring, for which humans are not the best). The overall conclusion was that at least some of them were still left to solve. More recently (2018), Strauch [3] revisited the ironies and came to largely the same conclusions - they are still valid and in many cases unresolved. Strauch went through a number of incidents, mainly from aviation, where the classical automation ironies were found to be as valid as ever. It was pointed out though that in general, safety has increased in the aviation sector as the knowledge about how to deal with the ironies has increased even if they are not solved. Strauch also notes that the development of new types of, and more autonomous, automation has introduced new ironies or issues. One of those is the autonomy conundrum described by Endsley [5], which relates to the operator's decreased Situational Awareness (SA) caused by high Levels of Automation (LOA) and reduced possibility to revert to manual control when needed. Endsley consider this a big challenge that must be addressed with good design. Trapsilawati et al. [6] points at engaging the operator in the work and using the automation as one way to counter the risk of reduced SA when automation increases. Endsley also emphasize the importance of shared SA between automation and human if the automation is highly autonomous and working in team with the human [5]. In this case, Endsley states that LOA is likely be high with low degree of adaptivity, and low granularity of control for the human, i.e. that the human's direct engagement in the automation's tasks are low.

In the SESAR Master Plan [7], which describes a development road-map for European ATM, increased automation is considered one of the key factors to achieve the goals of Single European Sky. Five different LOA are defined, where level five is full automation with no human involvement. For ATM, it is not expected that this fifth level will be fully achieved, at least not in the near future. Instead, the master plan points toward adaptive automation concepts. Automation is seen as able to initiate most tasks (but not all) and is not expected to be able to carry all of them fully. As long as full autonomy is not achieved, automation is described as a support.

Strauch [3] opens up for the possibility that maybe some of the ironies simply cannot be solved as long as we keep the premises the same, e.g. that the human shall be the one stepping in when automation gives up, and just look for countermeasures. Hence, are there other ways of tackling the problem?

2.2 Humans and Automation Working Side by Side

A suggestion on how to change paradigm on a conceptual level is given by Norman [8]. Norman describes how the thinking around automation in the control of highly autonomous cars should change from emphasis on LOA-based dividing of control tasks, to an approach where the focus is on the human-automation cooperation, thus getting away from the problem of the human having to constantly supervise the automation and acting as backup. Instead, the human and automation shall continuously cooperate. Shively et al. [9] also focus on cooperation when presenting how the principles of Crew Resource Management (CRM) can be used in Human-Automation Teaming (HAT) [10], [11], [12]. Bi-directional communication between automation and human is seen as a crucial factor. Roth et al. [13] focused on function allocation, but also emphasized in the conclusions that cooperation is an important aspect that must be considered. Earlier, Bradshaw et al. [14] brought up the need for focus on cooperation, though also including critique on the use of LOA for autonomous system, e.g. the underlying idea that work can just be switched between human and automation with no effect on the system. The critique has in turn been responded to by Kaber who means that it is a mix-up between levels of automation and levels of autonomy, not a problem with levels of automation as such [15]. Kaber also responds to concerns about that LOA may not be the right starting point for design of human automation coordination and cooperation (concerns raised in e.g. [16]). Kaber means that that sooner or later, the design reaches engineers for implementation and then it must be specified and made concrete who should do what in which situation, something for which a LOA framework is convenient. This is in turn replied on by Jamieson and Skraaning [17] who claims that LOA has lost its relevance as basis for human automation interaction design as the field of human automation interaction has developed over the years. For example, Jamieson and Skraaning points at that the effect on workload in complex work settings is badly predicted by LOA. LOA is not a concept with *one*, single definition. A multitude of variants and taxonomies have been presented since the first LOA was introduced by Sheridan and Verplank [18] in 1978, some more elaborated and some more simplified, as pointed out by e.g. Kaber [15] and Wickens [19]. Yet, the different LOA variants share the basic idea that tasks shall be divided between a human and an automation at different levels; either automation or human. Hence, it does not encourage a focus on cooperation. Further, just as implied by Jamieson and Skraaning, complex work and situations may not be possible to fully model using LOA, creating a mismatch between system design and reality.

There are proposed models that focuses on human-automation cooperation. Itoh and Pacaux-Lemoine [20] discuss the issue from the perspective of trust. A need for a common workspace for the automation and a human agent in order to understand each other and the cooperation process is identified. This common workspace is also discussed by Pacaux-Lemoine and Flemisch [21] and an example from Air Traffic Control (ATC) is discussed (first presented in [22]). Cooperation is mainly centered around tasks and takes place within any of three different layers of cooperation, which are linked together by the workspace.

Principally the same model is elaborated on by Flemisch [23], who combines it with theories of control sharing. Flemisch also introduces a temporal aspect, however, it is done by showing a snapshot of a situation with different control sharing states at two distinct points in time. Gutzwiller et al. [24] suggest to organize cooperation in human-automation teams by pre-defining task allocation under different conditions in a working agreement, also including transition points defining when to change the allocation of tasks and responsibilities.

Lundberg and Johansson's Joint Control Framework (JCF) [4] and its associated score notation enables description and analysis of how an agent interacts with a process over time, e.g. an ATCO controlling air traffic. By using it to model both a human and an automation as well as the interaction between the human and the automation, the cooperation can be described and understood. The JCF can be used both for analyzing existing systems and for analyzing non-existing, first-of-a-kind type of systems [25].

3 Designing an Autonomy Degradation Workspace

A few basic ideas were used as starting points upon which we built the ADW concept. The automation is considered a highly autonomous, cognitive agent that works together with a human agent in different roles at different levels of cognitive control; a different view from considering the automation a tool to be used by the human. The latter may be appropriate as long as the automation has limited autonomy and is controlled by the human, but less so when autonomy grows. Temporal aspects of the cooperation between the automation and the human is a key aspect, as is the shift from task sharing based on LOA to a more elaborate cooperation process. However, it may still be relevant to talk about LOA on a higher, overall system level even though we do not use it for specific interactions.

The ADW concept was elaborated on and developed by using a scenario-driven design process. First it was developed as a theoretical concept using an analysis framework and then the ideas were applied on a specific case. The case used was an ATM case, as ATM was considered a relevant domain. The case was built up as a scenario in an air traffic control real-time simulator, which enabled the temporal aspects of the ADW concept to be studied. Deliberately, only one case was used to enable a more in-depth analysis and description of the ADW in favour of more examples. Even though exemplified within a specific domain, the principles of the ADW could relate to any type of autonomous automation that needs to cooperate with a human. In the rest of the chapter, the ADW principles are described in more detail and the domain, including the case, as well as the chosen analysis framework are presented.

3.1 Application context: Air Traffic Management

Air Traffic Management (ATM) is a domain in which automation has been more or less constantly increasing, from the introduction of primary radar to today's complex systems with a multitude of tools and sub-systems that assist the ATCOs. The increase in automation is mainly driven by demands of increase in efficiency, while at the same time, the high safety standards must be maintained. However, automation is still relatively low and varies between different systems, and higher automation is seen as a key component for success of future ATM systems [7], [26]. Even though the systems may reach a higher level of automation and getting more autonomous than today, ATCOs are expected to work with the ATM systems for a foreseeable future, but the roles may change [7].

3.2 The Scenario

The ADW concept targets systems which are highly autonomous. As autonomy in most ATM systems of today are quite limited, a scenario in an imagined, future ATM system was created. Lundberg et al. [25] used a similar approach to evaluate a non-existing, first-of-a-kind system, though there, both the technology and the traffic situation were first-of-a-kind, while here, the autonomous system was fictitious but the traffic situation was not. In this scenario, a highly autonomous system works side-by-side with the ATCO. The main automation manages the traffic on a tactical level, solving conflicts and carrying out plans made by the ATCO. The ATCO has an active role, but on a more long-term tactical time horizon and on a a higher level; making plans, setting up goals, and coordinating with other stakeholders.

The scenario was built around a traffic situation with a high level crossing of two aircraft at the same altitude with different solutions available: The trajectories of two aircraft are conflicting, Fig.2. Aircraft SAS123 is crossing slightly behind the route of KLM456 at an acute angle. If nothing is done, the distance will be below separation minima (5NM). The automation calculates that the most efficient solution is to turn SAS123 slightly to the right until it is free of traffic, i.e. KLM456. However, this would get SAS123 too close to the adjacent sector. How close *too close* is may vary, but a typical rule is that required distance to a sector border is half the distance required as separation between two aircraft, e.g. 2.5NM if the minimum separation is 5NM. This ensures that aircraft are always separated even if they are at each side of a sector border. The second best option is to turn SAS123 left, but that results in a longer flown distance.

In this situation, an ATCO would have made an estimate whether it would be a good idea to call the ATCO in the adjacent sector to ask for permission to fly closer to the sector border than the rule prescribes. The automation establishes that probably, it would be no problem. However, the automation does not have jurisdiction nor means to coordinate this with the ATCO in the adjacent sector. Furthermore, the ATCO working together with the automation may possess additional knowledge about the situation. The conclusion by the automation is that it is time to consult the ATCO, to establish an ADW. These kinds of trade-offs are not uncommon, but a normal part of the ATCOs work and one of many skills an ATCO must possess.

For the understanding of the situation, it is important to remember that the example in the case in an excerpt from one specific situation to illustrate the ADW principles. It is assumed that other solutions, for example changing speed or altitude, have already been rejected by the automation due to other factors, which could be e.g. wind conditions and other traffic, leaving changes in the lateral route as the best alternatives. Furthermore, one might argue that an autonomous automation should be designed to be able to make these kinds of situations and judgements by itself. In a predictable reality where all cases can be identified and defined, that would probably be a good approach. However, in reality, the world is complex and even if this scenario might be possible to solve in other ways, there will always be situations where the automation is exposed to unanticipated situations or simply gets outside its own autonomy limits.

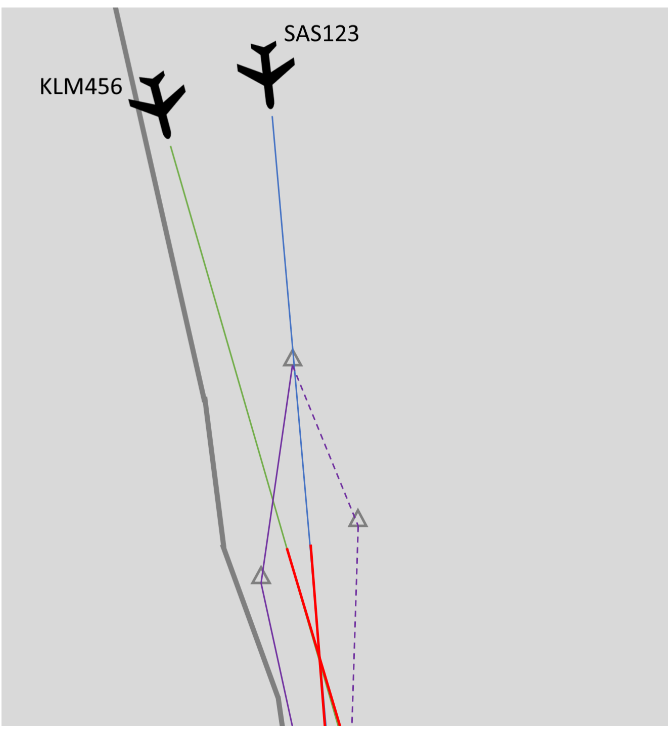

Fig. 2. Illustration of the traffic scenario (not to scale). The flight paths of SAS123 (blue line) and KLM456 (green line) will cross each other and the aircraft will come too close (red part of the flight paths). The automation calculates that the optimal solution is to turn SAS123 slightly to the right (full purple line), but that is too close to the border of the adjacent sector (thick grey line). The triangles show the turning points. The second best alternative would be a left turn (dotted purple line), however, a longer way for SAS123 to fly. The turning point where the purple lines begin is the latest point calculated by the automation when to give the turn command to the aircraft to achieve the intended effect. Hence, the response from the ATCO must be given earlier than that.

3.3 Analysis Framework

The Joint Control Framework (JCF) [4] was chosen as an approach to develop and model the ADW concept. The JCF combines ideas from several previous frameworks for modelling of control processes, e.g. the Extended Control Model [27]. In addition to bringing these ideas together into one framework, the JCF also provides a score notation that adds a temporal extension of the joints between control processes and controlled processes modelled in the JCF, which was an important reason of why it was chosen.

TABLE 1
Levels of Autonomy in Cognitive Control (LACC), adapted from [4].

Level	Description
6 Frames	Defines the situation and context, e.g. to maintain a safe and efficient air traffic in a certain airspace at a certain time.
5 Effects	Effect goals within the situation, e.g. keeping aircraft apart by a certain distance or striving towards minimum delay of traffic.
4 Values	Trade-offs between different criteria and effect goals.
3 Generic	Making up plans and how to implement them to achieve what has been decided upon on higher levels.
2 Implemenations	Actual implementation of plans, setting constraints and making implementation decisions.
1 Physical	Physical world object status, e.g. positions of aircraft on a radar screen.

In each process, a subject (e.g. an ATCO) is interacting with an object (e.g. an aircraft). Each point of interaction is called a cognitive joint and could relate to either perceiving information, making decisions, or performing an action. The notation for these interactions in JCF are perception points (PP), decision points (DP), and action points). While the PP:s and AP:s represent more explicit interactions with the object, the DP:s are more of a subject-internal nature, even though they still relate to the object in the control process and affects the object through the following actions. Each process interaction can take place at any of six different Levels of Autonomy in Cognitive Control (LACC) spanning from high level framing of the situation to physical interaction with interfaces, Table 1 (see [4] for details), and at a certain point in time.

JCF provides a score notation for this where each process is visualized by six parallel lines. Each line represents a certain LACC and their horizontal extension represents the time. The joints are depicted as dots in the score, like notes in a sheet music score. The horizontal position of the points in the score depicts the LACC at which the interaction takes place, and the horizontal position when it occurs. Johansson and Lundberg used a similar notation [28] with the cognitive joints distributed in time, but without different cognitive levels. Furthermore, the JCF enables modelling of not only principles, but specific episodes as well thanks to its temporal extension, which is in line with the ideas of ADW. The score notation also makes it possible to visualize multiple simultaneous processes, similar to a conductor's score. To make it easier to read, the example scores were enriched with arrows that depict the flow of cognitive joints (Fig.3 to Fig.6).

4 Resulting Designs and Analyses

The scenario was implemented in a real-time Air Traffic Control (ATC) simulator. However, since the case includes a highly autonomous automation that does not exist, the different solutions were scripted, which means that the aircraft flew as if they had got the instructions from the automation. The simulator was used as a player to visualize the traffic situation for modelling in the JCF score notation. Although using a scripted scenario, the real-time playback provided a realistic temporal progress of the traffic situation.

4.1 Transforming to and Presenting Information at the Right Level

A key component of ADW design is to identify the right level (Table 1) of the information to be provided by the automation to the human. The automation and the human works in parallel processes with the same objects and share the frames and goals set by the human, but the work is performed on different LACC. To illustrate the ADW principles, two JCF scores were used to start with. One score shows the process of the ATCO controlling the air traffic on a high level by setting goals, level 5 (Effects). A second score shows the automation working with the same process, but on a lower level, solving tactical problems and implementing the solutions. The figures 3 to 5 show step by step how the situation is identified by the automation, communicated to and acted upon by the ATCO, and finally how the automation receives the response and implements the solution. The work is mainly performed on levels 1-3 (physical, implementations, and generic), though it is governed by goals set on level 4 and 5 (values and effects). The automation identifies the conflict between the to aircraft (Fig.3, point 1), compares the rules of separation and the goal of optimizing the traffic (Fig.3, point 2). A decision is made that the ATCO shall be consulted (Fig.3, point 3 and 4) to sort out if it is a good idea to ask for approval to implement the most efficient solution.

When initiating the communication about the ADW, the automation must be able to transform the low level information into an information package that can be presented and understood at the level of which the human agent is working (Fig.4 point 4 and 5) to avoid that the ATCO must switch cognitive levels when prompted by the automation. Consider the differences between the two descriptions of the problem with the SAS123-KLM456 crossing presented in Table 2:

TABLE 2
Description of the same problem at different LACC.

LACC	Description
3 Generic	”I need to solve conflict between SAS123 and KLM 456, should I turn SAS123 left or right?”
5 Effects	”To solve a conflict optimally, I would like to let SAS123 fly 1.7NM closer to to sector X than allowed, no conflicting traffic detected in sector X. Is that OK?”

The first is a medium-low level description of the problem, level 3. It would force the ATCO to dig into the situation at level 1 or 2 to understand what the problem is, what the alternatives would mean, and maybe consider other alternatives and comparing them to the higher level goals, and finally, decide on which plan to implement. All of this takes time and may put the ATCO out of the loop of his/her work with the high level processes. The second description clearly states that it relates to a conflict between a rule (required distance to the sector border, safety) and a goal (optimization), i.e. relating directly to the goals and trade-offs at levels 4 and 5 where the ATCO is working, and that no issues with respect to other traffic than those in the original conflict are identified. The ATCO just has to decide whether or not it is feasible to coordinate with the adjacent sector to get a permission to make an exception from the rule to reach the optimization goal, a level 4 decision.

Note that when presenting the information to the human, the action to initiate the ADW (Fig.4, point 4), is taken on a lower level, 3, than the level of the information perceived by the ATCO, 5 (Fig.4, point 5). This transformation of the information is very important and one of the core ideas of the ADW. The human agent can then evaluate the information in a controlled manner starting from current level. A decision is made made on a slightly lower level, 4, while when providing the response, it is back on level 5 (Fig.4, point 6-7). The whole process can be paced by the ATCO and no extensive information gathering from lower levels is required.

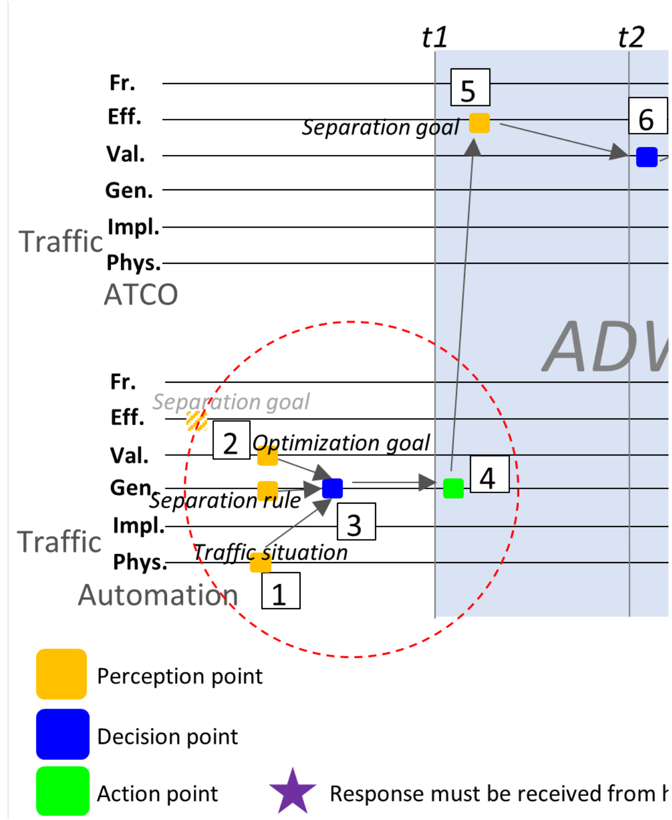

Fig. 3. Joint Control Framework (JCF) score to visualize the ATC example. The time is not to scale for readability reasons. The automation detects the conflict between the two aircraft (1). It takes into the rules and goals at hand (2) and concludes that an optimal solution would require an exception from the rule stipulating how close to a sector border an aircraft is allowed to fly and that is must consult the ATCO (3) which is done (4).

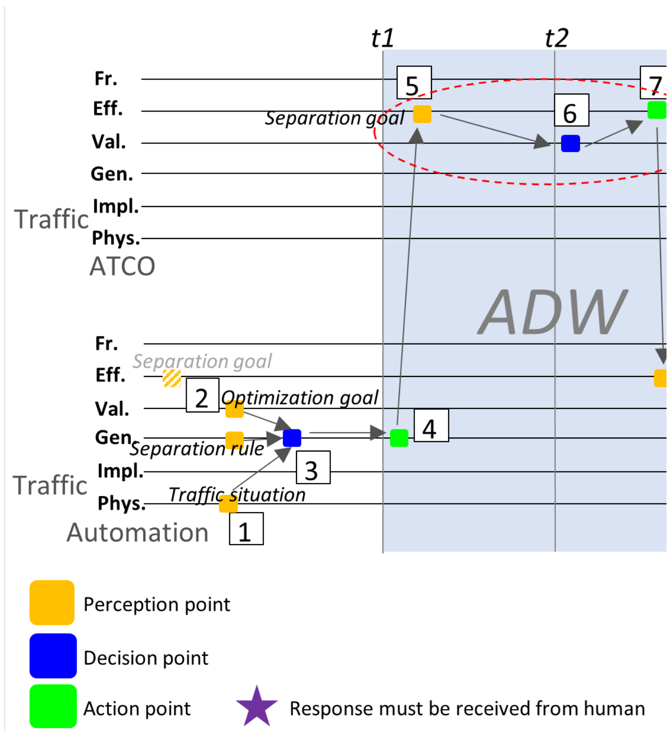

Fig. 4. The information about that the automation needs consultancy from the ACTO is presented to and perceived by the ATCO (5). Note that it is the level of the information that is indicated in the score, not how or from which HMI it is perceived. The ATCO makes a decision on how to act (6). In the example that would be coordinating with the adjacent ATCO. The result of this is given as response to the automation (7).

The response is given to the automation, which transforms it back to a low level solution that is compliant with the revised goal (Fig.5 point 8-10). If positive, i.e. that the ATCO has deemed it a good idea to make the coordination with the adjacent sector and have provided an affirmative answer, the optimized solution is implemented by the automation by modifying the route of SAS123 to include a right turn. If negative, the automation implements the second best alternative instead and gives the aircraft a left turn. By following this procedure, the ATCO has not had to get into details of possible solutions, but has been able to continue to work with the plans and goals. It is instead the automation that has made most transformations between different cognitive levels and spared the human that work. If looking at the ATCO working during the process, it may be the case that he or she casts an eye at the radar screen, i.e. perception of information at a low LACC, to get an overview of the total situation. That is however different from using it for getting detailed, low level information of each control process in order to compare them to high levels rules and goals as a basis for the decision making. And of course, as long as the information is present, nothing actively prevents the ATCO from taking in more low level information, but it is not needed for the ADW process, and if it is done, it is on the human's terms and not forced upon her.

This ability to transform the information to suite the ATCO is a crucial aspect of the automation's competence and should hence receive much attention in the design process. By always following the same procedure when initiating the ADW, the predictability in the communication is high - the human knows what to expect. Together with timing of the communication, it moves away from the unwanted, stressful handover situations and acts as a cooperating system rather than one trying until failing and then issuing an alarm.

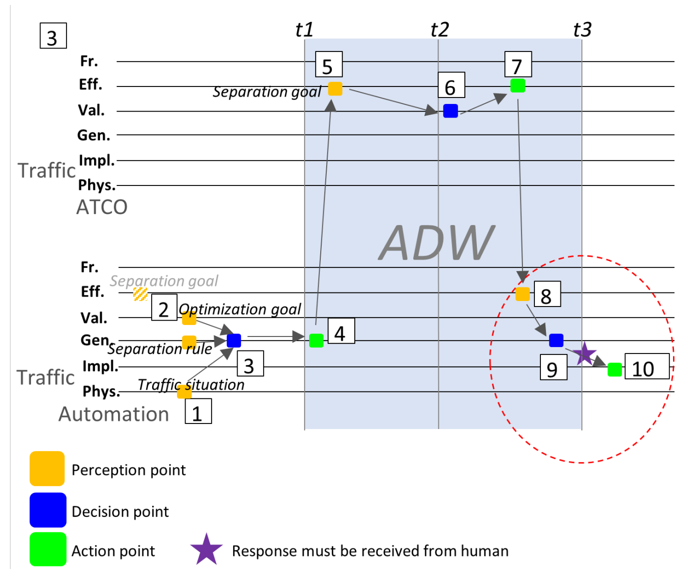

Fig. 5. The automation receives the response from the ATCO (8) and transforms this into a decision about which solution to implement (9) and then implements it (10). If the ATCO gave a positive response, the main alternative is implemented, otherwise the automation goes for the second best alternative. The solution can be implemented well before t3 if the ATCO responds earlier.

4.2 Rhythm and Timing by Adaptation

To avoid workload peaks, the number of ADW:s presented at the same time must be kept at a manageable level, and it would not be a good idea to present the ADW just as the ATCO is busy with something else. This is achieved by providing the autonomous automation knowledge about the situation of the human, e.g. by eye tracking and real-time analysis of system interaction. Using such information, the automation can distribute the presentation of the ADW:s to keep a steady pace and rhythm. Therefore, we introduce a third JCF score which visualizes the automation's process of knowing what the ATCO is doing (Fig.6, lowest score).

When the decision is made to establish an ADW (Fig.6, point 3), the automation also knows when it must get the response from the human to be able to implement the solution in due time (Fig.6, $t3$). To give the human a chance to reflect upon the ADW information and make a decision on if and how to respond, there must be a minimum time available (Fig.6, $t2$-$t3$). The minimum time required may depend on the application domain, but the idea is that it shall be a predefined, fixed amount of time in order to provide predictability. Even if a certain amount of time is needed to avoid surprises, initiating the ADW too early may result in an overload of ADWs. Furthermore, it is reasonable to assume that looking too far into the future will increase uncertainty due to the complexity of the real world, e.g. weather changes and other unforeseen, hard-to-predict

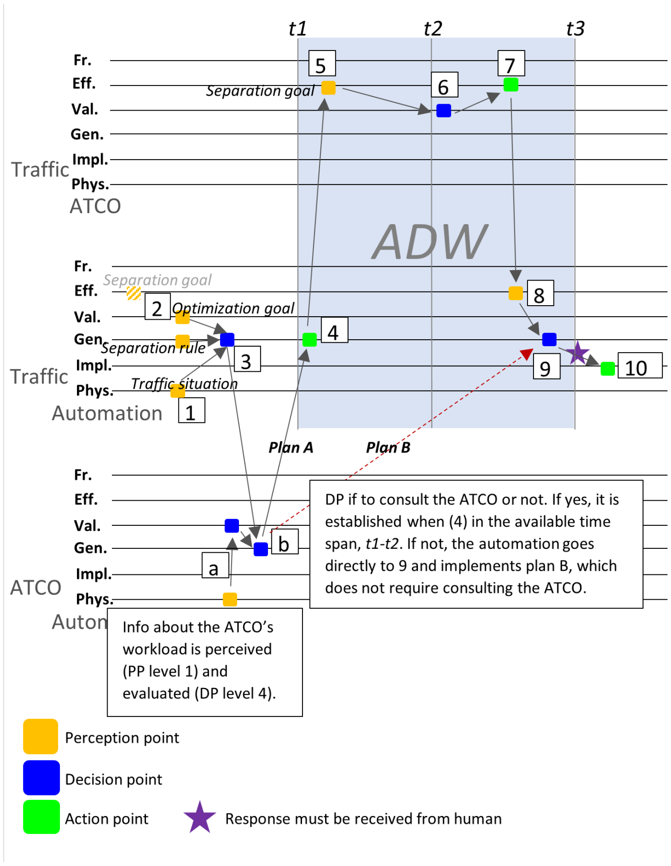

Fig. 6. Introducing the process of the automation looking at the ATCO to be able to adapt the ADW to the ATCO's situation. When the automation has decided that an ADW shall be established (3) it checks the whereabouts of the ATCO (a), e.g. workload and attention. It combines this information with when response is needed ($t3$) and decides on (b) when between ($t1$) and ($t2$) the ADW shall be initialised (4). If no convenient point in time is found, the conclusion is that the ADW shall not be established and the automation goes for implementation the second best alternative (9) without consulting the ATCO.

events. At some point early enough, it will be difficult to predict if something will actually develop into a problem at all. Therefore, an ADW horizon is suggested which defines the earliest time before needed response that the automation can initiate an ADW (Fig.6, $t1$). Consequently, an ADW can be presented between $t1$ and $t2$, exactly when depends on the situation. While $t2$ is supposed to be a fixed time, $t1$, the ADW horizon, is suggested to be adaptable to give the human a possibility to set the frames for the ADW. The automation uses the knowledge about the ATCO's situation (Fig.6, point a) to decide on when between $t1$ and $t2$ to initiate the ADW (Fig.6, point b, 4, and 5). Even if $t1$, the horizon, is set by the human operator, possible values may be depend on the implementation and domain specific prerequisites.

When looking at the different scores (Fig. 3 to 6), it might look like there is plenty of time, but the score only show these specific processes. Most likely, there are many more processes and events which could affect the situation. There will be other aircraft, coordinations with different stakeholders, strategic planning activities, and so on. All of this must be taken into account when the automation is deciding on when it is appropriate to establish the ADW, i.e. point a and b, leading to the action in point 4 for at $t1$ in Fig. 6.

Importantly, the ADW is supposed to be applied to non-critical situations. Hence, if by some reason no response is given, the only effect should be a decrease in efficiency, in this case some extra flown miles, not safety. However, if missed responses starts to appear more frequently, it is a clear signal about that something is not working optimally. Critical issues and situations though must be handled by e.g. alarms and back-up systems.

5 Discussion

The resulting JCF scores clearly show the parallel human and automation control processes of the analyzed scenario and how the processes are related to each other. It visualises the importance of the temporal aspects of the ADW and how the automation should work with it. All steps within the ADW follow the same pattern of perception, decision, and action. The critical step is to give the ATCO enough time between perception and decision (Fig. 6 point 5 and 6). To achieve this, the automation must possess knowledge about the workload of the ATCO. Though illustrated as points in the scores (Fig. 6 point a), this should be a continuously ongoing task for the automation to measure and evaluate the workload so that the information is always present when the need for an ADW occurs. When so, the automation must be able to answer three questions before establishing the ADW: Shall I consult the ATCO, and if so, when should I do it, and how? To answer the first question (Fig. 6 point b), the automation must evaluate the possible gains of getting the help from the ATCO against the risk of overloading the ATCO with too many tasks to perform. Hence, the automation must know not only what the workload is, but also have access to a calibrated workload limit to compare with.

By deconstructing the resulting ADW analysis piece by piece, the consequences of *not* following the ADW principles gets quite clear. If the first question, *if*, is not addressed, the automation will always present the information, regardless of the situation. What has then been created is, with respect to the temporal aspect, an alarm. Point 3 and 4 (Fig. 6) would take place practically simultaneous and the information would be presented to the ATCO earlier than needed. If the the first question, *if*, is addressed but not the second one, *when*, there would still be an alarm-like communication, but fewer occurrences. Finally, if raising and answering the first two questions but taking away the *how*, it would mean presenting the information at the same level of which the issue arose, moving point 5 (Fig. 6) in the analysis from LACC 5 to 3. This corresponds directly to the differences between the two LACC-dependent phrasings presented in Table. 2. Hence, even if he timing is seemingly right, it may create more work for the ATCO to understand the information in relation to the situation and the LACC of which he or she is working. This may in turn take time and make the estimation of the time needed harder for the automation to make, creating a viscous circle of self-reinforcing bad timing.

Though using the same notion of a workspace as in the models presented by Pacaux-Lemoine and Flemisch

[21], there are some important differences. In the Pacaux-Lemoine and Flemisch model, the workspace is separate from the controlled external process itself and only relates to the communication between the automation and the human, furthermore, the automation is still considered a support system. The results is then communicated through an interface, affects the process, and propagates back to the task goals in the process in a cycle with no defined extension in time.

5.1 Working Together

To show and analyze the consequences of the ADW, we used one specific scenario, but the principles are generic and similar scenarios should follow the same pattern for the human-automation cooperation. Similar situations could for example be trade-offs between local and global optimizations or incoming data that is not consistent, which may confuse the automation. In addition to making the extension in time clear, the parallel process JCF scores in the analysis also elucidates the cooperation between the automation and the ATCO and how they work with the same subject, the air traffic.

By working this way, side by side in a team, the human and the automation are continuously involved in controlling the same external processes, even though on different levels, not unlike musicians in an chamber music or jazz ensemble playing together without a conductor. A design goal for the ADW is that there should be no need for understanding exactly how the automation works, just as every musician does not need to know every detail of the other parts in the ensemble or be able play all instruments. Consequently it is important that the human can trust that the automation will ask for consultation when needed and that the automation will adapt the communication to suit the situation of the human. However, that is not different from how humans must trust each other when working together. To continue the music metaphor; the human and automation must learn to play together. The ADW design can contribute to this by providing predictable behaviour - the human shall not be surprised by issues communicated on a totally inconvenient level or with bad timing that interrupts the humans work or leaves too little time for response.

5.2 Future Research

We consider ADW a promising concept which should be suitable for any domain where human operators work with highly autonomous systems. Vessel traffic service (VTS) and train control closely related are domains sharing a lot of properties with ATM with respect both to control processes and the control room environments. In addition to extending it into different domains, it should also be addressed how the ADW principles can be used in more complex settings with teams larger than one human and one automation.

The ADW design concept must be supported by human computer interfaces (HMI) that facilitates the ADW principles. How it shall look and feel is a question for future research and will probably differ depending on the particularities of the domain, e.g. the visual context and processes on which it shall be applied.

Artificial intelligence (AI) is likely to to be a part of future systems and could provide the possibility for systems with learning abilities. However, if the automation can learn and improve its behaviour, it also means changing the behaviour. In what ways this shall be allowed and how it affects predictability and trust in the cooperation are important questions to address.

Finally, can the information about when, how often, and why ADW:s are established be used? One area to look into could be if it can be used as a performance indicator for the system as a whole. It could mean both using it for monitoring the system status or using the data for off-line analysis in order to gain deeper understanding of the human-automation cooperation. It could also be looked into if this could be used in an aggregated form by a supervisor role as an indicator of the overall status of the system of humans and automation.

6 Conclusions

Summarizing the results of the design of the *Autonomy Degradation Workspace* (ADW) concept and the analysis of a case applied in the *Air Traffic Management* (ATM) domain, we draw four main conclusions:

Firstly, to describe the ADW concept, a notation for the control processes and interactions over time is needed. The key characteristics are the adaptation of timing and the transformation of information levels in the interaction between the three processes.

Secondly, the cognitive joints in the modelled ADW can be grouped into four distinctive steps:

1) Identification by the automation of a need for an ADW
2) The automation's evaluation of if, when, and how to present the ADW
3) Perception of and response to the ADW by the human
4) Implementation by the automation of a solution based on the response from the human

Thirdly, temporal adaptation shall be done in real-time but the definition of target level for the transformation of information shall be done off-line. The rationale is that the level of cognitive control at which the human should work on is expected to be fairly constant, and if it is changed, there are no means of detecting and categorize it in real-time. By sticking to this, predictability can be maintained.

Last, but not least, the modelled ADW includes three processes: the Air Traffic Controller's (ATCO) work with the traffic, and the automation's work with the traffic, and the automation's monitoring of the ATCO's situation. This means that the need for the ATCO to monitor the automation to be able to take over is not there. If there had been a need to monitor the automation, it would have been shown in a fourth score, a process with the ATCO as subject and the automation as object to be monitored.

To summarize, the ADW concept is a potential way forward to solve some issues often encountered in human-automation cooperation by letting the automation answering if, when, and how the human shall be consulted in case of a degradation in autonomy. By doing so, the risk for overloading the human by initiating communication at an inconvenient time is reduced, while still providing enough time for a response. The risk for surprises is reduced by providing information at a consistent level of cognitive control which is tuned to match the level that the human is mainly working at. Finally, as the automation only consults the human if possible, the human should not have to continuously monitor the automation to try to foresee what the automation is up to.

Acknowledgments

The research was funded by the Swedish Transport Administration through project F AUTO (TRV 2018/41347).

References

[1] L. Bainbridge, "Ironies of automation," *Automatica*, 1983.
[2] G. Baxter, J. Rooksby, Y. Wang, and A. Khajeh-Hosseini, "The ironies of automation ... still going strong at 30?" in *ECCE '12: Proceedings of the 30th European Conference on Cognitive Ergonomics*, 2012, pp. 65–71.
[3] B. Strauch, "Ironies of Automation: Still Unresolved after All These Years," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 5, pp. 419–433, 2018.
[4] J. Lundberg and B. Johansson, "Cognition, Technology & Work - A framework for describing interaction between human operators and autonomous, automated, and manual control systems," *Cognition Technology & Work*, 2020.
[5] M. Endsley, "From Here to Autonomy: Lessons Learned from Human-Automation Research," *Human Factors*, vol. 59, no. 1, pp. 5–27, 2017.
[6] F. Trapsilawati, C. D. Wickens, X. Qu, and C. H. Chen, "Benefits of Imperfect Conflict Resolution Advisory Aids for Future Air Traffic Control," *Human Factors*, vol. 58, no. 7, pp. 1007–1019, 2016.
[7] SJU, "EUROPEAN ATM MASTER PLAN EXECUTIVE VIEW," 2020.
[8] D. A. Norman, "THE HUMAN SIDE OF AUTOMATION," pp. 1–6, 2015.
[9] R. J. Shively, J. Lachter, R. Koteskey, and S. L. Brandt, *Crew resource management for automated teammates (CRM-A)*. Springer International Publishing, 2018, vol. 10906 LNAI. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91122-9_19
[10] B. G. Kanki, R. L. Helmreich, and J. Anca, *Crew Resource Management*, 2nd ed. Academic Press, 2010.
[11] D. E. Murinho and P. S. Murray, "Handbook of Human Factors Aviation," in *Hanbook of Aviation Human Factors*, 2nd ed., J. A. Wise, V. D. Hopkin, and D. J. Garland, Eds. CRC Press Taylor & Francis Group, 2010, vol. 53, no. 9, pp. 224–243.
[12] G. E. Cooper, M. D. White, and J. K. Lauber, "Resource Management ob the Flight Deck," in *Proceedings of a NASA/Industry Workshop Held at San Fransisco, california June 26-28,1979*, San Franciso, CA, 1980, pp. 1–16.
[13] E. M. Roth, C. Sushereba, L. G. Militello, J. Diiulio, and K. Ernst, "Function Allocation Considerations in the Era of Human Autonomy Teaming," *Journal of Cognitive Engineering and Decision Making*, 2019.
[14] J. M. Bradshaw, R. R. Hoffman, D. D. Woods, and M. Johnson, "The seven deadly myths of 'autonomous systems'," *IEEE Intelligent Systems*, vol. 28, no. 3, pp. 54–61, 2013.
[15] D. B. Kaber, "Issues in Human–Automation Interaction Modeling: Presumptive Aspects of Frameworks of Types and Levels of Automation," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 7–24, 2018.
[16] K. Christoffersson and D. D. Woods, "1. How to make automated systems team players," *Advances in Human Performance and Cognitive Engineering Research*, vol. 2, pp. 1–12, 2002.
[17] G. A. Jamieson and G. Skraaning, "Levels of Automation in Human Factors Models for Automation Design: Why We Might Consider Throwing the Baby Out With the Bathwater," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 42–49, 2018.
[18] T. B. Sheridan and W. L. Verplank, "Human and Computer Control of Undersea Teleoperators," Massachusetts Institute of Technology, Cambridge, MA, Tech. Rep., 1978.
[19] C. Wickens, "Automation Stages & Levels, 20 Years After," *Journal of Cognitive Engineering and Decision Making*, vol. 12, no. 1, pp. 35–41, 2018.
[20] M. Itoh and M. P. Pacaux-Lemoine, "Trust View from the Human-Machine Cooperation Framework," *Proceedings - 2018 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2018*, pp. 3213–3218, 2018.
[21] M. P. Pacaux-Lemoine and F. Flemisch, "Layers of shared and cooperative control, assistance, and automation," *Cognition, Technology and Work*, vol. 21, no. 4, pp. 579–591, 2019. [Online]. Available: http://dx.doi.org/10.1007/s10111-018-0537-4
[22] M. P. Lemoine, S. Debernard, I. Crevits, and P. Millot, "Cooperation between humans and machines: First results of an experiment with a multi-level cooperative organisation in air traffic control," *Computer Supported Cooperative Work*, vol. 5, no. 2-3, pp. 299–321, 1996.
[23] F. Flemisch, D. A. Abbink, M. Itoh, M. P. Pacaux-Lemoine, and G. Weßel, "Joining the blunt and the pointy end of the spear: towards a common framework of joint action, human–machine cooperation, cooperative guidance and control, shared, traded and supervisory control," *Cognition, Technology and Work*, vol. 21, no. 4, pp. 555–568, 2019. [Online]. Available: https://doi.org/10.1007/s10111-019-00576-1
[24] R. S. Gutzwiller, S. H. Espinosa, C. Kenny, and D. S. Lange, "A design pattern for working agreements in human-autonomy teaming," *Advances in Intelligent Systems and Computing*, vol. 591, pp. 12–24, 2018.
[25] J. Lundberg, M. Arvola, C. Westin, S. Holmlid, M. Nordvall, and B. Josefsson, "Cognitive work analysis in the conceptual design of first-of-a-kind systems–designing urban air traffic management," *Behaviour and Information Technology*, vol. 37, no. 9, pp. 904–925, 2018.
[26] Federal Aviation Administration, "NextGen implementation plan 2018-19," Washington. DC, 2019.
[27] E. Hollnagel, *Joint Cognitive Systems*, 2005.
[28] B. J. Erik Johansson and J. Lundberg, "Resilience and the temporal dimension–the chimera of timely response," *Theoretical Issues in Ergonomics Science*, vol. 18, no. 2, pp. 110–127, 2017. [Online]. Available: https://doi.org/10.1080/1463922X.2016.1154231

Magnus Nylin Industrial PhD candidate at Linköping University, Sweden, for the Air Navigation Services of Sweden (LFV). The main focus of his research is human-automation cooperation and how adaptive automation can be used to improve this. He also has long experience from research and development within the air traffic management domain as a system engineer and simulation expert, mainly from verification and validation of novel air traffic control concepts.

Jonas Lundberg Biography text here.

Jimmy Johansson Westberg PhD (Linköping University, Sweden, 2008) in Visualization and Interaction. The main focus of his research is information visualization of high-dimensional and time-varying data, with focus on parallel coordinates. Another focus area is quality metrics for visual representations. He works in areas such as visualization of traffic management and unmanned aerial vehicles. He also has experience in visualization of climate data.