MODEL-BASED ANALYSIS OF HUMAN-MACHINE INTERACTION (HMI)

Names, Affiliations

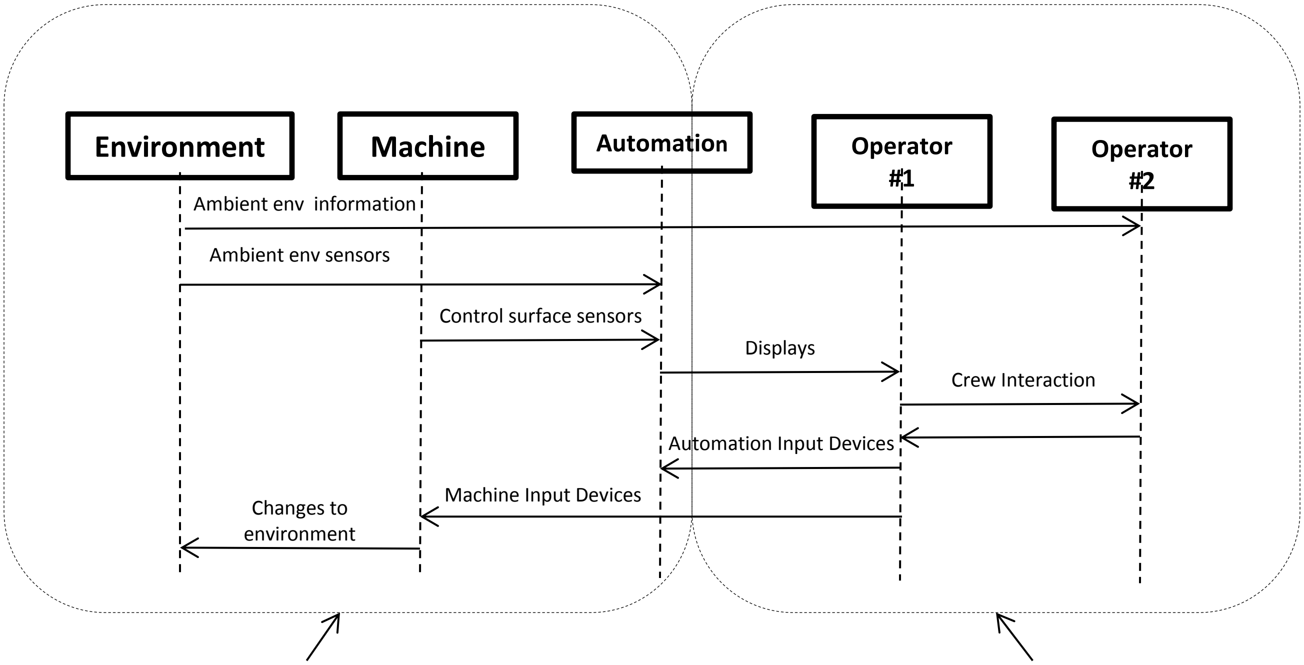
Abstract: The dominant effort in the development and deployment of a system is in the design and testing of the system technology and it's interaction with the operational environment. Advancements in engineering design processes have applied Computer-Aided Design and Model-based System Engineering (MBSE) concepts to ensure mating of physical components, functioning data interfaces, and appropriate system behavior and performance. Although increased emphasis has been placed on human factors and the design of human-centric user-interfaces, because the system boundary is the user-interface, the same level of CAD/MBSE rigor does not exists for the design and analysis of the *interaction* between the operator and the machine (i.e. the operational procedures).

This paper describes a method for specification and analysis of the procedures for Human-Machine Interaction (HMI). The HMI is documented using Sequence Diagrams (also known as Interaction Diagrams) tailored to specify the interaction between one or more crew members and the machine. The "grammar" of the tailored Sequence Diagrams can be used to analyze the performance of the HMI. The HMI Sequence Diagram can also be used in a Monte Carlo simulation to evaluate the HMI in a range of circumstances including the presence of stochastic machine performance and disruptions, and across a population of operators. A case-study is provided along with a discussion of the limitations, future work and implications of the approach.

INTRODUCTION

The dominant effort, time and cost in the development and deployment of systems is in the specification, design, analysis and testing of the system technology and it's interaction with the operational environment. This includes the physical interfaces with the environment of the sensors, actuators and their control surfaces, the user-interfaces, as well as temperature, vibration, radiation and other natural phenomenon. Advances in Computer Aided Design (CAD) and Model-based System Engineering (MBSE) have significantly improved the quality and productivity of this engineering design processes. Physical components can be mated in the virtual world of a CAD model, and systems and their functions can interact with each other in the virtual world of MBSE.

Although there has been increased emphasis on human factors and the design of human-centric user-interfaces, the same level of CAD/MBSE rigor does not exists for the specification, design, and analysis of the *interaction* between the operator and the machine, and between crew member operators (Figure 1). The root of the problem is that the human-machine interaction is not explicitly specified in the system requirements. Procedures, task analyses (e.g. GOMS) or cognitive engineering analysis are shunted off the design critical path or initiated after the design is complete (see Sherry & Feary, 1998). In part, the methods used for task analysis and human factors engineering may also not be compatible with the system/hardware/software engineering process, do not account for the operational factors such as the operational time window to complete a task, stochastic machine performance, operational disruptions, or a range of user performance (e.g. fatigue, experience).



CAD/MBSE and Development Process emphasize interaction between "system" and environment

CAD/MBSE and Development Process do not emphasize the interaction between "system" and crew members

FIGURE 1: CAD/MBSE rigor does not exists for the specification, design and analysis of the *interaction* between the operator and the machine (i.e. the procedures)

This paper describes the HMI Sequence Diagram (HMI-SD) to explicitly capture the human-machine interaction. The HMI-SD is based on the standard system engineering Sequence Diagram (also known as an Interaction Diagram) found in most MBSE Computer Aided System/Software Engineering (CASE) tools and is compatible with the engineering process. The standard SD is tailored to capture the interaction between human operators and the machine (including the automation) as well as between crew members. The HMI-SD can be used for analysis of the HMI weaknesses in the procedure as well as for performance measures such as time-on-task and probability of failure to complete the task in the operational time window. The HMI-SD can also be executed in a Monte Carlo Simulation to assess the performance of the procedure across a population of operators, across range of operational circumstances, in the presence of disruptions.

This paper is organized as follows. The next section provides an overview of operational procedures and the HMI/HCI process. The following sections describe the HMI Sequence Diagram, analysis of the HMI-SD and a case study. The paper concludes with a discussion of the implications of this method, limitations and future work.

OPERATIONAL PROCEDURES AND HUMAN-MACHINE INTERACTION

The "Command & Control center" of an enterprise, processing plant, or vehicle is composed of a team of operators (i.e. crew) and a suite of automation (see Figure 1). The automation includes sensors used for measuring the environment (e.g. ambient temperature, wind velocity), sensors for measuring the machine (e.g. engine temperature, control surface position), actuators for moving control surfaces, and decision-making logic. The crew interact with the machine (e.g. aircraft) either directly (e.g. a lever connected to the control surface) or indirectly through the automation (e.g. a knob or button that activates a control system that commands the control surfaces).

The mission, executed by the crew in conjunction with the automation and machine, is defined by a set of operational procedures (or "procedures"). The procedures outline the conditions under which they are activated and the sequence of actions that must be taken to achieve the objective of the procedure (Mauro et. al., 2016; Degani et al. 1997). The procedure typically divides the responsibilities between crew members (e.g. pilot flying and pilot monitoring) and includes "call outs" to ensure that all the crew members are coordinated in their individual tasks yet maintain a shared mental model of the state of the machine, it's automation, and their crew members.

Well designed procedures ensure that all the information is available to complete the procedure and that the procedures can be completed in a logical sequence to avoid overlaps and disruptions.

Procedures are generally categorized by the nature of the hazard that initiates the procedure: normal operations (i.e. frequent and non-hazardous), non-normal operations (i.e. infrequent but not hazardous), and emergency procedures (i.e. hazardous).

Due to the safety responsibility of government, for industries that are regulated, the procedures must be approved by the regulator. The method for approval of procedures is generally an *adhoc* approach relying on subject matter experts to inspect the procedures, and simulators to test the procedures. Due to the complexity of the conditions leading to the initiation of a procedure and the scenarios that may unfold during a procedure, it may not be possible to evaluate the procedure for all possible conditions and scenarios. Further it may not be possible to define every contingency within a procedure. In this way the procedures are only partially evaluated, leaving the crew to fill-in or improvise their way through the procedure. Every domain has a name for this improvisation such as "professionalism," or "airmanship." In all safety sensitive enterprises approval to operate the machine is achieved only after an extended apprenticeship period (e.g. a commercial pilots license is available after 1500 flight hours).

One of the key performance measures of a procedure is the Time-on-Procedure (ToP) relative to the Maximum Allowable Operational Time Window (MAOTW). Since a given procedure is conducted in various conditions both the MAOTW and the ToP exhibit a time distribution (see Figure 2). When the ToP is longer than the MAOTW (i.e. overlap in Figure 2), the procedure cannot be completed in the required time frame and the procedure is considered to be incomplete. The overlap of the distributions represents the Probability of Failure to Complete (P_FtC) and should be the key metric used by regulators and designers to assess the performance of the procedure in a way that acknowledges the stochastic nature of the process. For non-hazardous normal/abnormal procedures a probability of less than 5% may be acceptable. For emergency procedures a probability of 1% or).1% may be acceptable.

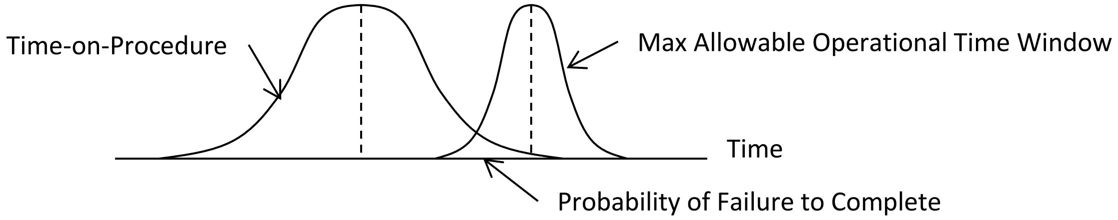


FIGURE 2: Time-on-Procedure (ToP) relative to the Maximum Allowable Operational Time Window (MAOTW). When the ToP is longer than the MAOTW (i.e. overlapping region), the procedure cannot be completed in the required time frame and the procedure is considered to be incomplete. The Probability of occurrence is key performance metric.

Human-Machine Interaction

Within the procedure there are loops of interaction that take place. These HMI- loops are characterized by three stages: (1) Observe and Orient, (2) Decide, and (3) Act (refs Boyd, Norman, Endsley).

The initiation of each HMI-loop is triggered by a sensory cue (i.e. visual, aural, tactile, or smell) or a memory cue (i.e. portion of a procedure trained and stored in Long-Term Memory). The cues may come from the environment, from the machine or from the automation. In modern "hermetically sealed" command and control centers, the cues are displays on the automation derived from environmental or machine sensors.

Next, the Decide step is made on the selection of the appropriate action(s). When the action is prompted directly by a cue (e.g. a label indicating the next action), or the decision is based on habit, the decision-making is trivial. Decisions made by habit are known as Automaticity. The decisions are part of well defined, detailed procedures, and are fast and reliable.

Rule-based decisions require the operator to used memorized IF-THEN rules from Long-Term Memory to fill-in the gaps in the procedure. In many cases, the operator will make the decision by trial-and-error (i.e. make a selection, realize it is the wrong selection and have to back-track. These decisions are slower than Automaticity and exhibit lower reliability.

Reasoning decisions are performed in the absence of instructions in the procedure. They rely on using first-principles, common-sense, and mental model building using trial-and-error to logically create the rules for the decision. These decisions are very slow and exhibit the lowest reliability.

Decisions can also be affected by information provided by the triggering event of the HMI-loop. This information is stored in Short-Term Memory (e.g. instruction for crew member) and subject to natural decay over a period of time and limits in capacity (i.e. 4+/-3).

The Act step involves manipulating the input devices on the automation or machine.

A typical procedure may involve between 7 and 50 HMI-loops. In general, the HMI-loops must be completed in the prescribed sequence (e.g. a display page must be accessed before an entry can be made). In this way delays in completing an HMI-loop ripple forward into the procedure and result in delays in completing the procedure.

Time Distributions in the HMI-loop

As described above the Procedure performance is measured by ToP. The ToP is the sum of the time to complete each HMI-loop (TtoHMILoop). The delays in performing each step in the HMI-loop can be defined based on specific characteristics of the observe and orient, decide and act steps.

Example time distributions for visual cues that trigger the HMI-loop are summarized in Table 1. HMI-loops initiated in the absence of visual cues (i.e. relying on LTM) exhibit the highest mean and variance. The time distribution is a function of the following properties of the visual cue:

(1) Visual cue absent and rely on Long-term memory
(2) Visual cue present but not in field of view

(3) Visual cue present and in field of view, but in the presence of competing cues (i.e. lost in the clutter)
(4) Visual cue present, in field of view, no competing cues, but label does not match the semantics of the procedure
(5) Visual cue present, in field f view, no competing cues, but label does match the semantics of the procedure

The distributions for the visual, aural and tactile cues are summarized in Table 1.

TABLE 1: Categories for Visual Cues and their associated Time Distributions

Visual Cue	Time Distribution N(μ, σ)
Not present, rely on LTM	N(15 secs, 21)
Visual cue present, but not in FOV	N(6.2 secs, 1.5)
Visual cue present, in FOV, but competing cues	N(5.5 secs, 1.3)
Visual cue present, in FOV, no competing cues, but not semantic match with task	N(4.2 secs, 0.7)
Visual cue present, in FOV, no competing cues, but not semantic match with task	N(0.2 secs, 0.3)

LTM = Long Term Memory, FOV = Field of View

A similar models exist for aural and tactile cues.

Time distributions for <u>Decision-making</u> are summarized in Table 2. There are two categories for of time distributions for decision-making: (1) Type of decision, and (2) Use of Working-memory. Decisions made by habit are known as Automaticity. The decisions are part of well defined, detailed procedures, and are fast and reliable. The time distribution for these decisions has one mode.

Rule-based decisions require the operator to used memorized IF-THEN rules to fill-in the gaps in the procedure. In many cases, the operator will make the decision by trial-and-error (i.e. make a selection, realize it is the wrong selection and have to back-track). The time distribution for these decisions is bi-modal. One portion of the population will make the decision rapidly as in the Automaticity. The other will have a longer distribution.

Reasoning decisions are performed in the absence of instructions in the procedure. They rely on using first-principles, common-sense, and mental model building using trial-and-error to logically create the rules for the decision. The time distribution for Reasoning has three modes.

TABLE 2: Categories and Time distributions for Decisions

Decisions	Time Distribution N(μ, σ)
Automaticity	N(0.1secs, 0.01)
Rule-based	N(0.1secs, 0.01) + N(3 secs, 1.5)
Reasoning	N(0.1secs, 0.01) + N(3 secs, 1.5)+ N(12 secs, 2.3)
+ WM overload or decay	μ + 3 secs

Decisions are also subject to a time penalty when they require use of Working Memory. When the HMI-loop is triggered by information that has to be stored in WM for longer than 7 seconds it is subject to a memory decay penalty of 3 seconds. Further, if more than 3 items are required to be held in WM, the time distribution is subject to a 3 second penalty.

Actions make a small contribution to the time distributions in the HMI-loop (Fitts, 19XX). Small additional time penalties are incurred when the device is not in range for a normal reach, the operation of the input device is confusing (e.g. unlabeled pull or push of knob), the input device is moded (i.e. works differently in different situations), and/or the input device does not acknowledge an input.

TABLE 3: Categories and Time distributions for Actions

Act	Time Distribution N(μ, σ)
Basic	N(0.1secs, 0.01)
+ Not normal reach	+ X secs
+ Input device manipulation confusing (e.g. pull of push knob)	+ Y secs
+ Input device is moded (i.e. works differently in different situations)	+ Z secs
+ Input device does not acknowledge entry	+ A secs

The time distribution for an HMI-loop is the sum of the steps defined as follows:

$$\mu_{\text{HMI-Loop}} = \sum \mu_i$$

$$\sigma_{\text{HMI-Loop}} = \text{SQRT}\ (\ \sum \sigma_i^2\ )$$

For the bi-modal distribution, the mean μ is weighted by p, p-1, the density of the two modes. For the tri-modal distribution, the mean μ is weighted by p1, p2, and p1+p2-1, the density of the three modes.

The time distribution for the Procedure is the sum of the steps defined as follows:

$$\mu_{\text{Procedure}} = \sum \mu_i$$

$$\sigma_{Procedure} = SQRT\ (\ \sum\sigma_i^2\ )$$

As above multi-model distributions are weighted by the density of the modes.

The probability of ToP exceeding the MAOTW is calculated as the Pr{ToP > MAOTW} which is equivalent to Pr {ToP – MAOTW}. For arbitrary distributions, this would be calculated as a convolution integral. For the special case that ToP and MAOTW are normally distributed and independent, then ToP – MAOTW is a normal distribution with $\mu = \mu_{ToP} - \mu_{MAOTW},\ \ \sigma^2 = \sigma^2_{ToP} - \sigma^2_{MAOTW}$. This reduces to find the probability that such a normal distribution > 0.

HMI SEQUENCE DIAGRAMS

The HMI Sequence Diagram is a Model-based approach to the specification and analysis of the human-machine interaction.

The HMI Sequence Diagram for a single operator is shown in Figure 2. The agents, shown in boxes across the top, represent the environment, the vehicle, the operator and the operator's working memory (WM) and Long-term Memory (LTM). Time increases from top down. Events that occur are shown by labeled arrows (or messages) between agents.

Figure 2 shows the HMI Sequence Diagram for a simple one HMI-loop Procedure from the field of aviation. A single pilot, sees traffic Out-the-Window (OTW). The location of the traffic is confirmed on the automation display. The decision is well rehearsed and made automatically to increase the rate of climb. The command to increase the aircraft rate of climb is made. The aircraft rate of climb increases and subsequently the traffic is no longer a threat.

The Allowable Operational Time Window (AOTW) is defined by the time distribution for the Near Mid-Air Collision (NMAC) with an aircraft at 1 o'clock with the specific relative velocities.

The time distribution for the Procedure is shown on the right in the table. This individual distributions are drawn from the Tables in Section 3 above. The Time-on-Procedure (ToP) is a Normal Distribution with mean of 26.4 secs and standard deviation of 21.8 secs. Due to the tails of the AOTW and ToP distribution overlapping, the probability of failure to complete the task is estimated at 0.15.

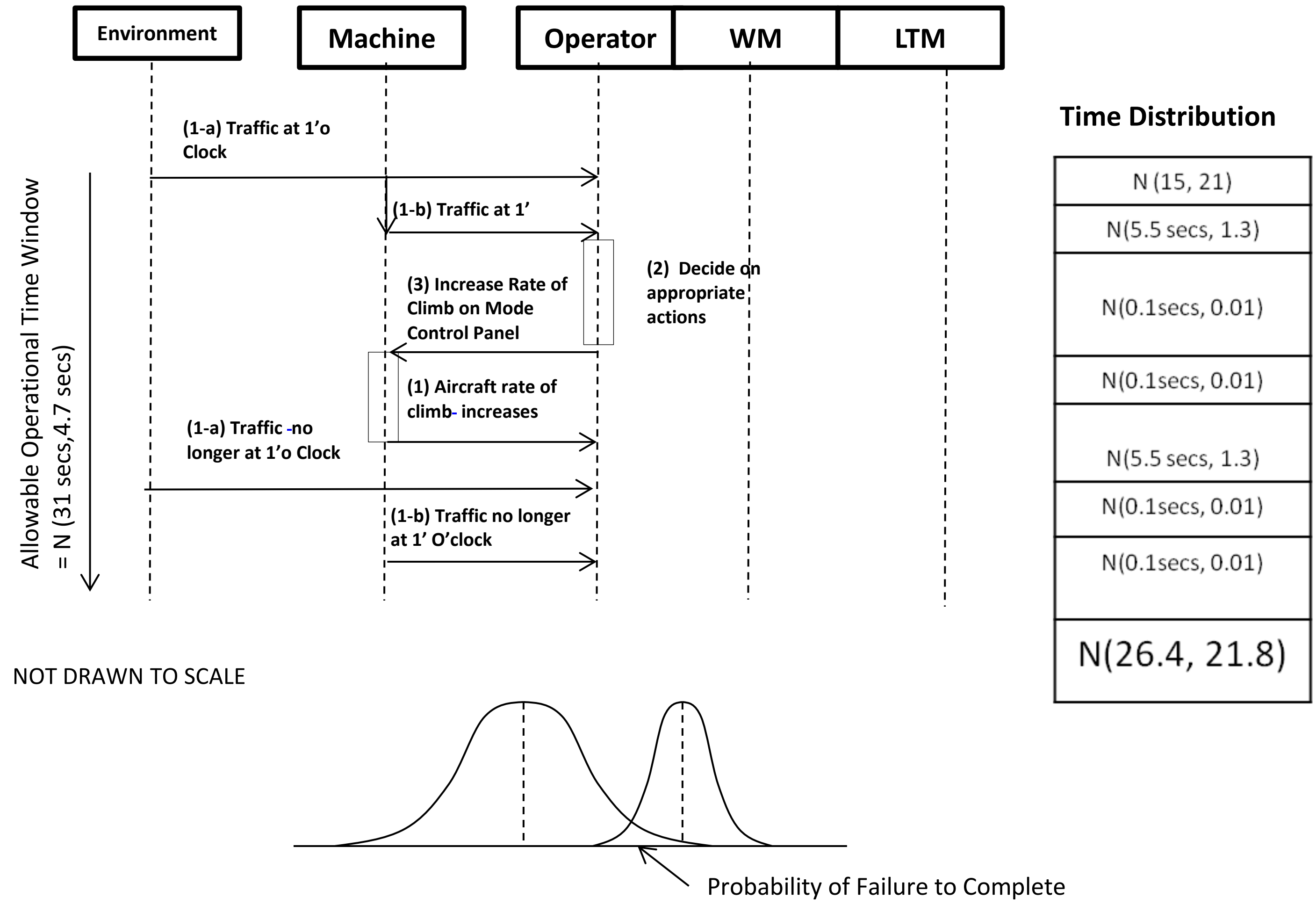


FIGURE 2: HMI-Sequence Diagram with one HMI-loop. Allowable Operational Time Window AOTW) on left, and time distributions for each event in the sequence diagram on right. The Time-on-Procedure (ToP) is shown in the bottom of the table on the right. The overlap between the AOTW and ToP is shown below.

Figure 3 illustrates an HMI Sequence Diagram that includes crew interaction between the Pilot Flying (PF) and the Pilot Monitoring (PM). The procedures is a portion of the "Takeoff" procedure for a commercial airliner. A previous procedure is completed to initiate acceleration down the runway. Once the aircraft achieves 80 knots (1) and a thrust setting (N1) greater than a specified threshold (2), the PM calls out "80 KNOTS, THRUST SET" (3) to alert the PF, who has his eyes focused out the window (not on the automation) that a key condition for takeoff has been met. The PF glances down at the automation (6,7) and confirms the conditions from the indicators on that side of the flight deck with a call-out "CHECK". Failure to achieve this condition or a discrepancy between information on the PF and PM instruments would be cause to abort the procedure. This pattern of cues and call-outs is used to double check critical conditions on the flight deck have been met.

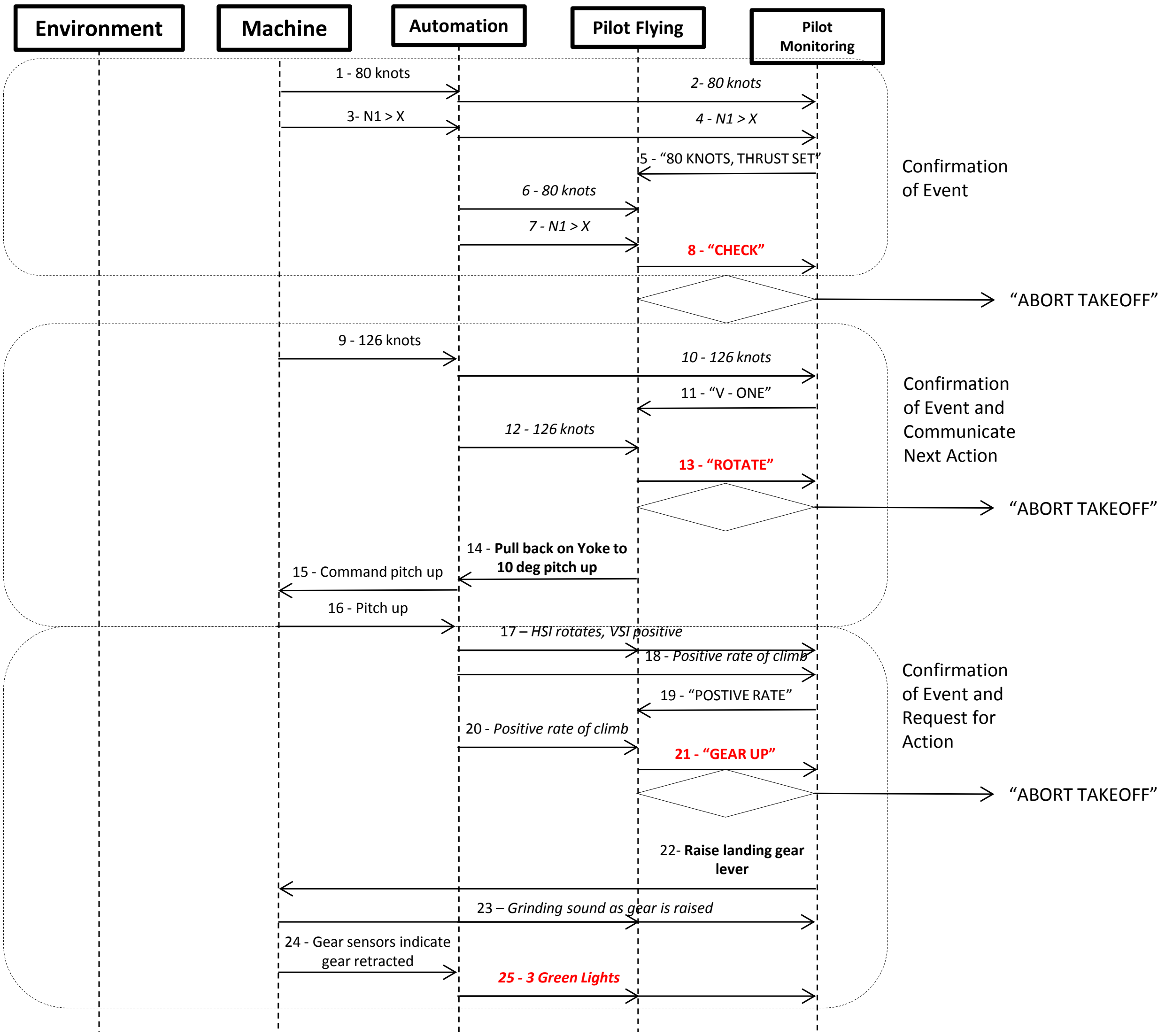


Figure 3: HMI Sequence Diagram for commercial airline "Takeoff Procedure" withcrew interaction: Pilot Flying (PF) and the Pilot Monitoring (PM).

The aircraft continues accelerating down the runway. When the aircraft reaches 126 knots (9), the PM makes this observation (10) and alerts the PF with a call-out "V – ONE" (11). The PF confirms this condition (12) and calls out the intention to perform the next critical maneuver with the call-out "ROTATE" (13). The PF pull back on the yoke (14) initiating a command from the automation to the aircraft (15) which responds by pitching up (16). This condition is reflected on the Horizontal Situation Indicator (HSI) and Vertical Speed Indicator (VSI) on the flight deck. In this sequence the call-out "ROTATE" does double duty by serving as a confirmation (i.e. "CHECK") and an indication of a next action. The call-out for confirming the pitch-up was deemed unnecessary as both crew members will feel the pitching sensation, see it out the windows and on the automation instruments.

In the next sequence, the PM observes a positive rate of climb (18) and calls-out “POSITUVE RATE” (19). This is confirmed by the PF (20), who requests the next action from the PM with a “GEAR UP” command (21). The PM raises the Landing Gear Lever (22) and the distinctive sound of the landing gear being retracted is heard by both crew members (23). Gear sensors indicate that the gear has been fully retracted (24) and the PM and PF observe three green lights above the Landing Gear Lever (25) indicating the completion of that task.

In this segment of the longer procedure, there are 4 events that are noted by the flight crew. These events result in 4 decisions all of which (in this example) are deemed critical discrepancies between instruments or failure to confirm the critical conditions would result in decision to abort the takeoff. Only two of the decisions result in physical actions on the input devices.

TABLE 4: Summary HMI metrics for the Take-off Procedure

Metric	Value
Events	4
Observe and Orient	4
Decide	4 (4)
Act	2
HMI-Loops	9
Shared Mental Model Blocks	3
Probability of Failure to Complete	

CONCLUSIONS

This paper described a formal time-based model for the design and analysis of operator procedures. The model is intended to allow the insertion of the design and analysis of the HMI and the HCI into the normal System and Software engineering process. The model is based on the traditional system engineering sequence diagram found in many CASE tools.

The method of analysis starts with the identification of operational hazards for each procedure and the Allowable Operational Time Window (AOTW) in which the procedure must be completed before the hazard is invoked (e.g. time to retract flaps before an over speed occurs). The AOTW need not be a discrete value can be represented as a distribution to take into account variance in the performance of the procedure (e.g. aircraft performance, wind).

The procedure is then defined in detail using the Observe-Orient, Decide, and Act loop. The HMI-loops include crew to crew interaction as well as crew to machine/automation interaction. The crew-crew interaction is crucial for establishing a Shared Mental Model of the evolving operational events. Each action in the loop is assigned a time distribution based on the properties of the action. The procedure is run in a Monte Carlo simulation to generate a time distribution for execution of the procedure.

When the right tail of the time distribution for the execution of the procedure overlaps with the left tail of the AOTW distribution, there is risk in completing the procedure in a timely manner. A threshold for an acceptable risk (e.g. 95%) can be established.

Disruptions (e.g. Air Traffic Control communication and other distractions) can be included in the procedure specification and used to determine the robustness of the procedure.

This method provides a way to formally document the procedure and it's HMI/HCI. Each procedure can then be assessed by formal measures of performance. Procedures can be modified and compared in a formal manner to the baseline procedure.

Limitations and Future Work

The accuracy of the analysis is dependent on the underlying time distributions assigned. On going work is collecting a data-base of these distributions from the literature, human-in-the-loop experiments, and FOQA and other live data.

Procedures are complex process to compare as they exhibit conflicting objectives. For example, a procedure with the fewest number of HMI-Loops is robust to disruption by creating a time buffer between the end of the execution of the procedure and the start of the end of the AOTW. However, additional HMI-Loops to include Shared Operational Situation (SOS) may reduce buffer time, but increase robustness through increased attention by all crew members through a Shared Mental Model. To address this conflicting objectives, Kourdali et. al. (2016) have proposed a Multi-attribute Utility (MaU) Model to capture the complexity and make the appropriate trade-offs.

ACKNOWLEDGEMENTS

The authors acknowledge the contributions of Robert Mauro (Decision Research Inc.), Julia Trippe (University of Oregon), Immanuel Barshi (NASA), Michael Feary (NASA). This research was funded by internal funds from the George Mason University Research Foundation.

REFERENCES

Boyd

Degani, A., & Wiener, E. L. (1997).Procedures in complex systems: The airline cockpit. IEEE Transactions on Systems, Man, and Cybernetics, SMC-27(3), 302-312.

Endsley, M. (1995) Towards a Theory of Situation Awareness. Human Factors. Volume 37, 1. Pages 32-64

Kourdali, H., L. Sherry (2016) A Comaprison of Takeoff Flap Retraction Procedures. In Proceedings ICNS 2016 Conference, Dulles, Virginia. April 20- 22, 2016.

Mauro, R. et. al (2016). FAA Report.

Norman

APPENDIX