Linköping Studies in Science and Technology
Dissertation No. 1249

Functional Modeling of Constraint Management in Aviation Safety and Command and Control

by

Rogier Woltjer

Department of Computer and Information Science
Linköpings universitet
SE-581 83 Linköping, Sweden

Linköping 2009
Abstract

This thesis has shown that the concept of constraint management is instrumental in understanding the domains of command and control and aviation safety. Particularly, functional modeling as a means to address constraint management provides a basis for analyzing the performance of socio-technical systems. In addition to the theoretical underpinnings, six studies are presented.

First, a functional analysis of an exercise conducted by a team of electricity network emergency managers is used to show that a team function taxonomy can be used to analyze the mapping between team tasks and information and communication technology to assess training needs for performance improvement. Second, an analysis of a fire-fighting emergency management simulation is used to show that functional modeling and visualization of constraints can describe behavior vis-à-vis constraints and inform decision support design. Third, analysis of a simulated adversarial command and control task reveals that functional modeling may be used to describe and facilitate constraint management (constraining the adversary and avoiding being constrained by the adversary).

Studies four and five address the domain of civil aviation safety. The analysis of functional resonance is applied to an incident in study four and an accident in study five, based on investigation reports. These studies extend the functional resonance analysis method and accident model. The sixth study documents the utility of this functional modeling approach for risk assessment by evaluating proposed automation for air traffic control, based on observations, interviews, and experimental data.

In sum, this thesis adds conceptual tools and modeling methods to the cognitive systems engineering discipline that can be used to tackle problems of training environment design, decision support, incident and accident analysis, and risk assessment.
Acknowledgments

During these years I have been very fortunate to work and interact with many people who have supported me in one way or another on the way to completion of this thesis. I want to express my gratitude to you all...

My advisors Professors Erik Hollnagel and Kip Smith for sharing your knowledge, experience, and enthusiasm for research in challenging discussions and support, and my former advisor Professor Sidney Dekker for your inspirational way of introducing me to this field.

My co-authors, for the pleasure of working together.

Fellow PhD students at IDA, IKP, and the HMI graduate school for interesting discussions, commenting on texts and other help, and good times.

Colleagues at IDA and IKP, as well as administrative staff at these departments, for support.

Participants in the Scandinavian functional modeling workshops, and other workshops and conferences, including the FRAMily, for challenging discussions and nice meetings.

Project partners at the Swedish Defence Materiel Administration, at the Swedish National Defence College, and in the ERASMUS project, as well as students doing FRAM projects and thesis work, for useful discussions.

Family and friends – especially my parents Trees and Jan, and Laura – for support and encouragement.

Johanna, for always being there.

Rogier Woltjer
Braskedamm, March 2009
Contents

1 Introduction 1
  1.1 Objectives and hypotheses 1
  1.2 Background 2
    1.2.1 Command and control 2
    1.2.2 Aviation safety 4
  1.3 Central issues and relevance 11
  1.4 Reading guide 12
  1.5 Appended papers 12
    1.5.1 Paper I 12
    1.5.2 Paper II 13
    1.5.3 Paper III 13
    1.5.4 Paper IV 13
    1.5.5 Paper V 14
    1.5.6 Paper VI 15
  1.6 Related work 15

2 Frame of reference 17
  2.1 Systems, functions, models, and methods 17
    2.1.1 Systems, organizations, complexity, and control 17
    2.1.2 Functions and joint cognitive systems 21
    2.1.3 Models of systems and their functions 23
    2.1.4 Methods to develop models 25
    2.1.5 Functional modeling methods 25
    2.1.6 Summary: Functional modeling 32
  2.2 Constraints 33
    2.2.1 Constraints facilitate control 33
    2.2.2 Constraints limit variety 33
    2.2.3 Constraints are behavior-shaping factors 34
    2.2.4 Constraints and affordances 34
    2.2.5 Constraints on and possibilities for action 35
    2.2.6 Actions shape constraints 35
    2.2.7 Constraints propagate 36
    2.2.8 Constraints and learning 37
    2.2.9 Constraints as a representational aid 37
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.10</td>
<td>Representing constraints in state spaces</td>
<td>37</td>
</tr>
<tr>
<td>2.2.11</td>
<td>Constraints, state spaces, and process control</td>
<td>38</td>
</tr>
<tr>
<td>2.2.12</td>
<td>Constraints open and close</td>
<td>39</td>
</tr>
<tr>
<td>2.2.13</td>
<td>Summary: Constraint management</td>
<td>39</td>
</tr>
<tr>
<td>2.3</td>
<td>Command and control</td>
<td>40</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Command and control</td>
<td>40</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Command and control (systems)</td>
<td>42</td>
</tr>
<tr>
<td>2.3.3</td>
<td>C2 as teamwork</td>
<td>43</td>
</tr>
<tr>
<td>2.3.4</td>
<td>C2 as decision making</td>
<td>44</td>
</tr>
<tr>
<td>2.3.5</td>
<td>C2 as a control process</td>
<td>46</td>
</tr>
<tr>
<td>2.3.6</td>
<td>C2 support systems</td>
<td>52</td>
</tr>
<tr>
<td>2.3.7</td>
<td>Summary: Command and control</td>
<td>55</td>
</tr>
<tr>
<td>2.4</td>
<td>Aviation safety</td>
<td>55</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Safety and accidents</td>
<td>55</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Accident models</td>
<td>58</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Approaches to safety</td>
<td>66</td>
</tr>
<tr>
<td>2.4.4</td>
<td>Analysis techniques</td>
<td>68</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Summary: Aviation safety</td>
<td>72</td>
</tr>
<tr>
<td>3</td>
<td>Methodology</td>
<td>73</td>
</tr>
<tr>
<td>3.1</td>
<td>Research settings</td>
<td>73</td>
</tr>
<tr>
<td>3.2</td>
<td>Functional exercise</td>
<td>75</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Electricity network resource management exercise</td>
<td>76</td>
</tr>
<tr>
<td>3.3</td>
<td>Microworlds</td>
<td>78</td>
</tr>
<tr>
<td>3.3.1</td>
<td>C3Fire</td>
<td>79</td>
</tr>
<tr>
<td>3.3.2</td>
<td>DKE</td>
<td>81</td>
</tr>
<tr>
<td>3.4</td>
<td>Case histories</td>
<td>83</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Norwegian Air Shuttle flight 541</td>
<td>83</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Alaska Airlines flight 261</td>
<td>85</td>
</tr>
<tr>
<td>3.5</td>
<td>Natural task environments and simulations</td>
<td>86</td>
</tr>
<tr>
<td>3.5.1</td>
<td>ERASMUS</td>
<td>86</td>
</tr>
<tr>
<td>3.6</td>
<td>Summary and progression</td>
<td>89</td>
</tr>
<tr>
<td>4</td>
<td>Results and analysis</td>
<td>91</td>
</tr>
<tr>
<td>4.1</td>
<td>Command and control</td>
<td>91</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Electricity network resource management exercise</td>
<td>91</td>
</tr>
<tr>
<td>4.1.2</td>
<td>Fire-fighting emergency management</td>
<td>93</td>
</tr>
<tr>
<td>4.1.3</td>
<td>Wargame DKE</td>
<td>98</td>
</tr>
<tr>
<td>4.1.4</td>
<td>Summary and progression</td>
<td>101</td>
</tr>
<tr>
<td>4.2</td>
<td>Aviation safety</td>
<td>102</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Norwegian Air Shuttle flight 541</td>
<td>102</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Alaska Airlines flight 261</td>
<td>103</td>
</tr>
<tr>
<td>4.2.3</td>
<td>ERASMUS ATM automation</td>
<td>105</td>
</tr>
<tr>
<td>4.2.4</td>
<td>Summary and progression</td>
<td>107</td>
</tr>
</tbody>
</table>
CONTENTS

5 Discussion 109
  5.1 Conclusions ................................................. 109
  5.2 Contributions ............................................. 111
  5.3 Continuations ............................................. 113

Bibliography 115

Index 133

Paper I 135

Paper II 153

Paper III 177

Paper IV 199

Paper V 209

Paper VI 237
Chapter 1

Introduction

This chapter defines the central objectives, issues, and research questions that this thesis addresses. The domains of application are described broadly, and scenarios of relevance to the objectives of the thesis are sketched to discuss the societal, scientific, and industrial relevance of the work presented here. This introduction concludes with an outline of the thesis presented in a reading guide, and listings of the papers appended to this thesis and related work by the author.

1.1 Objectives and hypotheses

The purpose of this research is the development and application of methods that enable the modeling (recognition, description, and representation) of constraints. The application of the methods aims to provide a new understanding of how complex socio-technical systems manage constraints in order to avoid loss of control.

In order to fulfill the purpose, this research examines the following hypotheses:

1. The performance of complex socio-technical systems is shaped by constraints, and the actions of complex socio-technical systems shape constraints in order to manage constraints.

2. Constraint management provides a basis for the analysis and improvement of (a) the safety of socio-technical systems (specifically, in aviation), (b) joint system performance and the design of support systems (specifically, in command and control), as well as (c) the training of teams in complex socio-technical systems (specifically, in emergency management).
3. Functional modeling of constraint management provides an adequate means to address constraint management in complex socio-technical systems for the purposes and domains outlined in (2).

1.2 Background

This section describes the background for the thesis, i.e., the practical problems that people have and how and why they can be explained or described as relating to functions and constraints.

1.2.1 Command and control

Emergency management has the purpose of limiting the negative consequences of harmful events, such as accidents, emergencies, and disasters. Military operations have the purpose of achieving own (political) objectives on a military battlefield while facing opposition from one or more adversaries. As a means to achieve these purposes, both emergency management and military activity employ a command and control ($C^2$) system, a combination of people and technology in a certain organization, to steer the behavior of the entire organization.

The January 2005 storm in southern Sweden

The devastating storm "Gudrun" that reached the west coast of Sweden on the 8th of January 2005 showed the need for Swedish utilities to learn how to cooperate and to train for emergency management and recovery operations. This storm is illustrative of the general command and control problems in emergency management. As a consequence of the storm, 20 people lost their lives (Renemark, 2007). Hurricane-force winds ripped countless trees out of the ground or broke them like matchsticks. Cellular phone masts were blown away. The wind and fallen timber disabled virtually all power and telecommunications in rural southern Sweden. Around 415 000 households were left without electricity or telephone service or both (Alexandersson, 2005). Some households lost electricity for as long as 45 days, around 75 million m$^3$ of timber worth 17 billion Swedish kronor fell down, and about 30 000 km of electricity line was damaged (Renemark, 2007).

In the aftermath of the storm, the utilities were unable to work together to formulate and implement a speedy response. It took several weeks for the utilities and their subcontractors to return all service to normal. Especially in southern Sweden, huge areas of forests were blown down, and the timber needed to be taken care of in order not to loose value. At the same time the fallen timber made the restoration of the electricity and telephone networks very difficult, because it blocked access to problem areas,
constraining physical movement. Resources (mainly manpower and machinery) for both timber transportation and utility restoration were scarce, posing constraints on restoration work. Additionally, many stakeholders were involved, with numerous companies, governmental organizations, and private persons being responsible and/or interested in restoring services and roads back to normal (Renemark, 2007).

The constraints between these stakeholders were a central issue in resolving the crisis. For example, the utilities had to wait for foresters to saw the timber in the right lengths, not to loose their value, government agencies didn’t share information in an effective way, and it was unclear who was financially responsible for various activities and materiel, especially concerning foreign work forces and materiel. This meant that preconditions to start the work were not met, constraining actions of individuals to start the restoration work. Existing and prepared-for cooperations and interdependencies proved to be partly insufficient, delaying the restoration of the electricity and telecommunications networks, thereby further constraining operations. As a result of creative problem solving by numerous parties, however, new functional couplings between parties were created, and constraints removed or otherwise overcome. Thus, constraints were managed through the adaptive performance of established and new functions. Training to manage constraints in various forms of exercises is paramount, as the only way to learn emergency management is through actual experience in real-life and staged situations. Paper I deals with this issue.

Fire-fighting

Another emergency management domain that is often studied and will be referred to later on is fire-fighting, or the control of wildfires. Its main aim is to constrain a fire so that life and property can be saved. Many constraints need to be overcome, related to, for example, the allocation and coordination of personnel, materiel (e.g., fire trucks, helicopters), and resources (e.g., water, fuel), and the time pressure induced by the progressing fire. Here too, command, planning, operations, and logistics form a network of interdependent functions that need to be coordinated in order to ascertain safe and efficient emergency management (e.g., McLennan et al., 2006). The devastating 2009 Australian wildfires provided many tragic examples of these challenges.

Challenges for the scientific community include how to analyze emergency management function performance, and how to design information technology to support these functions. Paper II deals with this issue.

Military activities

Military activities suffer from similar problems in the management of constraints. An army’s strategies, operations, and tactics ultimately have
the purpose of fulfilling own (political) goals through armed conflict (Von Clausewitz, 1832). This activity may be described as various functions being performed with the aim of constraining the adversary into a series of activities that are beneficial for the army’s own purposes. At the same time, the adversary aims to constrain the own army’s activities. An essential aspect of performing these functions in an appropriate way is knowing own and opponent’s forces strengths and weaknesses in order to select appropriate action (Sun Tzu, n.d.), in other words knowing the constraints on one’s own actions, and knowing how the opponents actions can be constrained. Success in strategies, operations, and tactics depends on the actions of collectives of units. These units mutually constrain each other, facilitating and limiting action. For example, artillery may provide cover for infantry, thereby both facilitating and limiting the infantry’s movement in a specific area of coverage. Interdependencies between these units in relation to military activities and goals are of decisive importance for the success and failure of military forces. Because of these interdependencies, complex endeavours of logistics are necessary to manage these constraints. Military historians have shown that these logistics (obtaining, maintaining, and transporting military resources, such as materiel and personnel), are both critical for success and difficult to manage (Van Creveld, 1977). This aspect is shared with resource- and personnel-intensive emergency management activities such as wildfire-fighting.

Currently, command and control is in a state of change. Both civilian emergency managers and military commanders seek to implement agile (quickly adapting) network-based organizations (Alberts, 2007). The drive for command and control agility stems from coordination problems inherent to the traditional military hierarchic command structures previously common in military activity (e.g., Alberts et al., 1999). The challenge to agility is exacerbated by the growing need for joint military operations with a variety of agencies (for example, rescue services, local and national governments, and private-sector agencies). Most visions of agile network-based command and control emphasize a high-performance information network that makes sensor information available to the entire command structure. Nodes in the network are foreseen to be able to communicate with all other nodes, so that informed and appropriate action can be taken locally. A key to C^2 agility is developing and updating a viable model that describes and predicts the constraints on each of the nodes’ actions, and how an adversary may be constrained to achieve the military’s goal, which may be done by modeling of the functions that these nodes perform and their interdependencies. Paper III deals with this issue.

1.2.2 Aviation safety

Commercial aviation is comparatively a very safe means of transportation. The current safety level of the aviation industry is on the order of one ac-
cident in a million or ten million flights, yet all involved in aviation safety invest great effort in making aviation even safer. Amalberti (2001) has recently argued that advancement beyond the current level of safety needs radically new models and methods than those currently used. Accidents and incidents occasionally bring to bare concerns and aspects that current theories and methods fail to address. One incident and one accident are presented here for further analysis of these theoretical and methodological shortcomings. Moreover, in addition to the lack in knowledge to address safety of currently operational systems, the demand for air transportation is expected to continue to grow in the coming decades. The current aviation system is, however, operating close to its capacity with regard to guaranteeing safe travel. In order not to let the safety level drop with increasing air traffic, new ways of organizing civil aviation, including the use of new technologies, are being developed. The third aviation safety example presents such a technology, automating part of the task of air traffic control.

Norwegian Air Shuttle flight 541

A Norwegian Air Shuttle Boeing 737-36N with callsign NAX541 was en-route from Stavanger Sola airport to Oslo Gardermoen airport. The aircraft was close to Gardermoen and was controlled by Oslo Approach. The runway in use at Gardermoen was 19R. The aircraft was cleared to descend to an altitude of 4000 ft. The approach and the landing were carried out by the co-pilot as "Pilot-Flying" (PF) and the captain as "Pilot Non-Flying" (PNF). Shortly after clearance to 4000 ft, the crew was informed that runway 19R was closed because of sweeping and that the landing should take place on runway 19L. The aircraft was instructed by air traffic control to land on 19L (a parallel runway to the left of 19R). The crew performed a quick briefing for a new final approach.

During the final approach, while the aircraft was aligned with the runway centerline (through signals of a device called the localizer) and on an appropriate glideslope (established with the help of a glideslope signal from the Instrument Landing System (ILS)) for runway 19L, the glideslope signal failed. It took some time for the pilots to recognize the glideslope failure, who had not yet switched to the air traffic control tower frequency from approach frequency after acknowledging the new frequency.\(^1\) Immediately after the glide path signal disappeared the descent rate increased to 2200 ft/min. The aircraft followed a significantly lower approach than intended and was at its lowest only 460 ft (140 m) over ground level at 4.8 nm (about 8.9 km) from the runway. The altitude at this distance from the runway should have been 1100 ft higher. The crew initiated a go-around, aborting the final approach, because the aircraft was still in dense clouds and it drifted a little from the extended runway centerline. However, the

---

\(^1\)See also the frequency change discussion in the Section on ERASMUS
crew did not notice the below-normal altitude during approach. Later a new normal landing was carried out. The executive summary of the Norwegian Accident Investigation Board (AIBN, 2004) explains that the investigation was focused on the glide slope transmission, its technical status and information significance for the cockpit instrument systems combined with cockpit human factors. The AIBN attributes the main cause of the incident to the pilots’ incorrect mental picture of aircraft movements and position. This “cause” however does not explain the accident in terms of contributing factors. The report concludes that the in-cockpit glide slope capture representation was inadequate. In addition, the report points to a deficiency in the procedure for transfer of responsibility between approach and tower air traffic control.

This incident points to deficiencies in the ability to adequately understand interdependencies between and constraints on the functions that are performed in the cockpit and the aviation system as a whole. For example, we see the propagation of constraints in the closure of a runway into the possibilities for action for an aircraft, the availability of information on information displays in the cockpit as sources of constraint on action, and the loss of the glideslope signal as an event that poses constraints on action. These constraints propagate through the network of functions that are performed by the various systems involved in guiding an aircraft through a safe final approach. Variability in functions occurs because of the effects of these constraints, so that critical functions, such as flying at an adequate altitude and glide angle towards the runway, experience variability in their performance. The constraints on action, and the consequences of actions on constraints therefore need to be understood in order to prevent such unwanted variability. Established models and methods have trouble modeling constraints, functions, and variability adequately. Paper IV evaluates an industrially established method and develops a new functional method, thereby addressing this issue.

**Alaska Airlines flight 261**

On the 31st of January, 2000, Alaska Airlines flight 261, an MD-83, crashed into the Pacific killing all 88 persons on board. Accident investigation (NTSB, 2003) revealed a wide range of human, technological, and organizational factors contributing to this tragic event. The National Transportation Safety Board (NTSB) determined that the probable cause of this accident was a loss of airplane pitch control resulting from the in-flight failure of the horizontal stabilizer trim system. The horizontal stabilizer is the wing at the back of the aircraft that enables the control of the nose-up and -down orientation of the aircraft. This entire wing can be moved to trim this orientation, letting the wing assume a desirable (“neutral”) position relative to the actual center of gravity of the aircraft over time. This wing is tilted by a system similar to a car jack, with a pin rotating with threads in an acme nut
to move the wing up and down. The accident aircraft was found to have completely worn threads in this assembly, such that the (auto-) pilots lost control of the horizontal stabilizer completely. In other words, the movement of this wing could not adequately be constrained so that the flight performance of the aircraft could be constrained to warrant safe flight.

The thread failure was caused by excessive wear resulting from Alaska Airlines’ insufficient lubrication of the jackscrew assembly. Alaska Airlines had extended maintenance intervals for lubrication and checking of the thread wear over several years of time. Each extension received the Federal Aviation Administration’s (FAA) approval. Thus, constraints on maintenance activity were loosened over a long period of time. This allowed the wear of the acme nut threads to progress to failure without detection. Numerous functions performed by various socio-technical systems were revealed to be problematic in hindsight. Functions such as certification of the design (guaranteeing that adequate constraints are in place), the (non-redundant) design itself (illustrating the lack of constraints on horizontal stabilizer movement), managerial processes at the airline and the regulator (the FAA), which were highly resource-constrained, not adequately monitoring and constraining maintenance operations, practices at the airline’s maintenance operations, interactions of constraints between design, operational, maintenance, and regulatory engineers’ work, and many more, were identified in the investigation and published in the report (NTSB, 2003). These contributing factors may be described as the inadequate management of constraints due to the complex functions involved, necessarily and continuously having to strike a balance between multiple competing goals.

The accident emphasizes the difficulty in the management of safety constraints in an environment of production and market pressure, and intransparency of interdependencies between functions, which develop over several years and even decades, that result in hazardous conditions that remain undetected and contribute to a catastrophic accident. It also shows that with existing methods of risk assessment, the aviation safety community is as yet incapable of addressing constraint management and uncovering interdependencies and long-term developments where constraints on aviation systems’ performance are degraded. This includes how couplings and interdependencies between various processes and functions emerge, disappear, and change. How the processes setting and changing constraints to safe and efficient performance can be monitored after being put into place by design and certification, to ensure safe and effective operations over long time periods, is another issue. New functional accident models and methods that are able to address these issues need to be developed, which is the central issue of Paper V.
1.2. Background

ERASMUS ATM automation

Envisioning and analyzing the potential effects of automation is an essential part of the system development process, in air traffic management (ATM) and other complex, safety-critical domains. Moreover, ATM automation is a currently active development area (e.g., Kuchar & Yang, 2000) considering the expected further increase in air traffic for the coming decades. Airlines, manufacturers, and system designers are eager to see the potential of data precision and computing capacity realized to be able to accommodate higher traffic levels while assuring safety and efficiency. The EU FP6 ERASMUS (En Route Air traffic Soft Management Ultimate System) project proposes to decrease the occurrences of aircraft conflicts by minor adjustments of their speed (Villiers, 2004). A simplified sketch of the current work of air traffic controllers and the suggested ERASMUS functionality exemplifies the concept.

The airspace is divided into many geographic sectors. After climbing to cruising altitude, aircraft fly from sector to sector until they approach their destination and initiate the descent. Air traffic control (ATC) is responsible for separations. Air traffic controllers guide aircraft from certain entry points, through the sector to certain exit points of the sector, by giving pilots instructions. ERASMUS concerns the en-route cruise phase of the flight in controlled airspace, when the aircraft is at altitude cruising towards its destination (Gawinowski et al., 2009).

Air traffic controllers continuously perform conflict search and monitoring as one of the essential functions they perform. Controllers use a radar screen where (among other information) aircraft are identified and their speed and altitude are shown. Depending on the area control center (ACC) where the controller works, various tools are available to the controller, mostly providing information on what currently goes on in the airspace and predictions of the future paths of aircraft. Typically a pair of controllers cooperate on guiding the aircraft in a sector, one communicating directly to the aircraft that are under their responsibility in the sector (called an executive or tactical controller), the other planning traffic and preparing information on aircraft that haven’t entered the sector yet (called a planning controller).

Air traffic control may be described as the management of constraints. Controllers communicate constraints to pilots, such as which heading to follow, at which altitude, and how fast to fly/descend/climb. These constraints change as the situation unfolds, clearances for headings or points on the flight plan change, sometimes aircraft speed is constrained, at other times it is not. Pilots also may affect the constraints that are put on their actions. They can for example request re-routings and altitude changes to save time and fuel. The airspace also provides constraints on aircraft, since it for example may contain military zones where commercial aviation is not allowed to fly, regulations of minimum altitudes over objects or terrain, and specified airways that aircraft need to follow. Another cate-
category of constraints are the characteristics of aircraft, as they are limited in their movement by the laws of physics, dictating restrictions in for example speed and climb and turn rates. Moreover, because of regulations for minimum distances (separation) between aircraft, constraints on one aircraft influence constraints on other aircraft. In this way, the control of aircraft through airspace is a continuous management of constraints.

Aircraft need to have a minimum lateral and longitudinal separation of 5 nm (9.3 km) or a vertical separation of 1000 ft (305 m). Depending on the usual traffic load in a sector and ACC, controllers choose to keep margins and will be comfortable to maintain minimum separations of 5-10 nm. Higher separations are inefficient, both for controllers and pilots. Again depending on the usual traffic load, detected conflicts are acted upon a few to around 10 minutes before they occur. Therefore controllers have an emerging feeling of urgency, they may wait to see if a potential conflict solves by itself or if action is needed. Controllers most often “see the solution” to the problem once they identify the potential conflict. That is, once a potential or certain conflict is identified, one or few solution(s) to the conflict are generated directly as part of a “solution library” that the controller has built up over years of training and experience, based on the type of conflict (conflict angle, climbing/descending, etc., see Flynn, 2002), aircraft speeds and characteristics, local airspace characteristics, particular and general rules that apply, etc. Controllers seldomly need to engage in more extensive “problem solving” behavior. Standard solutions to conflicts include instructing the aircraft to take a “shortcut” to the sector exit point or other waypoints on their flightplan, instructing to climb or descend to a new altitude, giving them a temporary heading or (seldomly) a speed restriction, etc. Solutions are issued as “clearances”, instructing pilots to perform a solution, with which they need to comply.

The ERASMUS system is envisioned to compensate for the higher workload of a busier airspace in the future, as by the year 2020 the number of flights is estimated to be 1.7 times the 2007 level (Gawinowski et al., 2009). A system called the ERASMUS Solver calculates minor speed changes to the aircraft in order to dilute future conflicts without creating new ones. The ERASMUS Solver detects potential conflicts in the aircraft trajectory data it processes at given time intervals (e.g., every 3 minutes), identifying potential conflicts that are candidate for speed adjustments within a preset speed change interval (dependent on the specific parameters

---

2 Requiring that the aircraft are flying between FL290 (29000 ft, 8.8 km) and FL410 (41000 ft, 12.5 km), are equipped with the right technology, and airlines and pilots follow certain procedures and manuals, to enable this separation.

3 A conflict is defined as a state in which the closest distance between the (probable) positions of an aircraft and a specific object is less than a minimum required legal separation plus a buffer (Vink et al., 1997, p. 9.3-48). A conflict is thus a state considering the expected future positions of aircraft. An actual violation of the minimum required legal separation is called a “loss of separation”. Here we will consider only conflicts between aircraft (and not between aircraft and airspace constraints). The buffer used by the air traffic controller is dependent on the specific ACC and sector in use.
that are chosen for the application, e.g. –6 to +3 percent). Based on the candidate potential conflicts, Controlled Time Over (CTO, an ATM-imposed time constraint over a point) proposals are calculated and verified by the ERASMUS Solver. The best verified CTOs are communicated to pilots for confirmation. Pilots may then examine and validate the ERASMUS solution. If the CTO is rejected or cannot be met, the flight continues as before, and the rejection is taken into account in the future calculation of CTOs by ERASMUS. If the CTO is accepted and can be met, the aircraft Flight Management System (FMS) is updated to fly according to the time constraint. Introduction of the ERASMUS system thus produces an additional source of constraint for the guidance of aircraft. Constraint management is then not only done by air traffic controllers and pilots, but ERASMUS puts constraints on aircraft movement, as well as on the possibilities for action for the air traffic controller.

Various implementation choices of ERASMUS have been considered during the project, including a fully “subliminal” application where controllers do not receive any information on ERASMUS actions. Two assumptions behind this idea are that (1) ERASMUS operates on conflicts 20 minutes ahead of time, much earlier than controllers would interfere, and (2) its speed changes are so small that controllers will not notice them and/or will not be disturbed by them (also because aircraft ground speed naturally varies somewhat due to wind changes).

Previously published studies demonstrate that the introduction of automation is rarely unproblematic from a human factors perspective. Negative consequences of automation have been reported such as lack of user acceptance, brittle performance when faced with unanticipated novelty, users’ over-reliance on the machine’s “expertise,” and biasing users’ cognitive and decision processes (e.g., Bainbridge, 1983). Also, controllers themselves are generally very concerned with safety and control centers often have a mature approach to safety and implementing new features or new systems in their daily operations (e.g., Ek et al., 2007). Paper VI in this thesis reports on the investigation of risks in relation to a possible introduction of ERASMUS in the current ATM setting. Existing methods for risk assessment have largely proven to be effective for technical systems but encounter problems when trying to address the full range of socio-technical implications of new designs. Potential interdependencies and couplings between functions that people and machinery jointly perform (such as discovering and diluting threats to safe separation between aircraft) still are extremely difficult to anticipate. Thus, new methods of risk assessment are required. Paper VI addresses this issue.

---

4ERASMUS is now seen as an input to the SESAR (Single European Sky ATM Research) programme, and has also investigated various ERASMUS applications that would fit with how the SESAR programme envisions that the ATM world would look like in the year 2020, or even further into the future. Since these are more speculations than specifications, however, the future scenarios of ERASMUS are not investigated explicitly here.
1.3 Central issues and relevance

The central issues in the systems, settings, and scenarios described above are how to discover, describe and/or assess the constraints on functional units in complex socio-technical systems, and how these systems can best manage these constraints in a dynamic high-risk environment. This issue involves understanding functions and constraints that have emerged in the past (as in accident and incident analysis and learning from exercises), that are currently valid (as in understanding the impact of these couplings on crisis managers, military commanders, pilots, or controllers in an actual situation) and that may emerge in the future (as in training, risk assessment, and accident prevention).

The societal relevance of these issues is clear from the descriptions of the cases and scenarios above. Many lives are lost, much material is lost, and generally resources are wasted because of unsafe and inefficient systems and the consequences of their use during their often long operational life. It is therefore of the utmost interest for society in general and numerous agencies, companies, and private persons specifically to invest in the development of safe and efficient processes in relation to safety-critical activities.

The scientific relevance of these issues is that there is a general consensus among scholars that current theories, models, and methods for analyzing and assessing complex socio-technical systems are still not in line with the present-day or near-future complexity of these systems. New insights and advancements are made on a regular basis, but the science of the behavior of complex socio-technical systems is still young, as are the systems themselves. Continuous improvement and striving for increased understanding is necessary. Moreover, the development of complex systems proceeds at a very high pace, new technologies are taken into use and more complex organizations are being constructed continuously. As theories and methods go hand in hand, the development of both needs to be furthered continuously to match this development in the operational reality.

The relevance for industry\(^5\) is that the only way to advance the level of safety at the operational level is through continuous analysis of how complex systems work in reality, trying to learn from past experience, and trying to anticipate problems and opportunities, based on societal demands and scientific knowledge. One important aspect in this endeavour is method development, as methods may concretely implement theories and scientific findings and enable industries to systematically work toward improvement.

This thesis aims to provide both the scientific community and industrial stakeholders with concepts and methods in order to meet the goal of

\(^5\)Industry here means companies, government agencies and non-profit organizations that are stakeholders in safety-critical work.
improving complex safety-critical systems and processes, for the benefit of a safer society at large.

1.4 Reading guide

After this Introduction, Chapter 2 of this thesis goes on to describe the theoretical frame of reference. Readers who do not wish to read the detailed treatment of (parts of) this theoretical background are referred to the last subsection of each of the four sections of Chapter 2, which summarize each section. Chapter 3 discusses the methodology that has been used in the studies that are presented, with a summary of methods in the last section of the Chapter. Chapter 4 discusses the results in a summarizing analysis of the studies. Chapter 5 concludes with the conclusions, contributions, and potential future continuations of the research presented in this thesis.

1.5 Appended papers

References and abstracts of the appended papers are presented here.

1.5.1 Paper I


This paper addresses the roles of Information and Communication Technology (ICT) in training for effective emergency management and inter-organisational coordination. Collocation can encourage the development of common ground and trust and, in turn, result in greater efficiency and effectiveness. We expect to find communication and artefact use during collocated training that cannot readily transfer to the ICT used to link distributed work settings. This expectation makes the reliance on ICT and distributed work during emergency management operations suspect. To test these claims, we observed a large-scale, real-time exercise designed to facilitate cooperation among electricity and telecommunications companies. The exercise scenario was similar to the January 2005 windstorm that left much of southern Sweden without electricity or telephone service and revealed the need for better cooperation among utility providers. The observations suggest that while collocation is clearly beneficial, a mismatch in ICT use between collocated training and distributed emergency management operations is likely to be detrimental for preparedness.
1.5.2 Paper II


This chapter describes a method for generating ecological representations of spatial and temporal resource constraints in network-based command and control, and illustrates its application in a command and control microworld. The method uses functional and goals-means task analysis to extract the essential variables that describe the behavior of a command and control team. It juxtaposes these variables in ecological state space representations illustrating constraints and regions for opportunities for action. This chapter discusses how state space representations may be used to aid decision making and improve control in network-based command and control settings. Examples show how state space plots of experimental data can aid in the description of behavior vis-à-vis constraints.

1.5.3 Paper III


A critical element to successful command and control (C²) is developing and updating an accurate and lucid model of the interdependencies between functional units, e.g., multiple platoons of artillery and tanks. Two of the challenges to this understanding are (1) the adoption of a detailed description of interdependency and the associated understanding of interdependent functions (Brehmer, 2007) and (2) the application of that description to both own and opponent forces’ opportunities and vulnerabilities to provide for agility (Alberts, 2007). This paper documents an approach to modeling functional interdependency that addresses these challenges. The Functional Resonance Analysis Method (FRAM; Hollnagel, 2004) is shown to describe the C² functions of the DOODA loop (Brehmer, 2007) and the tactical and operational functions of military activity. FRAM models are applied to own and opponent forces in a computer-based dynamic war-game (DKE) to reveal and characterize both agile and unsuccessful C² practice.

1.5.4 Paper IV


Accident models and analysis methods affect what accident investigators look for, which contributing factors are found, and which recommendations are issued. This paper contrasts the Sequentially Timed Events Plotting (STEP) method and the Functional Resonance Analysis Method (FRAM) for accident analysis and modelling. The main issues addressed in this paper are comparing the established multi-linear method (STEP) with the systemic method (FRAM) and evaluating which new insights the latter systemic method provides for accident analysis in comparison to the former established multi-linear method. Since STEP and FRAM are based on a different understanding of the nature of accidents, the comparison of the methods focuses on what we can learn from both methods, how, when, and why to apply them. The main finding is that STEP helps to illustrate what happened, whereas FRAM illustrates the dynamic interactions within socio-technical systems and lets the analyst understand the how and why by describing non-linear dependencies, performance conditions, variability, and their resonance across functions.

1.5.5 Paper V

On January 31, 2000, Alaska Airlines flight 261, an MD-83, crashed into the Pacific Ocean; after airplane pitch control was lost as a result of the in-flight failure of the horizontal stabilizer trim system jackscrew assembly’s acme nut threads (NTSB, 2003). Accident investigation revealed a wide range of human, technical, and organizational factors contributing to this tragic event, providing a case where popular linear models and methods have difficulty addressing the full complexity of the processes leading up to the accident. This paper treats each of the steps of analysis according to the Functional Resonance Accident Model (FRAM; Hollnagel, 2004), a systemic non-linear modeling method, and discusses how functional resonance occurred through the variability in functions performed by joint human, technical, and organizational systems. It thereby aims to facilitate a better understanding of how functional variability in design, certification, limited and inadequate maintenance, negligent safety culture, economic factors, and human performance together can resonate and contribute to accidents. In this way it aims to contribute to accident prevention and the engineering of more resilient complex dynamic systems.
1.5.6 Paper VI


The ERASMUS project proposes to reduce the number of aircraft conflicts by minor adjustments of their speed. Various versions of applications are under consideration, one issue being whether to inform controllers and involve pilots or to let automation act autonomously. The Functional Resonance Analysis Method (FRAM) provides a framework and a method for systematically describing and evaluating functions and performance variability. This method is used as a means to indicate and evaluate the effects and impact on controller and pilot work resulting from ERASMUS automation. Various instantiations of a partial model resulting from the application of FRAM are presented, illustrating how air traffic management automation human factors and risk assessment issues may be addressed with this method.

1.6 Related work

These articles, presentations, and reports are related to the research presented in this thesis, and/or contain earlier versions of the material presented here, but are not included as appended papers:


Chapter 2

Frame of reference

The studies presented in this thesis are based on a frame of reference that consists of theories of systems, cognition, action, and decision making, their relations to constraints, related modeling methods, and models of command and control and aviation safety. This theoretical frame of reference is described in this Chapter.

2.1 Systems, functions, models, and methods

Models of the performance of systems, of function performance, and methods of analyzing systems and their performance, form the basis for the studies presented here. These and other relevant associated concepts are described here.

2.1.1 Systems, organizations, complexity, and control

Systems and environments

A commonly adopted broad definition of a system is the definition by Hall & Fagen (1968). In their words, a system is "a set of objects together with relationships between the objects and between their attributes" (p. 81). Objects are the parts of a system. Objects can be physical such as a steering wheel or a computer, or abstract such as a variable, process, or a non-governmental organization. Attributes are properties of objects. Relationships tie the objects and attributes together, for example through causal connections, interdependencies or interactions.

A variable is a quantity or quality which may be assigned a value. A system often embodies a large number of variables, of which typically only a few are of interest and many necessarily must be ignored, as determined by an observer/experimenter. For example, a control engineer will be interested in aircraft performance variables such as thrust, pitch, or flaps set-
ting, whereas a linguist will be interested in utterances and dialogue in pilots’ communication. Ashby (1956) used the term essential variables for the variables that are to be kept within assigned limits for an organism to survive, or, one could say, for a system to be able to function.

"The state of a system at a given instant is the set of . . . values which its variables have at that instant" (Ashby, 1960, p. 16, numerical values in original). Systems behave in the sense that their state changes over time. A system that changes state over time is called a dynamic system (Ackoff, 1971). "A line of behavior is specified by a succession of states and the time-intervals between them" (Ashby, 1960, p. 20). For example, a 4-dimensional aircraft trajectory may be called a line of behavior, linking three-dimensional position states over time, as well as an emergency management team’s time-stamped phone calls during a day may be called a line of behavior. Lines of behavior thus describe system dynamics.

The environment of a system consists of all objects that (1) affect the system when their state changes or (2) are affected by changes of the system state (Ackoff, 1971; Hall & Fagen, 1968). It follows from these definitions that it can be difficult to distinguish between a system and its environment, and that systems may be nested: systems are often parts of larger systems. Open systems can be influenced by their environment, closed systems cannot (Ackoff, 1971).

As Jagacinski & Flach (2002, Chapter 1) explain, the open/closed system perspective can also be applied to the study of human behavior and cognition. Behaviorism (e.g., J. B. Watson, 1913) considered animals (including humans) to be "black box" systems with a very clear-cut system boundary between human and environment. The environment provided stimuli, and the individual reacted to these stimuli with a response. This was only a partial open systems view, as the "black box" was not investigated, and only the link stimulus → response was explored, disregarding the full system ↔ environment relation. The human mind, modeled as a "black box", was neither decomposed into subsystems.

As a reaction to behaviorism, information processing (e.g., Wickens, 1992) investigated that "black box" in relation to stimulus → response, seeing information processing steps largely in isolation, that is, as closed systems. Thus, information processing steps such as perception, memory, decision making, and motor control could be studied as isolated components of human cognition, with a tendency to study one of these components and regard the others as its environment (Jagacinski & Flach, 2002).

Ecological perspectives on cognition (Brunswik, 1955; Gibson, 1986; Neisser, 1976; Von Uexküll, 1957) recognize the flow between systems and environments in an open systems manner, and "focus on higher order properties of the perception-action dynamic, rather than on the local transfer functions of component stages" (Jagacinski & Flach, 2002, p. 4). The ecological approach focuses on that "action is linked to perception through the situation" rather than that "perception is linked to action by a brain" (Flach,
Cognitive systems engineering (Hollnagel & Woods, 1983, 2005; Woods & Hollnagel, 2006) takes such an ecological perspective, and is concerned with open dynamic systems with relative boundaries defined by their functions. For example, one could see a pilot as a closed system, not considering crew, the aircraft systems, air traffic control, weather, the airline, etc. However, to properly understand pilot actions, the pilot’s environment needs to be taken into account, and pilots and cockpit systems can be treated as a joint system influenced by and affecting its environment. Similarly, a firefighter, and even a fire-fighting team must be described as an open system, because the people they work with, the people they rescue, the smoke, the fire, which form their environment, are affected by their actions, and affect their performance.

In this thesis we are mostly concerned with a special class of systems, called purposeful systems:

A purposeful system is one which can produce the same outcome in different ways in the same (internal or external) state and can produce different outcomes in the same and different states. Thus a purposeful system is one which can change its goals under constant conditions; it selects ends as well as means. …Human beings are the most familiar examples of such systems. (Ackoff, 1971, p. 666)

Goals are defined here as sets of system states for which certain conditions are met. Ends are synonymous to goals, and means aid the system in behaving towards ends.

Organizations

Organizations can be seen as a special class of systems. First, we will define an organization from a systems perspective:

An organization is a purposeful system that contains at least two purposeful elements which have a common purpose relative to which the system has a functional division of labor; its functionally distinct subsets can respond to each other’s behavior through observation or communication; and at least one subset has a system-control function. (Ackoff, 1971, p. 670)

Organizations can be classified along many dimensions. Perrow (1984) suggests the tightness of the coupling between the components of an organization, and the complexity of interactions between these components, as criteria for the comparison between organizations.
2.1. Systems, functions, models, and methods

Complexity
Definitions of complexity are often omitted in the literature of disciplines related to CSE. Definitions and descriptions that have been posed range from the very practical to the very abstract. For example, a system may be called complex if (a) the components of a system are tightly coupled and may interact in unexpected ways (Perrow, 1984). Coupling refers to the time-dependency of a process, the flexibility of action sequences, the number of ways to achieve a goal, and the degree of operational slack in resources. Complexity of interactions refers to the number of variables and causal relations in the system’s processes and interconnected subsystems, limited substitutions, and interactions in unexpected sequences that are not easily observed or understood (Perrow, 1984). From a CSE perspective, Woods & Hollnagel (2006) state that (b) “the more intertwined the relationships between structure and function, the more complex the system operationally (and the less the system is decomposable into almost independent parts)” (p. 55). Other definitions state that systems are complex if (c) system behavior is difficult to formulate even when almost complete information about its components and their interrelations is given (Edmonds, 1999), or (d) the system needs structured methods of analysis for thorough and valid assessment (FAA, 1988). Note the difference in focus in these definitions on complexity in an epistemological sense (the description of a system is complex) or an ontological sense (the actual system is complex) (Hollnagel, 2008a; Pringle, 1951).

The characterization of complex systems into intertwined structural and functional coupling suffices here, although it is still rather imprecise and abstract. Hollnagel (2008a) suggests a more pragmatic concept, manageability or tractability, instead of complexity: “a system or a process is intractable if the principles of functioning are only partly known or even unknown, if descriptions are elaborate with many details, and if the system may change before the description is completed” (Hollnagel, 2008a, p. 7).

Control
For the definition of control we refer to the science of control, cybernetics1. Cyberneticians consider that to control a process is to steer the behavior of that process. Systems and processes are closely related in systems science and cybernetics: Ackoff (1971, p. 666) describes a process as “a sequence of behavior that constitutes a system and has a goal-producing function”. Thus, process behavior is displayed by a system. Related to purposeful systems, cybernetics pioneers Rosenblueth et al. (1943) have defined purposeful behavior as behavior that can be interpreted as goal-directed, on the basis of feedback and prediction.

1The term ‘cybernetics’ stems from the Greek κυβέρνης (kybernetes) for steersman, governor, pilot, or rudder.
2.1.2 Functions and joint cognitive systems

Cognitive systems engineering (CSE; Hollnagel & Woods, 1983, 2005; Woods & Hollnagel, 2006) addresses questions such as "(1) how cognitive systems cope with complexity, …(2) how we can engineer joint cognitive systems, …and (3) how the use of artefacts can affect specific work functions" (Hollnagel & Woods, 2005, p. 24). Related to distributed cognition (Hollan et al., 2000) and macrocognition (Klein et al., 2003), cognitive systems engineering takes an ecological view regarding the importance of context when addressing cognition.

Cognition in the CSE view is not exclusively about the cognitive processes in the individual human (cognition in the head). Cognition always needs to be addressed in terms of people and their cognitive tools, situated in an environment with a certain organization. The concept of tools ranges from post-it notes to a blind person’s cane to ‘artificially intelligent’ decision support systems. People include tools and organizations in their individual cognitive processes and these three therefore constitute the joint cognitive system. Furthermore, cognition is adaptive: people adjust to the artifacts in and the organization of the environment and adjust and organize their cognitive tools and environment to themselves.

Joint cognitive systems

CSE is concerned with systems (1) constituted of people and machines in a certain organization (2) that control processes, and (3) that are defined by their functions. A cognitive system (Hollnagel & Woods, 2005) is a system that can control its behavior, on the basis of experience, towards its goals. The term joint cognitive system means here that control is accomplished by an ensemble of cognitive systems and (physical and social) artifacts in a specified organization that exhibit goal-directed behavior. In the areas of interest to cognitive systems engineering, typically one or several persons (controllers) and one or several support systems are part of a joint cognitive system, which in a complex environment are jointly engaged in some sort of process control. The boundaries of joint cognitive systems are relative, and defined by their functions. For example, depending on the purpose of the analysis, the functions of a pilot-cockpit system may be in the focus of analysis, or the functions of a system consisting of many aircraft and air traffic controllers, or the functions of a system consisting of pilots in the cockpit and the mechanics at the maintenance department of an airline. Function is thus a central concept in the study of joint cognitive systems.

Functions

The concept of function is a concept of diverse definition in science. Based on the concept of function in biology, Mahner & Bunge (2001) suggest several aspects of function, including the difference between internal activity,
the mechanism of functioning, and external activity, the functioning or role that is displayed, the distinction between valuable activity and malfunction or dysfunction, and functions that evolve and are reproduced because of their value (adaptations). The social sciences, following Mahner & Bunge (2001), add purpose, intention, or goal to these notions of function. Technical functions, they argue, may have value for a larger system, and rather than having purpose in themselves, have an intended purpose by a designer and serve a purpose for a user, which may not be the same. For example, a hammer has no relevant internal activity of itself, its purpose is only relevant in relation to its use, and it may be used for purposes other than the designer intended. Mahner & Bunge (2001) also distinguish between a functional explanation and a functional account. Functional explanations include mechanisms explaining functions in a causal, probabilistic, or mixed fashion, whereas functional accounts describe external activity and do not go into mechanism.

Connecting these concepts of function to the earlier discussion about open and closed system descriptions of people as cognitive systems in Section 2.1.1, cognitive systems engineering is interested in the purposeful functions of cognitive systems (people), meaning in what people do and why they do it, in order to achieve which goals. CSE is interested in external activity of people, and in adaptations in the sense of functions for coping with complexity. CSE is thus interested in giving functional accounts of cognitive systems (because it does not go into cognitive mechanisms). Joint cognitive systems can be seen as social systems, of which CSE is interested in internal and external purposeful activity, in successful adaptations as well as malfunctioning and maladaptation. Regarding JCSs, CSE aims not only at providing functional accounts, but also functional explanations.

In mathematics and computer science, a function is an abstract entity that associates a given input to an output, in principle in a deterministic way (the same input always gives the same output). In this discipline a function is thus the combination of internal and external activity. In the context of systems theory, Ackoff (1971) defines the function of a system as the production of outcomes that define the system’s goals, adding the purposefulness aspect. Lind (1994) goes further by stating that functions represent the roles of a system as intended by the designer relative to the goals of the system of which it is a part, taking a technical function perspective. The value of a function in its use (which may be different from what a designer intended), adequacy in the fulfilment of goals over time, and evolution of the function through adaptation are important to CSE (Woods, 1998), and excluded by most definitions. Other critical issues for the current thesis are polytyely, the multiplicity of goals (Brehmer, 1992), which may even be con-

---

Footnote: In contrast, one may note that information processing psychology is less interested in the purposefulness of activity, and focuses on internal activity or mechanisms of functioning, including memory, attention, decision making, etc., of individual people, thereby seeking to give a functional explanation of cognitive systems.
flicting, and multiviability, multiple ways of achieving a goal (Hollnagel, 1986). The following definition of function is used here:

A function is a set of actions that a system performs or is used for, which are valuable for the achievement of a set of goals.

Woods & Hollnagel (2006) emphasize the importance of functional synthesis and functional modeling in cognitive systems engineering: "Functional syntheses provide models of how systems are adapted", and "behavior of JCSs is adapted to some purposes, potential variations, and constraints in the world of work" (p. 55). Woods & Hollnagel (2006, p. 56) discuss some essential characteristics of functional syntheses, of which functional models are a product: they are context-bound (but generic), tentative (and can be overturned or re-interpreted), emphasize multiple goals, purposes, trade-offs, and dilemmas, concern dynamic processes emphasizing change, are inherently multi-level, involve the "mutual interaction and adaptation of agent and environment (ecological)", "support projections of how systems will respond when changes are introduced", "reveal vulnerabilities for under- versus over-adaptation", and do not imply "correct", "optimal" or "best" strategies, but rather capture variations and constraints that behavior is adapted to and help in expanding the adaptive power of the JCS. Functional models, system purposes, variations, constraints, and adaptations are thus central concepts in the understanding of joint cognitive systems.

2.1.3 Models of systems and their functions

A model is defined here as "a representation through which some features of a system can be characterized or described" (where a representation is "something that stands for something else" (Palmer, 1978, referred in Jorna & Van Heusden, 2003)). A model is therefore always a simplification of a system, and can be seen as having the same status as a theory or hypothesis, where some models are more supported by empirical tests than others. In his discussion of models of man, Warr (1980) discusses some interrelated parameters that may be seen as a model’s differentiating characteristics:

**Intended reference** refers to the features of reality that the model aims to cover, and thereby to its boundaries.

**Parsimony** refers to the sparing use of concepts and assumptions and is related to the Law of Parsimony or "economy of explanation in conformity with Occam’s razor"3, the least complex explanation for an observation is preferred.

---

3Occam’s razor states that "plurality ought never be posed without necessity": In psychology, this principle is reflected in the more recent writings of Broadbent (1980) where he calls for the minimization of models of man, which is echoed in the consequent "minimal modeling manifesto" of CSE (Hollnagel & Woods, 2005).
Internal consistency refers to the agreement or absence of contradiction between the parts of a model. This characteristic is related to verification, assuring that "the actual model that has been constructed is indeed the one … intended to be built". (Miser & Quade, 1988, p. 528).

Degree of elaboration referring to the detailed description of assumptions and relationships between parts.

Heuristic value referring to a model’s influence on empirical and conceptual developments in a scientific discipline.

Actual and potential empirical support refers to the amount of supporting evidence for the validity of a model that has been collected and that may potentially be collected (i.e. the model’s potential for generating empirically testable expectations). This characteristic is related to validation, assuring that "a model is a representation of the phenomena being modeled that is adequate for the purposes of the study of which it is a part" (Miser & Quade, 1988, p. 529).

Implied research methods and types of evidence refers to the types of empirical investigations that a model entails.

Degree of quantification refers to the extent to which mathematical terms are presented and applied in a model.

Compatibility with other perspectives which we interpret as the agreement or absence of contradiction with other models.

Causal assumptions refer to the type of causal processes that a model assumes.

Warr (1980) goes on to note that it is inappropriate to prescribe any desirable level or content for each of these characteristics, and that the importance of each characteristic is subjective to (communities of) scientists.

Validation is a process that often receives great attention in evaluating and reviewing models. The validation process can however be very difficult, because of the (un)availability of data or laboriousness of the application of the model relative to the available time. Also, as models are in principle theories or hypotheses, they can never be proven, independent of the amount of empirical evidence they have been subjected to, but only illustrated for certain (sets of) cases. Especially in the social sciences, where the set of social systems to which a model might be applied is gargantuan, this makes that the validity of models is often a matter of dispute. Models of socio-technical systems, as qualitative descriptions in the social sciences, can never be proven, or validated, but can be verified or calibrated in the face of empirical data, and benefit from triangulation using various methods and data sources (Bloor, 1997). Hodges (1991) states that models that
are not easily validated or have not been validated, may be used legitimately, among other uses, as a means to improve data quality or condense data, as an aid to thinking and hypothesizing, or as an illustration of an idea.

### 2.1.4 Methods to develop models

A method is defined here as "a systematic process employed in order to achieve something". In this report we are concerned with models of socio-technical systems and their behavior, and methods with the purpose of establishing these models.

In order to evaluate and compare methods (or techniques), criteria may be specified. Shorrock & Kirwan (2002), while comparing human error identification techniques, specify the following criteria: Comprehensive- ness, structure and consistency, range of life cycle stage applicability, inter-rater reliability, predictive accuracy, theoretical articulateness and plausibility, ability to capture contextual circumstances, flexibility in levels of analysis, usefulness in generating countermeasures, resource efficiency in terms of training, analysis time, and expertise, usability (ease of use), and auditability. Some of these criteria are used later to evaluate the methods employed in this thesis.

### 2.1.5 Functional modeling methods

In engineering, functional modeling is an established way of modeling technical systems. Multilevel flow modeling (MFM) was developed by Lind (1994, 1999). MFM aims to represent goal structures and their relationships to underlying causal mechanisms in process control plants. It is concerned with processes where mass and energy are manipulated. It distinguishes between the levels of objectives, functions, and components, and may thus be called a functional modeling technique. It is however explicitly aimed at the functional analysis of technical systems, and the application of its principles to socio-technical systems falls beyond its purpose (Lind, personal communication). Other functional modeling engineering methods exist, such as value analysis and engineering (Miles, 1972). There are however not many techniques that are meant to model socio-technical system functions. Goals-means task analysis, structured analysis, cognitive work analysis, and functional resonance analysis may be called functional modeling methods for socio-technical systems, and they are treated briefly in this section.

#### Goals-means task analysis

Goals-means task analysis (GMTA) reveals the goals in the task, the specific subtasks that need to be executed to fulfill the goal, and the preconditions
that need to be met to accomplish these specific subtasks (Hollnagel, 1993). GMTA has three strengths (Hollnagel & Woods, 2005). First, it has a recursive structure of preconditions for the fulfillment of high-level goals that form the new lower-level goals. This gives it flexibility, and avoids imposing a fixed number of levels of analysis. Second, it adheres to the description of joint cognitive systems in that behavior results from goals (rather than reactions to stimuli). There is no forced function allocation of tasks to human and machine. GMTA specifies goals, not a fixed sequence of tasks, as a basis for function allocation and behavior. The GMTA thus illustrates the subgoals, tasks, and preconditions for achieving a top-level goal. In this way, the goal-directedness of the joint cognitive system is addressed. The analysis specifies the goals and how to achieve them, without the need to specify which agent (person or automated system) has to perform the tasks to achieve goals. Third, it links goals and tasks, resulting in models that are as simple as possible. The Test-Operate-Test-Exit (TOTE) unit (Miller et al., 1960) and the means-ends analysis in the General Problem Solver (GPS; Newell & Simon, 1963) are early examples of goals-means task analysis. More recently, the principles of GMTA have successfully been applied to the modeling, design, and evaluation of process control tasks (Dreivoldsø et al., 2007; Hollnagel, 1993; Hollnagel & Bye, 2000; Lind & Larsen, 1995), indicating the method’s usefulness.

Structured analysis

The structured analysis and design technique (SADT) (Marca & McGowan, 1988; Ross, 1977, 1985; Ross & Schoman, 1977), later standardized into the modeling language IDEF0 (IEEE, 1998) is an elaborate and well-defined functional modeling language. IDEF0 defines a function as "an activity that, typically, takes certain inputs and, by means of some mechanism, and subject to certain controls, transforms its input into output, . . . , a function name shall be an active verb or verb phrase" (IEEE, 1998, p. 13). Functions are thus described in terms of inputs (what is transformed or consumed by the function to produce the output), controls ("conditions that must be met before the function can produce correct output"), outputs ("what is produced by the function"), and mechanism ("the means that carry out the function") (IEEE, 1998, p. 13). The name ‘constraint’ has been used instead of ‘control’, and ‘resource’ instead of ‘mechanism’. The representation of a function in diagrams is typically a box, arrows model physical and data objects that may exist in the tangible world, and the four function descriptors as box/arrow attachments, as in Figure 2.1. Thus outputs of one function can form inputs, mechanisms, or controls of another function.

---

4According to The Merriam-Webster Online Dictionary, http://www.m-w.com/, accessed 25 June 2008, recursion means "the determination of a succession of elements (as numbers or functions) by operation on one or more preceding elements according to a rule or formula involving a finite number of steps".
SADT started as a method for designing factory management systems, but the range of possible application areas is enormous as expressed by the method’s originator Douglas T. Ross: “[it] is applicable to any interesting subject” (Ross, Foreword to Marca & McGowan, 1988, p. xiv).

Although the transformation that a function performs from input to output can be a time transformation, changing the temporal location of an input object (such as a storage function), time cannot be formally modelled in SADT. Real-time SADT extensions have though been developed (Moore et al., 1988) and discussed (Peters & Peters, 1997). SADT has been applied to various aspects of safety, such as modeling safety management systems (Hale et al., 1997), safety of manufacturing systems (Niel et al., 1993; Nowak et al., 1996), surgical incident management (Van der Hoeff, 1995), and dependability evaluation (Cauffriez et al., 2006). Various other applications exist, such as in process assessment (McGowan & Bohner, 1993) and for decision making information technology (Kupryuhin et al., 1999). IDEF0 has been applied in emergency response (Liu & Fang, 2006), and command and control processes have been modeled with SADT (Grant, 2005; Grant & Kooter, 2005; Grant et al., 2007).

Cognitive work analysis

Cognitive work analysis starts with analyzing environmental constraints and works towards cognitive constraints, at the following five levels of analysis (Vicente, 1999): Work domain, control tasks, strategies, social-organizational, and worker competencies. Constraints can thus be identified as properties of each of these levels, each in turn reducing the degrees of freedom, the possibilities for action. With each level of analysis, Vicente (1999) describes a method for analysis.

Work domain analysis (see also Naikar et al., 2005) is based on the so-called abstraction-decomposition space. There are five levels of abstraction, in decreasing abstraction functional purpose, abstract function, generalized function, physical function, and physical form. There are five levels of decomposition, in decreasing order: Total system, subsystem, function
2.1. SYSTEMS, FUNCTIONS, MODELS, AND METHODS

unit, subassembly, and component. The result of the analysis of the work domain is represented by functions in a fixed 5 times 5 level abstraction-decomposition space with the order of functions noted by arrows between the functions.

Control task analysis is based on the abstraction-decomposition space and the decision ladder. The decision ladder is basically an information processing model of human cognition with links between information processing activities and states of knowledge with so-called shortcuts, quick shunts that bypass the traditional steps of processing. Novices are said to follow the information processing steps in sequential order, experts in their skillful behavior use shortcuts.

Strategies analysis determines the strategies and information flow that experts use to solve a problem. Vicente (1999) provides no clear method for identifying strategies, but mentions descriptive field studies as the best method to elicit the strategies that experts use. The result of the analysis of the strategies is represented by information flow maps, one for each strategy.

The fourth step in cognitive work analysis is the analysis of social organization and cooperation. In this stage, workers' role allocation, organizational structure, and possibly responsibilities are mapped onto the abstraction-decomposition space with the work domain, the decision ladders describing control tasks, and the information flow maps describing strategies. The social-organizational constraints are therefore illustrated in each of these representations.

The fifth and last step in cognitive work analysis is the analysis of worker competencies. This analysis uses the skill-based, rule-based, knowledge-based (SRK) levels of cognitive control. Behavior is normally a combination of skill-based, rule-based, and knowledge-based behaviors, the distribution of which is mainly dependent on the specifics of the task and the level of expertise of the worker.

Cognitive work analysis may also be used as the basis for interface design. Such interfaces are called ecological interfaces (Rasmussen, 1999; Vicente, 1999; Vicente & Rasmussen, 1992) for their emphasis on visualizing constraints in the operators' environment or ecology (cf. Gibson, 1986), or function-based interfaces (Dinadis & Vicente, 1999; Lintern et al., 1999; Potter et al., 1992) for the method's description of functions at various levels in the abstraction-decomposition space. Ecological interfaces essentially visualize constraints to improve operator performance. Seppelt & Lee (2007) for example show that an ecological interface showing the limits of automation (adaptive cruise control in their case) may improve operator (driver) performance in several respects.

Applications of (partial) cognitive work analysis to command and control include military function allocation (Jenkins et al., 2008), agent-based modeling (M. Watson & Lui, 2003), team design (Naikar et al., 2003), interface design and decision support systems (Bennett et al., 2008; Cummings
& Guerlain, 2003), improvement of tactical decision making (Flach & Ku-
perman, 2001), and command and control work domain analysis (Chin et 
al., 1999), documenting the method’s usefulness.

Cognitive work analysis has been applied extensively to the aviation 
domain. Most applications can be found in the area of cockpit display de-
sign and flight control (e.g., Flach et al., 2003). Applications of (partial) 
cognitive work analysis to air traffic control include design of weather dis-
plays for terminal control (Ahlstrom, 2005) and explorative functional task 
and practice effects analysis (Niesse n et al., 1997).

There are however problems with this framework. Lind (2003), who 
witnessed the development of the abstraction hierarchy from close dis-
tance, discussed problems with the abstraction hierarchy. The lack of pro-
cedures to conduct an analysis is a major shortcoming. Lind (2003) fur-
thermore points to conceptual problems such as confusion and modeling 
difficulty in the use of the fixed five levels of abstraction and the five levels 
of decomposition. Some entities lie at several abstraction levels at the same 
time, a fact which the abstraction hierarchy cannot handle. System bound-
daries are also problematic with respect to the levels of decomposition (Lind, 
2003).

Functional resonance analysis method

The functional resonance accident model (FRAM; Hollnagel, 2004, see Sec-
tion 2.4.2) describes socio-technical systems by the functions they perform 
rather than by their structure, and aims to capture the dynamics of such 
systems by modeling non-linear dependencies and variability with which 
functions are performed. FRAM postulates that both normal performance 
(success) and failure are emergent phenomena that cannot be attributed to 
specific system components. Performance variability is natural in socio-
technical systems, enabling people to cope with complexity and uncer-
tainty. Thus, every function has a normal weak variability. In FRAM, 
functional resonance is the detectable signal (an undesirable event) that 
emerges from the unintended interaction of the weak variability of many 
signals. This model was coined by Hollnagel (2004) for accident model-
ing and complex system analysis purposes under the acronym FRAM, but 
the acronym has over time also become to mean functional resonance ana-
lysis method, for the method associated to the model. The steps of FRAM 
are presented here in their most basic form, as outlined in the initial book 
by Hollnagel (2004). Note that, as a method related to CSE, FRAM aims 
to fulfil the characteristics of functional synthesis by Woods & Hollnagel 
(2006) outlined in Section 2.1.2. Alterations and further developments of 
the method that have come forth from the studies presented in Papers II-VI 
are discussed in Chapter 4.
Step 1. Identifying functions

Table 2.1 describes the six aspects that a FRAM-module addresses for each function that is identified (input, output, preconditions, resources, time, and control), and Figure 2.2 presents the FRAM representation of a function. To find the FRAM-modules, one may start with the top-level goal, which may translate into the top-level function, or one may start with any function and move on to related functions.

<table>
<thead>
<tr>
<th>function aspect</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>That which the function uses or transforms</td>
</tr>
<tr>
<td>Output</td>
<td>That which the function produces</td>
</tr>
<tr>
<td>Preconditions</td>
<td>Conditions that must be fulfilled before the function can be carried out</td>
</tr>
<tr>
<td>Resources</td>
<td>That which the function needs or consumes when it is carried out (e.g. matter, energy, information, manpower)</td>
</tr>
<tr>
<td>Time</td>
<td>Time available, as a special kind of resource or constraint</td>
</tr>
<tr>
<td>Control</td>
<td>That which supervises or adjusts the function (e.g. controller, guideline, plan, procedure)</td>
</tr>
</tbody>
</table>

Table 2.1: FRAM-module description of function aspects (Hollnagel, 2004).

Figure 2.2: Hexagonal representation of function and function aspects (Hollnagel, 2004).

Step 2. Characterizing variability

Eleven common performance conditions (CPCs) are identified in the FRAM method to be used to elicit potential or actual variability:

1. availability of personnel and equipment,
2. training, preparation, competence,
3. communication quality,
4. human-machine interaction, operational support,
5. availability of procedures,
6. work conditions,
7. goals, number and conflicts,
8. available time,
9. circadian rhythm, stress,
10. team collaboration, and
11. organizational quality.

These CPCs address the combined human, technological, and organizational aspects of each function. After identifying the CPCs, the variability needs to be determined in a qualitative way in terms of stability, predictability, sufficiency, and boundaries of performance.

**Step 3. Defining functional resonance**

The output of the functional description of step 1 is a list of functions each with their six aspects. These functions may be linked together through their aspects. For example, the output of one function may be an input to another function, or produce a resource, fulfil a pre-condition, or enforce a control or time constraint. When the links between functions are found, through thorough analysis of functions and common or related aspects, these links may be combined with the results of step 2, the characterization of variability. That is, the links together with the common performance conditions specify where the variability of one function may have an impact, or may propagate. Once variability propagates, common performance conditions of other functions affect if the variability that their aspects are exposed to is damped or amplified. This analysis thus determines how a (stochastic) resonance can occur of variability across functions in the system. For example, if the output of a function is unpredictably variable, another function that requires this output as a resource or an input may be performed unpredictably as a consequence, if that function’s performance conditions allow for the variability to propagate to its output. Many such occurrences and propagations of variability may have the effect of resonance; the added variability under the normal detection threshold becomes a ‘signal’, a high risk or vulnerability.

**Step 4. Identifying barriers and indicators**
Barriers are hindrances that may either prevent an unwanted event to take place, or protect against the consequences of an unwanted event (Hollnagel, 2004). Barriers can be described in terms of barrier systems (the organizational and/or physical structure of the barrier) and barrier functions (the manner by which the barrier achieves its purpose). In FRAM, four categories of barrier systems are identified (each with their potential barrier functions):

**Physical barrier systems** block the movement or transportation of mass, energy, or information. Examples include fuel tanks, safety belts, and filters.

**Functional barrier systems** set up pre-conditions that need to be met before an action (by human and/or machine) can be undertaken. Examples include locks, passwords, and sprinklers.

**Symbolic barrier systems** are indications of constraints on action that are physically present. Examples include signs, checklists, alarms, and clearances. Potential functions encompass preventing, regulating, and authorizing actions.

**Incorporeal barrier systems** are indications of constraints on action that are not physically present. Examples include ethical norms, group pressure, rules, and laws.


Besides recommendations for barriers, FRAM is aimed at specifying recommendations for the monitoring of performance and variability, to be able to detect undesired variability. Performance indicators may thus be developed for every function and every link between functions.

**2.1.6 Summary: Functional modeling**

This section of the theoretical background has positioned this research in the stance of cognitive systems engineering, based on concepts from systems theory and cybernetics. Definitions of essential concepts have been provided. This thesis is about joint cognitive systems, defined by their functions, acting in complex settings with the aim to control processes. Models may be used to describe socio-technical systems and their performance. Methods may be employed with the purpose of establishing these models, of which some examples have been given. Evaluation criteria for these models and methods have been described.
2.2 Constraints

Cognitive systems engineering draws on several disciplines that offer perspectives on the concept of constraint. This section explores these perspectives and draws upon them to elucidate the definition of constraint. The following definition of constraint is used here:

A constraint is an actual and/or perceived feature of a specific system or its environment that defines limits and/or opportunities for the performance of that system’s functions.

2.2.1 Constraints facilitate control

Systems theory (Checkland, 1981) treats systems as hierarchically organized components at multiple levels that interact. Control is an emergent property that arises when the system components interact. Systems theory views constraint as control laws, laws that define and delimit the nature of interaction between the components of a system. Loss of control results from a lack of appropriate constraints (control laws) on component interactions. Accordingly, system design must enforce the appropriate constraints to ensure control and must continue to do so as changes occur. Interactions between hierarchically organized components form adaptive control loops at various control levels. In order to provide control, communication has to occur. Information to enforce constraints is communicated downward in the hierarchy of components; feedback to ensure that constraints have effectively been enforced is communicated upward (Checkland, 1981; Leveson, 2004).

Control is always associated with the imposition of constraints, and an account of a control process necessarily requires our taking into account at least two hierarchical levels. ... Any description of a control process entails an upper level imposing constraints upon the lower. The upper level is a source of an alternative (simpler) description of the lower level in terms of specific functions that are emergent as a result of the imposition of constraints. (Checkland, 1981, p. 87)

2.2.2 Constraints limit variety

In cybernetics, constraints are recognized to play a significant role in the control of systems (Ashby, 1956). One of the most fundamental principles of control and regulation in cybernetics is the law of requisite variety, which states that only "variety can destroy variety" (Ashby, 1956, p. 207). This means that the controlling system needs to have at least as much variety

---

Note that the definition of constraint has been revised since the publication of the appended papers.
(behavioral diversity) as the controlled system has variety, for the controlling system to be able to control the controlled system. In other words, the controller of a system needs to be a model of that system in order to control it (Conant & Ashby, 1970). Systems are said to behave in terms of a continuous series of actions, in terms of a process. Thus there are two processes going on in process control: The process to be controlled and the process of controlling. One process is used to control another process (Brehmer & Allard, 1991).

Constraint in cybernetics "occurs when the variety that exists under one condition is less than the variety that exists under another" (Ashby, 1956, p. 127). Both sets of processes have variety, which is delimited by boundaries in the form of constraints. Processes or systems that are interdependent mutually define these boundaries of constraints on variety. The processes in the environment are subject to constraints, such as the laws of gravity and thermodynamics. Similarly, the processes to be controlled put constraints on the actions of controllers. If there are constraints on the actions of the controller, the variety of the controller is constrained, giving him less opportunity to meet the variety of the process. Therefore, in order to maintain control it is important that the operator knows that there are constraints that affect the set of possible actions, and that the variety that can be met as a consequence changes accordingly. The cybernetic view agrees with the systems-theoretic view that constraints are laws that define and delimit behavior.

### 2.2.3 Constraints are behavior-shaping factors

In his discussion of the method he calls cognitive work analysis (discussed in detail in Section 2.1.5), Vicente (1999) discusses constraints in terms of behavior-shaping factors. A constraint, he says, is a relationship between, or limit on, behavior. Moreover, he states that constraints remove degrees of freedom, which are possibilities for behavior. Environmental constraints are associated with factors external to the worker. Cognitive constraints are the limits on behavior associated with the cognitive limitations of the human information processing system. The limited capacity of short-term memory (Miller, 1956) and the attention bottleneck in the processing of perceived information (see, e.g., Best, 1995; Wickens, 1992).

### 2.2.4 Constraints and affordances

Norman’s (1998) work on the design of artifacts treats constraints as the opposites of Gibson’s (1986) notion of affordances. Affordances are perceived or actual fundamental properties of artifacts/environments that determine their possible functions, with reference to a specific agent and a specific artifact/environment (Fajen & Turvey, 2003; Gibson, 1986; Norman, 1998). This means that affordances and also constraints, refer to perceived or ac-
tual possibilities for action with regard to both the properties of a specific agent and a specific artifact/environment. Similarly, Miller & Woods (1997) state that constraints can be due to the nature of the domain and task, the design of the artifacts, and to the individuals in a setting.

For example, a chair affords support and the dimensions of a well-designed chair afford sitting for most people, but probably not for all people. Conversely, constraints control the set of possible actions for specific combinations of agents and environments. Norman defines four types of ‘everyday’ constraints: Physical constraints, which limit physical operations (e.g., the shape of a chair); Semantic constraints, which rely on the meaning of the situation and people’s knowledge of the world (e.g., an arrangement of chairs for a workshop or a keynote speech); Cultural constraints, which rely on cultural conventions (e.g., the traditional ways of sitting in Europe versus Japan); and Logical constraints, which follow from logical rules (e.g., assembling an Ikea chair entails logical constraints that all parts have to be used and have to be assembled in a certain order).

2.2.5 Constraints on and possibilities for action

In daily use of the phrase ‘constraints on action’, the term possibilities for action seems its direct opposite and therefore deserves some attention here. Possibilities for action are determined by a specific agent in a specific environment. The concept of an agent’s possibilities for action in a setting encompasses both capabilities and opportunities: Capabilities are the states that a specific agent is able to obtain in a range of settings; opportunities are the states that can be obtained by any agent in a specific setting (Lind, 2000). Lind (2000) calls the intersection of opportunities and capabilities the reachability set. The reachability set is thus a function of internal constraints (capabilities) and external constraints (opportunities), similar to the kinds of constraints that are identified in cognitive work analysis (see Section 2.1.5. Lind (2000) thus focuses on how constraints shape action in terms of which states can be reached given internal and external constraints.

2.2.6 Actions shape constraints

Not only do constraints shape action in many ways, as is pointed out in various other sections, people can also influence constraints. This has been described as the difference between acting under constraints and acting upon (influencing, affecting) constraints (Huguet et al., 1996). That is, people can change the situation they are in, change the constraints, by acting on these constraints. An example is the available time, which is a constraint on any action. If people notice that the time available for a particular job is too short at the current way and pace of working, one can work harder or more efficiently (acting under constraints) or try to enlarge the time available by

---

6This follows from the principle of agent-environment mutuality (Gibson, 1986).
extending a deadline or change the requirements on the actions to be performed and cut down the amount of work (acting upon constraints). Sometimes the former is easier to attain, sometimes the latter, sometimes neither is possible, or sometimes both. Woods (personal communication) refers to this as the "stickiness" of constraints, meaning that some constraints are harder than others (violation has more serious consequences), or may be negotiable. Constraints may thus not only be changed, but also be ignored. An example of this can be found in the launch of STS-26 on September 29, 1988, the first space shuttle mission after the Challenger accident. The launch was delayed to fix the crew’s flight pressure suits, and due to lighter than expected upper atmospheric winds. The "countdown continued after a waiver of wind condition constraint was issued" (NASA, 1988), indicating that constraints may be knowingly dismissed in pursuit of goals that are not unifiable with actions adhering to constraints.

The possibility of changing or removing constraints is also reflected in the writings of management cybernetician Beer (1981), who writes about actuality, what a system is managing to do at a certain point in time, under existing constraints, capability, what a system could be doing under existing constraints, and potentiality, what the system "ought to be doing by developing [the system’s] resources and removing constraints, although still operating within the bounds of what is already known to be feasible" (Beer, 1981, p. 163).

2.2.7 Constraints propagate

The verb ‘to propagate’ means ‘to spread’, or ‘to produce offspring’. In the literature on distributed cognition and human factors, the act of propagating has been ascribed to constraints. We interpret this as the statement that constraints create new constraints or affect other constraints.

Hutchins (1991) provides an illustrative example. He observed operations on the bridge of a ship when an important navigation device had broken down. The ship’s navigation team maintained control of the ship through self-organization of their problem solving and organizational structure. The team did not have a global problem description, but were able to come to a satisfactory global solution by setting constraints on each other’s actions by presenting partial solutions and solving problems locally. Hutchins showed that a "system of mutually adaptive computational sub-parts" (Hutchins, 1991, p. 35) can solve problems efficiently in organizationally unanticipated situations by propagating constraints on each other’s actions.

A similar description of constraint propagation has been provided in the air traffic management domain (Chapman et al., 2001).
2.2.8 Constraints and learning

Learning processes may be described as processes that let the learner explore constraints, bringing their behavior closer and closer to constraints that limit behavior. Ericsson & Lehmann (1996) emphasize this for cognitive constraints. Vicente (1999) and Flach (2000) express this for both cognitive and environmental constraints. Woods (personal communication) acknowledges this idea by stating that experts, compared to novices, can be characterized by a superior understanding of constraints, and when to abide by them.

2.2.9 Constraints as a representational aid

Woods (1986) emphasizes the importance of spatial representations when providing support to controllers. Representation in representation design denotes a pictorial or symbolic abstract description of something. Representation design (Woods, 1995) and ecological interface design (Vicente & Rasmussen, 1992), offer design guidelines concerned with constraint: Systems that support cognitive work should facilitate discovery of constraints, represent constraints in a way that makes the possibilities for action and resolution evident, and highlight the time-dependency of constraints. One representation scheme for such discovery is the state space, which is discussed in the next section.

2.2.10 Representing constraints in state spaces

One way of representing constraints is through state spaces. This discussion of state spaces draws upon the definitions and descriptions in Ashby’s cybernetic investigation of the nature of adaptation (Ashby, 1960). Ashby notes that lines of behavior can be illustrated in tabular form, in graphical form (where variables’ values are plotted over time), or in Cartesian spaces where the axes represent tuples of variables. The latter have been called both state spaces and phase spaces in the literature (Port & Van Gelder, 1995; Stappers & Flach, 2004). We shall use the term state space rather than Ashby’s (1960) phase space, to avoid confusion with Moray’s (1986) description of phase plane as a coordinate space of error (a variable’s deviation from a desired value) and the first derivative of error. State space representation is a graphic method where variables are represented by axes and where states have representative points in the plane whose coordinates reflect the values of variables in that state.

Recent work in cognitive systems engineering, human factors, and ergonomics reflects a renewed interest in the state space representation. Vicente’s description of constraint spaces (in cognitive work analysis, see Sections 2.1.5 and 2.2.3) in which actors can chose many trajectories has much in common with Ashby’s (1960) discussion of state spaces. Stappers & Flach (2004) describe state spaces as a promising, richer visualization
method of the behavior of cognitive systems. They avow that it steps away from the typical box and arrow diagrams of perception, mental processing, and motor control that hinder appropriate thinking about dynamics. They describe the metaphor of state spaces as "that of the cartographer’s map, showing a landscape with roads and pitfalls, mountains and rivers, i.e., opportunities and threats, but leaving the reader’s mind free to roam and imagine developing states and possible routes through the territory" Stappers & Flach (2004, p. 825). State spaces ignore processing stages and emphasize constraints.

Constraints can be illustrated by lines in state spaces beyond which a specific opportunity is not present or a particular threat exists. Beyond specified thresholds that represent constraints, safe behavior may be jeopardized, efficiency compromised, or beneficial actions made impossible. Constraints may therefore delimit regions of opportunities for action and indicate when the achievement of goals is threatened. Paper II goes further into the theory and application of state spaces representing constraints.

### 2.2.11 Constraints, state spaces, and process control

De Jong & Köster (1974) describe chemical process control operators’ behavior of checking information sources with the help of constraints represented as thresholds in state spaces. This behavior is governed by three control principles: feedback constraint control, anticipatory constraint control, and optimizing control. In feedback constraint control, operators attempt to detect if a constraint (a safety limit on the value of a variable) has been violated, and act upon this information if this is the case. Often operators use additional safety margins. In anticipatory constraint control (De Jong & Köster, 1974), the operator tries to anticipate the future behavior of the process. The various displays of variables may allow the operator to extrapolate and anticipate early warnings of imminent trouble.

Displaying state spaces, constraints, and regions of opportunities for action may therefore enhance control. The objective with displays that present constraints is that their character goes from an abstract, formal, mathematical representation to a representation that has meaning to the perceiver, much like Gibson (1986) describes the differences between formal plane geometry and the ecological geometry of surfaces. Formal mathematical planes merely present data. The ecological geometry of the things that we perceive around us, in our environment, is meaningful to us; we know what to do when we perceive them, according to Gibson. Displays may present state spaces and constraints, and thereby regions of opportunities for action. If these displays become part of the tools that operators have access to and learn to interpret in relation to the strategies they employ, the displayed content carries a meaning that make the opportunities for action evident. They help operators to know what to do.
Only few applications of state spaces in displays have been reported. Knecht & Smith (2001) for example have developed the ‘manoeuvre space’, a constraint-based aircraft tactical display that displays collision risk in free-flight airspace. Experimental trials (Knecht, 2007) suggest that the display may be useful for pilots to maintain aircraft separation in a self-organizing manner without continuous control of air traffic control.

2.2.12 Constraints open and close

Constraints have also been recognized as an important aspect in problem solving strategies in information processing psychology. Reitman (1964) describes problem solving as getting from an initial "problem" state to a solution state (the goal). Open constraints are parameters that as yet have unspecified values. Open constraints have to be eliminated if one wants to transform an ill-defined problem into a well-defined problem, or, to get from a problem to a solution. Thus, a problem gets easier to solve when constraints are closed, simply because less options are open that need to be decided upon. This may be recognized in daily life, for example, when booking a meeting between three very busy people matching their agendas (closing constraints) subsequently leaves very few meeting dates to be chosen between. Conversely, closed constraints may have to be opened to gain access to an opportunity for action. In the example, if there are no overlapping free times in people’s agenda’s, at least one of the attendees has to reschedule a schedule entry (open a constraint) in order to find a solution to the meeting planning problem.

2.2.13 Summary: Constraint management

Constraints are essential for gaining and maintaining control, because constraints limit variety of system performance. For joint cognitive systems (JCSs) to control a process, exercising constraints on that process is paramount. Moreover, constraints shape the performance of functions of joint cognitive systems, as there are features of specific JCSs and environments that set restrictions to and opportunities for what a joint cognitive system can do. Because there can be a difference between actual and perceived constraints, representing constraints with the help of models, descriptions, and support systems, may help joint cognitive systems to take appropriate action. But not only do constraints shape action, actions may also be taken in order to open or loosen constraints, actions may shape constraints. In learning environments, participants learn about the constraints on their actions and how their actions can influence constraints. Constraint management is the overarching concept that describes the relationship between JCS functions and constraints, because for adequate function performance, JCSs need to be able to recognize the constraints on their actions, both internal within the system and external from their environment, how
their actions can influence these constraints, and how constraints can be put on the processes they aim to control. The studies, models, and methods developed in this thesis aim to facilitate the understanding of how constraint management works in practise and how it may be improved.

2.3 Command and control

Emergency management and adversarial as well as general command and control concepts, models, functions, and systems are discussed in this section.

2.3.1 Command and control

The principal goal of emergency management is to produce countermeasures to limit negative consequences of emergencies (Carter, 1991; Waugh, 1998). The aim or goal of military activity is to control the adversary, or, according to Von Clausewitz (1832, Book 1, Chapter 1), to compel the opponent to fulfil our will, ultimately through disarming the enemy. As the Introduction has sketched, both emergency management and military activity have become so complex that they need an elaborate command and control function to ascertain that the often large organizations involved in these activities can meet their goals. Emergency management describes control activity in levels of operational functions (the lowest level, or sharp end) then tactical, then strategic and higher levels. The military describes control activity with the tactical level being the lowest, then operational, then strategic and higher functions (Atkinson & Moffat, 2005). The function of command, in both domains, lays at the second lowest level.

Command and control, according to Van Creveld (1985), entails (1) gathering information on own and enemy forces, and the environment, (2) finding means to store, retrieve, filter, classify, distribute, and display information, (3) assessing the situation, (4) laying down objectives and means to achieve them, (5) deciding on action, (6) planning, (7) issuing orders and ascertain their correct understanding, and (8) monitoring the execution by means of feedback, and returning to (1). It is clear that the principles of control theory and systems theory described in earlier sections can be found in this popular definition of command and control.

Coordination between emergency management organizations, companies, and government is therefore a difficult issue that is high on the research agenda (Smith & Dowell, 2000). Similarly, a new paradigm of coordination has emerged within the command and control community, as the military is developing capabilities according to a network-based control philosophy.
Network-based command and control

The concept of network-centric command and control (Alberts et al., 1999; Alberts & Hayes, 2003; Brehmer & Sundin, 2000, 2005; Cebrowski & Garstka, 1998; Tjøstheim et al., 2004) is envisioned to enable forces to organize in a bottom-up fashion, to self-synchronize, to meet commander’s intent. This concept steps away from the traditional ‘platform-centric’ hierarchical command structure where combat power is inherently lost in top-down command-directed synchronization. Network-centric warfare envisions a high-performance information grid that makes sensor information available to the entire network. Filtering, aggregation, and presentation is envisioned to be done by computer systems, and nodes in the network are meant to communicate to any other node if necessary. There is therefore a high reliance on information technology. The focus in network-based command and control is on agility. Within the command and control community, agility has recently been defined as follows:

Agility . . . is the synergistic combination of robustness, resilience, responsiveness, flexibility, innovation, and adaptation. Each of these attributes of agility contributes to the ability of an entity . . . to be effective in the face of a dynamic situation, unexpected circumstances, or sustaining damage. Effectiveness without agility is fragility. (Alberts, 2007, p. 23)

Modeling the adversary

The major differences between the control processes of military command and control (in battle) and civilian command and control or "operations other than war" (ACT, 1995) (such as emergency or crisis management), is that the process to be controlled (the adversary):

- may try to deliberately hide its intent and goals, and
- may try to deliberately mislead other parties concerning its intent and goals, both through spreading false information and performing feints, and, most importantly,
- will try to deliberately hinder its adversaries from achieving their goals.

Although crisis and emergency management may also attempt to control highly unpredictable processes, the aspects above make the goal of military command and control particularly difficult, as the following statement by well-known 19th century strategist Moltke illustrates: "As Moltke remarked to his aides, the enemy always seemed to have three alternatives open to him and he usually chose the fourth." (Van Creveld, 1985, p. 8).
2.3. Command and control

Literature searches in the cognitive systems engineering, organizational psychology, and command and control literature have not revealed descriptive studies of how exactly commanders attempt to infer adversary intent in actual battle. There have however been attempts to formalize and implement adversary intent. The field of artificial intelligence (AI) has investigated the recognition of plans for a number of years now, and specifically some attempts to recognize or infer the intent of adversaries have been reported. The inference of adversary intent is a plan recognition problem in AI. Plan recognition algorithms roughly match a set of observations with entries in a repository of hypothetical plans (Carberry, 2001; Mulder & Voorbraak, 2003), which is very difficult to do reliably.

Another perspective is to analyze the possible future moves for the own and adversary side to perform. Game theory (see, e.g., Camerer, 2003; Nash, 1951) is concerned with the mathematical modeling of strategic interaction. It provides powerful tools for mathematical modeling of decisions in so-called turn-taking situations, where the possible outcomes of actions, their probability and utility, are combined to find the optimal move. It assumes that all parties involved are rational and act according to the maximization of expected utility, which is very demanding for people to perform. Simon noted in the fifties while comparing decision making models: "[game theory] requires of economic man even more fantastic reasoning powers than does classical economic theory" (Simon, 1959, p. 266). Game theory may also be applied to command and control, as was recently done by Brynielsson (2006). From this and other work (Bracken & Shubik, 2001) it becomes clear however that the application to complex problems such as actual command and control situations often is too laborious and complicated to implement.

2.3.2 Command and control (systems)

A command and control system may be defined as: "The facilities, equipment, communications, procedures, and personnel essential to the commander for planning, directing, and controlling operations of assigned forces pursuant to the missions assigned." (US Joint Chiefs of Staff, quoted in Builder et al., 1999, p. xiii). A systems thinking perspective of joint human-machine systems in an organization may be recognized in this definition. This way of thinking is however not that new. Adelson (1961) describes a command control center as a place where decisions are made, consisting of "the assemblage of men and machines that constitute the system" (p. 729). He stated that "command centers are typically nodes in networks constituting command control systems" (Adelson, 1961, p. 726). This suggests that the network-based environments (see section 2.3.1) envisioned today are not as conceptually revolutionary as one may think.

Adelson goes on to enumerate a large number of important factors that the "decider" must know about including the recent history of rele-
vant variables’ behavior, the present state of the world, predicted future states, including the opponents’ behavior, the own system’s and environment characteristics, alternatives for action, predicted outcomes, and objectives. Adelson (1961) even mentions war games as tools to simulate and predict processes.

There are many different definitions of command and control (C2), and it seems that the command and control community still is in search for a sound theoretical foundation of command and control7 (Lawson, 1981; Pigeau & McCann, 2002; Wohl, 1981). Because there is no agreement about a definition and a model of command and control, and because command and control systems are under continuous development, efforts of improving command and control are consequently directed at different aspects or expected future problems of the military. Diagnostic listings of actual operational problems are scarce. Wohl (1981) lists some operational problems for which decision-aiding may be needed. Two of his categories of problems are ‘battlefield perception’ and ‘force planning and commitment’. Interestingly, Wohl notes the problems of knowing both which opportunities and which constraints exist on adversary actions and on own forces’ actions.

2.3.3 C2 as teamwork

Teams are social entities composed of members with high task interdependency and shared and valued common goals (Dyer, 1984). They are usually organized hierarchically and sometimes dispersed geographically; they must integrate, synthesize, and share information; and they need to coordinate and cooperate as task demands shift throughout a performance episode to accomplish their mission. During a performance episode, team members engage in taskwork processes and teamwork processes. Individual taskwork is defined as the components of a team member’s performance that do not require interdependent interaction with other team members. In contrast, teamwork is defined as the interdependent components of performance required to effectively coordinate the performance of multiple individuals. Team performance is conceptualized as a multilevel process (and not a product) arising as team members engage in managing their individual- and team-level taskwork and teamwork processes (Kozlowski & Klein, 2000). (Salas et al., 2008, p. 541)

7C2 is used here to include the many variations of this acronym, from C2 to what could be aggregated to C6I²SAR (Command, Control, Communications, Consultation, Computers, Combat systems, Interoperability, Intelligence, Surveillance, Target Acquisition and Reconnaissance), and beyond.
Team performance taxonomies have been devised by a relatively large number of scholars, in order to capture the essence of what aspects of group work constitute team work. More than 130 models and frameworks of team performance or its aspects have been reported (see, e.g., overviews by Militello et al., 1999; Salas et al., 2008). Relevant applications include a study by Suparamaniam (2002) who evaluated team performance taxonomies in the light of emergency management. The team performance taxonomy of specific relevance here is the team performance taxonomy used in Paper I, developed, applied, and refined over several decades by Fleishman & Zaccaro (1992). The taxonomy contains seven categories of team functions. The categories are orientation, resource distribution, timing and pacing, response coordination, motivation, systems monitoring, and procedure maintenance. Paper I goes in further detail which specific functions are included in each category.

2.3.4 C² as decision making

Among the many theories of decision making, normative decision theory and naturalistic decision making are particularly applicable here for their history of being applied to command and control, and being central theories in the discussion of decision making within cognitive systems engineering.

Normative decision theory

The term decision making refers to a mental process that precedes the execution of action. The classical prescriptive model of the decision making process is expected utility theory. This theory includes the steps of definition of the goal, the generation of all possible options that lead to the goal, the assessment of the probability of success for each option, the evaluation of the utility of these options, the multiplication of probability and utility of the outcomes of the options, and the selection of the option with the highest expected utility for execution (see, e.g., Von Neumann & Morgenstern, 1953). Expected utility theory describes decision making as a well-defined sequence of mathematical operations. It is normative, in the sense that whenever it is possible to meaningfully use its mathematical rigor it should be used to determine rational choice from a set of options. Largely because of this mathematical rigor, expected utility theory has proven to be difficult to apply to dynamic complex environments. Expected utility theory assumes that numerous constraints exist in the domain in which the decisions are made. Often, goals, options, probability and utility of options, and outcomes cannot be readily identified or evaluated in actual complex dynamic work environments. The remainder of this section goes through the difficulties that expected utility theory has in describing decision making in dynamic complex environments, mainly to highlight that
the constraints that expected utility theory assumes are not in place in these environments.

Klein & Calderwood (1991) question whether goals can be isolated in the way that classical decision making theory assumes, emphasizing that it is risky to lift goals out of the larger context of other and higher-order goals. As analysis methods such as goals-means task analysis (Hollnagel, 1993) show, the fulfillment of the top goal in complex tasks involves the achievement of numerous subgoals. Research on dynamic decision making (Brehmer, 1992) illustrates the characteristic of polytely, that often not one but many, possibly conflicting, goals have to be satisfied to some extent. Many systems have the goals of maintaining high levels of efficiency and safety (or efficiency and thoroughness) simultaneously (e.g., Hollnagel, 1999, 2004). There may be priorities given for these goals, that make it easier for people to choose appropriate action, but trade-offs in goal fulfillment are often necessary and can be difficult. Also, short-term goals may conflict with long-term goals, a trade-off that may force or demand operators to make ‘satisficing’ (Simon, 1955) or ‘sacrificing’ (Woods, 2006) decisions.

Goals are often ill-defined. When one of the goals is to get as far away from a given criterion as possible (for example when trying to make profit), it may be true that the goal (maximize profit) is explicitly stated but hardly well-defined. So not only deviation-counteracting, but also deviation-amplifying behavior (Maruyama, 1968) has to be considered in the modeling of goals. When another goal is to maintain safe operation, this means that unintended harmful events such as incidents and accidents should be avoided over a long period of time. The goal is thus to prevent these events from happening. Having a non-event as a goal may also pose ill-defined goals in terms of what to do.

The generation of options is not less problematic. The generation of options in a board game may be relatively easy, but in complex environments the options are often too numerous to generate within a reasonable period of time. The characteristic of multiviability, expressing that multiple ways exist to achieve a goal (Hollnagel, 1986) seems to apply to a wide range of tasks in complex environments. In process control environments with many variables that can be controlled, or when variables can be manipulated over a wide range of values, the various ways of achieving high-level abstract goals seem practically innumerable. The generation of options in a situation and possible subsequent situations is an almost ungraspable task in such cases.

People have severe difficulties in estimating probabilities, even in non-stressful situations (Tversky & Kahneman, 1974, 1981). To make matters worse, the situations that people working in dynamic complex environments are exposed to rarely offer probability estimates of events that they have experienced before, making learning and getting a feel for probability estimation virtually impossible (Klein & Calderwood, 1991). Utility assess-
ment in the classical sense of assigning numbers expressing desirability to outcomes does not cover the many kinds of senses of implications of outcomes that expert decision makers often have. Moreover, the outcomes are often hardly separable from other outcomes, which makes it difficult to assign a utility to a single outcome of a single decision (Klein & Calderwood, 1991).

Time is a constraint on any decision making process (e.g., Volta, 1986). Normative decision theory assumes unlimited time to generate and evaluate options. In complex dynamic environments, time constrains both the evaluation of feedback and the prediction of future behavior. Another important aspect is that process control often involves parallel processes with different time scales (Brehmer & Allard, 1991).

### Naturalistic decision making

Numerous studies of decision makers in so-called "naturalistic", high-stake, time-critical, complex settings have lead a group of researchers lead by Gary Klein to propose a model for decision making termed recognition-primed decision making (see, e.g., Klein & Calderwood, 1991). The critical insight of this method is that 'decision makers’ in naturalistic settings actually “do not make decisions” and only rarely consider more than one option. The recognition-primed decision making model states that decision makers generate an option based on recognition of familiarity, evaluate whether that option will work, and implement it if this is the case. If not, an other option is generated and evaluated by mental simulation, a process that continues until a satisfactory action is found. Adequate assessment and recognition of the situation are thus the keys to successful "decision making" according to this model. In truly dynamic, complex environments of military command and control, recognition may be difficult because situations are likely to be unprecedented in their characteristics. It is thus disputable if relying on supporting recognition (e.g., through training) alone will be a sufficient strategy in supporting the commanders of the future.

### 2.3.5 C² as a control process

Besides modeling command and control as a decision-making activity, it may also be seen as a control process, as will be discussed below.

### Dynamic decision making

Dynamic decision making (Brehmer, 1992) focuses on the functions served by decision making in order to gain control, or to achieve some desired state of affairs, rather than on decisions in themselves. Furthermore, dynamic decision making involves tasks with a dynamic character. This implies that a series of interdependent decisions in real-time is required to
reach the goal, and that the state of the decision problem changes both autonomously and as a consequence of the decision maker’s actions. Dynamic decision making describes the system that the decision maker aims to control and the means for achieving control as control processes, so that one process is used to control another process (Brehmer & Allard, 1991). Microworlds, with some of the essential features observed in applied dynamic settings, form the main experimental platform that has been used to study dynamic decision making (Brehmer, 1992, 2005b), as is described further in Section 3.3.

With respect to our discussion of constraints, it is useful to discuss the related concept of ‘possibilities for action’ in terms of dynamic decision making. Brehmer (1992) notes that control theory is a useful framework in this context, since achieving control is the goal of dynamic decision making. Brehmer (1992) discusses four pre-conditions in control theory: (1) there must be a goal, (2) the system state must be observable, (3) it must be possible to affect the system state, (4) a good regulator of a system must be a model of that system (Conant & Ashby, 1970). Dynamic decision making research in Brehmer’s view “is concerned with people’s formulation of goals and models as a function of the observability and action possibilities of the system to be controlled” (Brehmer, 1992, p. 218).

Thus the question of what constitutes observability and action possibilities is a research problem. The discussion of constraints above is explicitly concerned with this problem. Constraints influence the action possibilities. Hence, it is important that a controller’s model of the system to be controlled is adapted to the possibilities for action and, in turn, to the constraints on action. To be able to know the possibilities for action, constraints or possibilities need to be observable. As was shown in the discussion of normative decision theory above, exhaustive option generation is not efficient in many complex environments with conflicting goals and multiple ways to achieve these goals. Therefore, to be able to adapt one’s model to the possibilities for action, it may be more useful to be able to clearly and more directly observe the constraints on action.

Contextual control model

For cognitive systems engineering the ability to maintain equilibrium or to be in control is paramount. Hollnagel & Woods (2005) state that being in control entails knowing what has happened and what will happen. These cybernetic concepts are incorporated in the control models of cognitive systems engineering.

The contextual control model (COCOM; Hollnagel & Woods, 2005) looks at cognition in practice, where the context of the situation determines the actions and therefore the performance of people. COCOM is based on three main concepts: Competence, control, and constructs (Hollnagel, 1998b; Hollnagel & Woods, 2005). Competence refers to what a
person can do in a given situation. Control describes how actions are chosen and executed. Constructs refer to what a person knows, including the understanding a person has of the situation. Constructs, actions, and events follow each other in the cyclical model of human action. The cyclical model of human action (Hollnagel & Woods, 2005) is the heart of cognitive systems engineering, and is based on Neisser’s (1976) perceptual cycle. This model, also called CAEC loop, basically states that the current understanding of an operator (or construct, C) directs the operator’s next action. This action (A) produces some information or feedback. This information (or event, E) modifies the operator’s current understanding, etc.: Construct → directs/controls → Action → generates → Event/feedback → modifies → Construct → directs/controls → Action → … (Hollnagel & Woods, 2005, Figure 4.2, p. 74). This does not mean that these concepts always strictly occur in this order: Event evaluation and action often occur in an intermingled or iterative fashion (Hollnagel & Woods, 2005).

The cyclical model of human action also presents the temporal constraints that joint cognitive systems must adhere to (Hollnagel, 2002). In this cycle, it takes a certain time to evaluate the events and modify understanding, a certain time for the operator to select an action, and a certain time to perform the action. All three must be completed within the available time to do so, which is the time between the point in time where action becomes necessary and the latest finishing time for the action. In most dynamic complex environments, the available time is constrained by the environment or the process to be controlled. In these cases the task at hand is force-paced rather than self-paced, and the operator is possibly forced into trade-offs. If available time is too short to evaluate events, select and perform action, the operator is temporally constrained to such an extent that the processes of event evaluation, action selection, and/or action execution have to be expedited. The opportunities for adequate reaction to feedback and anticipation of future events are thus constrained, and control may be lost. If the estimated performance time exceeds the available time, a trade-off between efficiency and thoroughness (Hollnagel, 2004) has to be made, and both human strategies of and technological systems may be used to cope with this situation (Hollnagel & Woods, 2005).

Extended control model

Also the extended control model (ECOM), focuses on control of joint cognitive systems (Hollnagel & Woods, 2005). The ECOM comprises four parallel control loops, similar to the cyclical CAEC loop, which make ECOM a multi-layered model of cognition and human action. The control principle is based on the idea that behavior towards goals, and thereby control, is a combination of feedback (compensatory) and feedforward (anticipatory) control. The control loops range from compensatory control to anticipatory
control. At the most compensatory level, the tracking loop performs immediate feedback control with a very short time cycle. The regulating loop is a combination of feedback and feedforward control at the short term. The regulating loop provides goals and criteria to the tracking loop. The monitoring loop is a feedback loop testing if goals are met, has a relatively long duration time, and provides the regulating loop with plans and objectives. Finally, the targeting loop is a feedforward loop that sets goals that form the input for the monitoring loop.

Distributed supervisory control processes

Military command and control may be modeled as a distributed supervisory control process (Woods & Shattuck, 2000), a cybernetic process (Lawson, 1981) in which the state of the process to be controlled is monitored, is compared to a desired state, and compensatory action is taken in case of a discrepancy. Command and control systems, as joint cognitive systems (Hollnagel & Woods, 2005) of people and equipment in a physically distributed organization, have the purpose of exercising control of a process that is dynamic (it changes over time) with a high potential for surprise (Woods & Shattuck, 2000) and multiple ways of achieving a goal (multivariability, see e.g. Hollnagel, 1986).

This combination of control process characteristics shapes the particular nature of information and plans in command and control. Information stems from series of observations and measurements from physically distributed observers and sensors, and information is not seldomly unreliable, uncertain, incomplete, or even contradictory. Plans and instructions must be interpreted with respect to the circumstances that govern the situation at hand, which is often not exactly the same as the situation that was anticipated at the time of plan generation (Woods & Shattuck, 2000). Plans may therefore be seen as resources for action (Suchman, 1987), rather than strict instructions that should or can be executed meticulously.

C² as constraint management

As a result of a study employing grounded theory, Persson (1997) defines command and control as “the enduring management of internal and external constraints by actors in an organization in order to achieve imposed and internal goals” (p. 131). He argues that the core of command and control is “constraint management”, and defines constraint as something that is an obstacle to action or goal-achievement. He applies this concept to factors such as resources, organization, rules, procedures, missions, time, socio-cultural factors, communication, and language. To exercise command and control, therefore, means to manage the resource-related, structural, mission-related, temporal, organizational, socio-cultural, communicative, and language-related constraints on an officer’s actions. Persson does not systematically develop how each of these constraints may be-
come manifest or may be resolved, but he does indicate the importance of the recognition, resolution, and thereby management of constraints in command and control.

The OODA loop

Boyd (1987) developed the Observe-Orient-Decide-Act (OODA) loop to explain why the American F-86 pilots outperformed the North Korean Mig-15 pilots in the Korean war. The basic loop, Observe → Orient → Decide → Act → Observe → …, basically dictates that if one is to outperform the adversary, one’s own OODA loop needs to be performed faster than the adversary’s OODA loop, in other words one needs to “out-OODA” the adversary. Boyd also developed a more elaborate OODA loop, adding a feedback link Decide → Observe, and “implicit guidance and control” links Orient → Observe and Orient → Act, and specifying the Orient function further. The OODA loop is the dominant model of military C2, albeit many (however similar) alternative models (Mayk & Rubin, 1988), and numerous criticisms of the OODA loop, among others for giving little guidance for how to improve command and control systems (Brehmer, 2005a, 2006; Grant & Kooter, 2005).

The Dynamic OODA loop(s)

Brehmer (2005a, 2006, 2007) recently developed a cybernetic model of command and control called the Dynamic OODA (or DOODA) family of loops. Brehmer identifies two major shortcomings of the OODA loop: (1) the absence of a representation of the effects of the ACT stage, rendering representation of delays in C2 impossible, and (2) the lack of detail in its description of the requisite functions for effective C2. Brehmer further identifies two major shortcomings of existing cybernetic models of C2 (which he claims also apply to the OODA loop): (1) their depiction of the C2 process as an essentially reactive process, leaving no room for initiative, and (2) their generality rendering a poor representation of the requisite functions for effective C2. The DOODA loop aims to overcome these problems. The basic DOODA loop for research in normative C2 science is the Function-DOODA loop, or F-DOODA, which represents the C2 process with the functions that need to be performed to accomplish a mission. It is illustrated in Figure 2.3.

The C2 system is embedded in a mission loop that connects C2 and environment. The C2 system comprises three functions and the sensors. The sensemaking function produces an action-oriented understanding of the mission and the situation at hand in the form of a Course of Action (COA), and takes accordingly as its inputs the mission and the data acquired by the function data collection. The planning function translates the COA into orders, which is the most important output of the C2 system. The orders are translated into military activity, which at higher levels will be lower level
DOODA loops, and at the lowest level actual moving and firing. These actions are then filtered through frictions (uncertainty and disturbances deranging plans and actions, see Von Clausewitz (1832) military theory), and result in effects, which are picked up through sensors by the data collection function. The relations between these functions are to be seen as logical relations in the form of preconditions, rather than causal or temporal relations.

Brehmer emphasizes that sensemaking is a collective process of commander and staff together testing hypotheses (explicitly or implicitly), using data, guided by the mission. In the initial (planning) phase of C² work, data will be sought to test hypotheses about if the COA will work, in the execution phase of C², data will be sought to examine if the current plan needs modification and if new COAs need to be generated. In his view there is no objective truth or ultimate understanding of the truth, the sensemaking process refines the understanding that is continually subject to revision through testing it against data. The value of the understanding can only come from its consequences, the outcome of the actions taken, and therefore what the consequences of action should be must be part of understanding, action is a component of sensemaking, according to Brehmer (2006). According to this model it is therefore very difficult or impossible to assess sense directly, since much of it is implicit, and since the outcome of actions may also be successful (towards accomplishing the mission) with incorrect understanding or unsuccessful for reasons other than truth of understanding, because of numerous factors external to the C² sys-

\[8\]This statement is in fierce contrast with the assumptions behind the Endsley (1995) model of situation awareness.
tem. Brehmer sees the COA as a potential but imperfect indicator (but not measure) of achieved sense.

2.3.6 C² support systems

For decades the trend in decision support systems has been to automatically evaluate the situation that operators are in, intelligently generate options and evaluate these options, and even to take over (part of) the selection and execution of options for action. This trend was based on a similar conception of the decision making process in these distinct steps. It is the thesis of this chapter that this perspective on decision support merely constrains the decision maker, rather than supports cognitive activity, in complex dynamic process control environments. Representing constraints in support systems, so that constraints can be recognized by operators that are in need of support, is the alternative support perspective that is championed here. Support systems that enable the recognition of constraints facilitate joint man-machine systems to control a process toward their goals.

Decision support by constraining

The classical view on decision making, normative decision theory, assumes numerous constraints on the decision environment. The related classical view on decision support as intelligent generation, evaluation, and selection of options for action constrains the decision maker more than that it supports cognitive work. This section elaborates on this argument.

Although normative decision theory has proven to be difficult to apply to dynamic, complex environments like command and control (Brehmer, 1992; Hollnagel, 1999; Klein & Calderwood, 1991), it has dominated the expert system paradigm for decision support systems and automation (c.f. Parasuraman et al., 2000). Like normative decision theory generally, decision support systems based on the option generation and evaluation paradigm can perform well only in limited, well-defined and predictable domains. The literature is replete with reports about the pervasively negative consequences of support systems based on decision theoretic designs including, but not limited to, lack of user acceptance, brittle performance when faced with unanticipated novelty, surprising operators with unexpected automation behavior, users’ over-reliance on the machine’s “expertise”, and biasing users’ cognitive and decision processes (Bainbridge, 1983; Billings, 1997; Dekker & Woods, 1999, 2002; Hollnagel, 1999; Hollnagel & Woods, 2005; Norman, 1990; Lee & See, 2004; Parasuraman & Mouloua, 1996; Roth et al., 1997; Skitka et al., 1999; Winograd & Flores, 1986; Woods, 1986).

It is not surprising that this normative decision theory-based paradigm of decision support suffers from these problems when looking at the roles of decision support system and operator. By making an automated assess-
ment of the situation, these support systems impose an interpretation layer between the situation and the operator, and thereby constrain the operator’s view of the process in terms of ‘objective’ facts that are to be made sense of by each operator individually. By automatically generating options for action, the operator is constrained to considering only these options. Moreover, the combination of automated assessment and automated option generation ensures that it is virtually impossible for the operator to challenge or add to the options generated. The automated evaluation of these options in terms of probability and utility either highly biases or confuses the operator, in which case the operator is constrained to think in terms of the given evaluation and unable to compare self-generated options because of dissimilarities in the multiplication process, if there is one at all. Automated selection of the action to be performed is the ultimate constraint on the operator’s action: The operator is given only one alternative. The operator is thus increasingly constrained in the ability to make sense of the situation, assess it, and generate the most appropriate action at each of these steps of intelligent decision support, or levels of automation. The operator is constrained more by the decision support system than by the process to be controlled or by the environment in which the task is to be performed. Rather than facilitating the operator’s efforts at control, the decision support system controls the operator.

Constraint management support

Hollnagel & Woods (2005) note that especially the intelligent evaluation of alternatives, and choice of most appropriate action given performance conditions (constraints), is difficult for machines, because they are bad at finding appropriate criteria for evaluation. The complexity of goals and existence of multiple ways to achieve these goals, treated later in this text, makes evaluation even more difficult and dependent on the specific context. These reflections result in the CSE view that states: "From the joint cognitive system perspective, the intelligence in decision support lies not in the machine itself (the tool), but in the tool builder and in the tool user" (Woods, 1986, p. 173). This, in combination with the work of Neisser (1976) and Klein & Calderwood (1991), among others, makes that support for people doing their work should focus on assuring a satisfactory understanding of the situation and an effective execution of action toward the goals of the joint cognitive system.

In the many tasks where the option generation and evaluation paradigm of intelligent decision support systems seems inappropriate, other ways of supporting the joint cognitive system to gain and maintain control may be found. For the domain of air traffic control, Dekker & Woods (1999) mention an alternative form of supporting controllers: "In one situation, controllers suggested that telling aircraft in general where not to go was an easier (and sufficient) intervention than telling each individual where
to go” (p. 94). This statement suggests that operators may necessarily and sufficiently be informed about the constraints on their actions, rather than all possible actions or the supposedly best option. Fundamental similarities between the air traffic control and command and control tasks indicate that this idea may transfer well to the command and control domain. The approach to decision support that is investigated here is to support control through the recognition of constraints.

Thus, instead of support systems that pose additional constraints on the decision makers, a useful strategy would be for support systems to present characteristics of the environment and of the process to be controlled in relation to the constraints on action, in order to support control and the ability to make trade-offs in meeting multiple conflicting goals within available time.

When aiming to support control, which may be described as a combination of compensatory and anticipatory control loops, one may suggest that feedback and feedforward capabilities should be supported. In combination with constraint management support, this could, theoretically, lead to support for compensatory and anticipatory constraint management control, similar to the concepts developed by De Jong & Köster (1974), discussed in Section 2.2.11. Generally there is consensus on that feedback is important, in human-computer interaction (e.g., Shneiderman, 1998), human factors (e.g., Norman, 1998), and CSE, and feedback is the central concept in most interfaces.

Supporting anticipation is though not as easily implemented, mainly because of the uncertainty of the future. The demand for anticipatory aid has been articulated before. Ackoff for example reports on managers commenting the following about decision support: "We do not need your help in deciding what to do. Just tell us what is going to happen and we’ll take care of the rest" (Ackoff, 1983a, p. 59). Woods has described a similar concern as 'Avery’s Wish', as a response to a usability engineer’s query about the functionality of an artifact to be designed: "I yearn for a system that shows me an image of what the system and the room are going to look like, ten minutes from now" (Woods, 2001, Laws §2). The idea behind these concerns is that when future states of a process to be controlled can be predicted, one can prepare for anomalies and avoid the loss of control.

There is however an obvious danger in prediction and associated preparation. Predictions are often difficult to make reliably, and if preparations rely and focus too much on false predictions control may be lost. This dilemma has been researched at least since the 1960s in supervisory control, where predictive displays have been evaluated for various tasks and with varying success (see, e.g., Warner, 1969). This may have effects that are worse than if preparation was not focused on the predicted problems. In a sense, the feedforward information is used as input to control actions in the same manner as feedback control, as less reliable feedback about the future.
Elaborating on this trade-off, Ackoff then points to another dilemma: “The more accurately one can predict, the less effectively one can prepare; and the more effectively one can prepare, the less need there is for prediction” (Ackoff, 1983a, p. 59). To summarize, there is a danger in preparation without recognizing that predictions may be wrong, and predictions and preparations are not always equally feasible or necessary. This may result in an intermediary decision support concept that relies on people’s ability to predict when provided accurately with feedback on the process being controlled.

2.3.7 Summary: Command and control

Command and control (C²) processes are used to steer emergency management or military activity, in order to accomplish the goal of constraining a process to be controlled, such as the restoration process of an electricity network, a developing wildfire, or activity of a military adversary. The process to be controlled is often highly unpredictable, and command and control may be a complex task. Thus a complex command and control system, consisting of a number of human-machine systems, is often necessary in emergency management and military command and control. In order to address the functions that are involved in C², various perspectives on command and control have been outlined, such as the perspective of C² as an effort of teamwork, C² as a decision making task, and C² as a process control task, as well as C² as a process of constraint management. Various models within each of these perspectives have been outlined, as well as the consequences of these perspectives for the design of (technical) support systems for command and control.

2.4 Aviation safety

Many definitions and descriptions of safety exist, and accordingly there is little consensus on the concept. Some definitions and models are treated here to sketch the pillars and breadth of the concept. Although portions of the thesis focus on aviation safety, research about safety generally in complex socio-technical systems is discussed interchangedly with the specifics of aviation, where applicable.

2.4.1 Safety and accidents

ICAO defines safety in its safety management manual as follows: “Safety is the state in which the risk of harm to persons or of property damage is reduced to, and maintained at or below, an acceptable level through a continuing process of hazard identification and risk management.” (ICAO, 2006, p. 1-1). Thus, safety can be seen as freedom from unacceptable risk.
This introduces a threshold of acceptability of risk, which is why hazard identification and risk management are needed: These processes determine which risks there may be and if they are acceptable. Thus, risk needs to be explained.

ICAO describes risk as two-dimensional: "Evaluation of the acceptability of a given risk associated with a particular hazard must always take into account both the likelihood of occurrence of the hazard and the severity of its potential consequences" (ICAO, 2006, p. 4-1, emphasis as in original). Likelihood and severity are thus to be defined carefully if the definition of risk has any practical applicability. Acceptable risk is defined in aviation, for example by the Federal Aviation Administration (FAA) in Federal Aviation Regulations (FAR) section 25.1309. The current document describing various means for compliance with these regulations (FAA, 1988) states that minor failure conditions can be probable, major failure conditions must be improbable, and catastrophic failure conditions must be extremely improbable. Severity is defined qualitatively and likelihood quantitatively in some detail. Minor failure conditions do not significantly reduce aircraft safety and are within crew capabilities. Major failure conditions entail for example significant or large reduction in safety margins or functional capabilities of aircraft or crew, significant increase in crew workload impairing efficiency or effectiveness, discomfort to or adverse effects on occupants. Catastrophic failure conditions prevent continued safe flight and landing. Probable means anticipated to occur once or more often during the entire operational life of each aircraft, or a probability on the order of \( p > 10^{-5} \). Improbable means anticipated not to occur during operational life of a single random aircraft of a type, but occasionally during the life of an aircraft type, or \( 10^{-9} < p \leq 10^{-5} \). Extremely improbable means not anticipated to occur to any aircraft of a type during that type's operational life, or \( p \leq 10^{-9} \).

Important factors in determining if risks are "as low as reasonably practicable" (ALARP) are cost and benefit. Taking risks has certain benefits, but may involve cost if accidents occur, yet mitigating risk also incurs costs and benefits.

The terms of accidents and incidents also need to be defined. ICAO uses the following definitions, focusing on the consequences of an event:

An accident is an occurrence during the operation of an aircraft which entails:

1. a fatality or serious injury;
2. substantial damage to the aircraft involving structural failure or requiring major repair; or
3. the aircraft is missing or is completely inaccessible.

\(^9\)Similar regulations are determined in Europe by the Joint Aviation Authority (JAA) in JAR-25.1309.
An incident is an occurrence, other than an accident, associated with the operation of an aircraft which affects or could affect the safety of operation. A serious incident is an incident involving circumstances indicating that an accident nearly occurred. (ICAO, 2006, p. 4-3, emphasis as in original)

Looking back at the Introduction, these definitions make the Alaska 261 case an accident, and the Norwegian 541 case a serious incident.

Weick has described safety as a "dynamic non-event." Thus, safety is concerned with events (or occurrences, in the ICAO wording above) that should not happen, and something is safe when accidents do not occur. Moreover, safety is dynamic, that is, it is an event prevented from happening over time. The difficulty of measuring safety is made clear from this definition: If we want to measure safety, how to measure events that do not happen? And even if non-safety may be measured in terms of measuring the events that we do not want to happen, how should these be quantified (are they all equally serious indications for non-safety), and to what other quantification should they be compared? As Hollnagel describes the comparison of a company CEO with the captain of a ship, using control-theoretic terms, "in the case of safety [it] is rarely clear where one is to steer (target), how course and speed can be changed (control options), how the propulsion takes place (process), or how to measure the position (performance indicators)" (Hollnagel, 2008b, p. 77).

Rochlin (1999) has described safe operation as a social construct, meaning that safety is not something that exists or does not in a system, but needs to constantly be created by people in the system. The field of high-reliability organizations (HRO) has studied the strategies and behavior patterns of people in safety-critical domains. They identified, among other factors, operator’s preoccupation with failure, reluctance to simplify, sensitivity to operations, commitment to resilience, and deference to expertise to be critical strategies in creating safety (Roberts, 1990; Roberts & Bea, 2001; Rochlin, 1999; Weick & Roberts, 1993; Weick & Sutcliffe, 2007). The following statement about the responsibility for safety by ICAO illustrates the actors involved in the collective construction of safety in aviation:

The responsibility for safety and effective safety management is shared among a wide spectrum of organizations and institutions, including international organizations, state regulatory authorities for civil aviation, owners and operators, service providers for air navigation services and aerodromes, major aircraft and power plant manufacturers, maintenance organizations, industry and professional associations, and aviation education and training institutions. (ICAO, 2006, p. 2-1)

Referring to the definitions of systems and organizations in Section 2.1.1, the entire aviation system can thus be seen as being responsible for
creating and maintaining aviation safety. Human and technical aspects of safety have been the focus of safety for at least a century now. During the last decades, organizational and management factors have been stressed (Flin, 2003; Reason, 1997), as the responsibility statement above clearly indicates to be essential.

### 2.4.2 Accident models

Accident models are conceptions of how accidents occur, often implicit in the minds of accident investigators or in their guidelines. As these conceptions determine what accident investigators look for during investigation, and thereby which contributing factors are found (Lundberg et al., in print) and prescribed to be fixed, accident models are important to recognize, identify, and reflect upon when discussing system safety and the control of risk. Early accident models (still in use), often attribute accidents to a root cause. However, even for domains that are less complex than the ones examined here, the root cause concept is problematic. Compare the concept of a single root cause to the complex courses of events that precluded for example the Alaska 261 accident and the conclusion can only be that the concept of root causes is not useful when talking about complex socio-technical systems. Other accident models are thus needed to adequately model and thereby understand accidents. The concept of "cause" is psychologically and philosophically problematic (see, e.g., White, 1990, for an overview), basically since David Hume’s apt finding that causation is not observable, but must be inferred.

Safety science has come up with a wide range of accident models since the 1930s. Accident models have been classified by Hollnagel (2004) into simple linear, complex linear, and systemic accident models. Simple linear models, such as the Domino model (Heinrich, 1931), model socio-technical systems by their physical and organizational structure and focusing on linear cause-effect relationships between independent components. Complex linear or epidemiological models, such as the Swiss Cheese model (Reason, 1997), also decompose socio-technical systems by their structure and consider linear relationships, but of interdependent components. Latent conditions (e.g., fatigue, bad design, management production pressure) affect how active failures (such as unsafe acts or human error) can sneak through deficiencies in the barriers (e.g., safety regulations and procedures, supervisor checks) in the system to cause an accident. As scholars have recently argued, the linear models of accident causation, and the view on safety as a hunt for human error, do not suffice to be able to model and understand the complex nature of contemporary accidents, and more "systemic" models of accidents and safety are necessary (Amalberti, 2001; Dekker, 2004; Hollnagel, 2004; Leveson, 2004; Reason et al., 2006; Rochlin, 1999; Woods & Cook, 2002).
Systemic models treat safety as an emergent property of systems as a whole, and try to find system-wide vulnerabilities rather than flawed system components. The models of systems-theoretic accident model and processes (STAMP) and functional resonance accident model (FRAM) as two such systemic approaches.

**Models of human error in hindsight**

Some accident models, such as the Domino model (Heinrich, 1931) and the Swiss Cheese model (Reason, 1997), basically state that an unsafe act by an operator is the start of a sequence of hazardous events. This view states that error is the result of flawed decision making, imperfect information processing, fatigue, carelessness, inattentiveness, recklessness, noncompliance with procedures, complacency, and the like (Helmreich, 2000). This easily leads to the view that human error is a common source of accidents. This view is reflected in the frequent reporting that about 70 percent of aviation accidents are caused by human error (Helmreich, 2000), a view that is still current by judging recent reports classifying accidents by causes, including pilot error (Baker et al., 2008).

A similar category of error classification is maintenance error. Marx & Graeber (1994) present definition of maintenance error: "Maintenance error can be defined as an unexpected aircraft discrepancy (physical degradation or failure) attributable to the actions of an [aircraft maintenance technician]" (p. 88).

Much research has been undertaken into aircraft maintenance, which is understandable since maintenance is a particularly complex and complicated human endeavour regarding human, technical, and organizational factors (e.g., Gramopadhyea & Drury, 2000; Reason & Hobbs, 2003). Large numbers of diverse aircraft parts need to be checked, removed, dismantled, repaired, replaced, and/or reassembled, often under difficult working conditions. Various concurrent job schedules and processes may involve blended groups of maintenance personnel and may take a long time across shifts, entailing various managerial and organizational issues. Some of these factors were addressed and identified as problematic in the NTSB (2003) investigation of Alaska 261 as well (not surfacing in the accident report summary above), showing the complexity of trade-offs in aircraft maintenance work as it is actually performed.

Several perspectives have been taken in order to understand and improve aircraft maintenance work, classification of errors and failures being one of them, which has occupied researchers since the dawn of safety and human factors research. Several researchers have cited accident statistics of the 1980s and 90s, which indicate that maintenance error is a factor (typically among others) contributing to 12-18 percent of commercial aviation accidents (Gramopadhyea & Drury, 2000; Marx & Graeber, 1994; Rankin et al., 2000). Classifications of maintenance error often differentiate between
2.4. AVIATION SAFETY

Various kinds of maintenance error. The British Civil Aviation Authority published a list of most frequent maintenance problems (cited in Marx & Graeber, 1994), listing physically observable technical problems, such as incorrect installation of a component, installation of the wrong component, electrical wiring, and inadequate lubrication. A Boeing/Air Transport Association (ATA) study found various specific factors more remote from the maintenance technician, such as manufacturer/vendor maintenance/inspection error, airline policy, and poor design to contribute to maintenance-related accidents to a large extent (cited in Rankin et al., 2000). Others have classified maintenance errors in classes of cognitive activity. Hobbs & Williamson (2002) for example classified maintenance activities and errors in the categories of skills-, rules-, and knowledge-based, measuring percentages of working activity and errors in each category.

Error classification and counting may give some indication of which categories of errors may need to be addressed. However, theoretical error classification schemes without clear links to remedial action are of little help to the maintenance communities (Marx & Graeber, 1994) and categorizing human errors does not provide an understanding as to why and how failures occur, but rather disembodies, decontextualizes, and relabels data instead, possibly misleading efforts of safety improvement (Dekker, 2003, 2004). The traditional human error perspective treats a socio-technical system as a technical system where parts fail to perform as designed (Dekker, 2003, 2004; Hollnagel, 2004). This is in essence a normative/prescriptive approach, or at best a descriptive approach comparative to a norm (Rasmussen, 1997), approaches that Amalberti (2001) has argued to be able to improve safety up to, but not beyond, the one accident per $10^{-5}$ events level (calling higher levels "ultra-safety"). This corresponds to a sequential or epidemiological accident model perspective. Moreover, failure/error counting schemes give indications of system performance only at isolated points in time, under-appreciating long-term effects such as gradual changes in operational or managerial practice and cultural aspects that may play a role.

As Ernst Mach already argued in 1905, "knowledge and error flow from the same source, only successes can tell one from the other" (as cited in Rasmussen, 1986, p. 151). At least since the 1980s scholars have debated whether human error really is a useful concept and can be seen as a cause of accidents. These authors state that human error is the result of trouble somewhere deeper in the system, a consequence or symptom rather than a cause (Dekker, 2002, 2003, 2004; Hollnagel, 1983; Woods et al., 1994). Investigations should therefore dig deeper and continue their analysis, rather than stop, at the conclusion that somewhere in the system, a person did something wrong. The often reported "loss of situation awareness" has similar problems as an explanatory concept. Situation awareness theory (Endsley, 1995) views the assessment of the situation as a passive information digestion process, in contrast to well-established models of cognition such as
Neisser’s (1976) perceptual cycle where the current understanding directs perception and information gathering, not only perception shapes current understanding. The theory does not explain what fills the “cognitive vacuum” that is supposedly created by the “loss of situation awareness”, but rather demands an explanation itself (Dekker, 2002) (see the description of Norwegian 541 in the Introduction). As Woods & Hollnagel (2006) note, “there is no such thing as a cognitive vacuum – the work system is adapted to some purposes, variations, constraints, even if they are not the ones that outsiders see or desire” (pp. 55-56).

Trying to understand and model the behavior of joint cognitive systems in a descriptive manner is one of the goals of cognitive systems engineering. This is reflected in the CSE views on accident investigation and human error, and decision making, see Section 2.3.4. For now, we may remark that this way of thinking emerged during the 1970s and 1980s, not only in the human factors discipline, but also in management.

Rationality is a judgment made by one or more persons of one or more others (subjects). It involves the relationship between the researcher’s and the subject’s models of the reality faced by the subject(s). For this reason, the attribution of irrationality by one to another is often a consequence of mismatching models. Therefore, [...] researchers are better off attributing rationality to others (even if this implies their own irrationality) and trying to uncover the model(s) of the others because it leads to better understanding of them and an increased ability to serve them. (Ackoff, 1983b, p.719)

Ackoff’s statement reflects the cognitive systems engineering viewpoint, that trying to understand human behavior is the focus of research, rather than developing normative models of rational man that people should emulate and developing systems that try to correct people’s behavior’s deviance from the rational ideal. Rationality of people’s reasoning is not an assumption in cognitive systems engineering. Moreover, rationality in the normative decision theory sense is seen as intellectually and practically unnatural (Hollnagel, 2005). Instead, cognitive systems engineering emphasizes that people always try to make trade-offs between thoroughness and efficiency (Hollnagel, 2004) when trying to accomplish their goals and meet changing demands.

Moreover, knowing the outcome of events affects judgment under uncertainty. People ascribe increased likelihood to the actual outcome of an uncertain course of future events when knowing that outcome, and overestimate what they would have known and what others did know without that outcome knowledge (Fischhoff, 1975). Hollnagel & Amalberti (2001) present an alternative view on human error, stepping away from hindsight bias, stating that two aspects of actions are essential and sufficient in the evaluation of people’s “errors”. First, once an action is taken, the only eval-
uation that the operator(s) can make is whether the outcome of the action was as intended. Unintended outcomes may thus be detected or go unnoticed. Second, once an unintended outcome is detected, the operator(s) may have the opportunity and ability to recover from the unintended situation or not. Note that unnoticed unintended outcomes, as well as noticed but unrecovered unintended outcomes, would be classified as human errors in the classical view on human error. The unintended outcome detection and recovery perspective thus evaluates human actions in a non-judgmental fashion.

Models of drift and migration

More organizationally oriented work may be illustrated by continuing the description of maintenance work, in connection to the Alaska 261 case described in the Introduction. A more descriptive, dynamic, and organizational account of aircraft maintenance is provided by McDonald et al. (2000), describing an abstract self-regulatory model of safety management systems (SMSs), focusing on human and organizational aspects. In this model, the system is described in terms of policy, standards, work organization, and normal operations, where policy affects standards, which affect work organization, which affects normal operations. Safety management provides monitoring of operations, on which feedback is based, leading to change in all parts of the aforementioned system. The study gives interesting insight and methodological support on the investigation of organizational and safety culture, reporting a rather homogeneous professional culture within four aircraft maintenance organizations, and differences in SMS use and safety climate. Guidance on philosophies or concrete methods to use to achieve the various functions it implies remains a challenge. Theories and assessment methods of safety climate/culture are numerous, but the concept is not yet well-understood, as indicated by the diversity in and lack of agreement on theories, concepts and methods (see, e.g., Flin et al., 2000; Guldenmund, 2000, for extensive reviews). Moreover, Alaska 261 underlines the fragility of this self-adjusting safety management process, craving a more detailed understanding of the processes and functions involved.

Two metaphors that aim to integrate the dynamics and multiplicity of complex socio-technical systems are migration toward boundaries and drift toward failure. Rasmussen (1997) proposes a model about how variable system performance is likely to migrate towards boundaries under the influence of various pressures. In this model, boundaries are put up by economic and workload constraints, and perceived and functionally acceptable (safe) performance define the space of possible action, similar to the "field of safe travel" suggested by Gibson & Crooks (1938). The system exhibits variable Brownian-motion-like performance, while being pressured towards the various boundaries by pressures of management for efficiency,
of safety-improvement initiatives, and a tendency of minimizing effort. A field of possible action is thus navigated, in a fashion which bears similarity to Gibson’s (1986) ecological theory of perception. A hierarchical modeling philosophy and several analysis techniques describing organizations in various levels are associated to this model, enforcing constraints down the hierarchy and communicating information about the enforcement of constraints upward (Leveson, 2004; Rasmussen, 1997; Svedung & Rasmussen, 2002).

Amalberti has developed Rasmussen’s model further as a model of ecological safety, or the model of migration and transgression of practices (Amalberti, 2001; Amalberti et al., 2006; Polet et al., 2003). An initial safe space of action is defined during system design stage is defined by boundaries of rules, protocols, in-depth defences, and barriers, and Rasmussen’s (1997) boundaries of economic and workload constraints. From that the system is taken in operational use, under pressures of improved performance and individual gains, performance migrates to also include borderline tolerated conditions of use, “violations” outside of the “legal” designed specifications of use, but within the tolerance and sometimes even requirements of staff and management. In a later phase, performance may migrate further (tolerated only under extreme pressure or conditions) and transgress the boundary where violations take the form of close calls, and eventually transgress the universally unaccepted boundary of safe performance, leading to incidents and accidents. This model stipulates that migration of the system is almost unavoidable, including violations of boundaries set in the design phase, in order for the system to survive under increasing pressures of economic and individual gain. The migration model thus stipulates that performance varies in a field of safe behaviour between boundaries, where violations (transgressions of boundaries) after some time become normal and cannot be eliminated but should be managed instead. However, as Amalberti et al. (2006) note, violations are as yet incompletely understood by the research community. To gain understanding, and to improve safety, a descriptive approach of normal and actual performance is necessary, rather than merely a more strict enforcement of safety constraints.

The concept of drift toward failure has been suggested as a similar metaphor for describing a system that slowly develops more risk-filled behaviour and eventually breaks down, while all looked well during the process at the decisive operational or managerial levels. It provides a metaphorical description of this phenomenon, which has many similarities to the migration toward boundaries model discussed above. Dekker (2004) calls drift toward failure “the greatest residual risk in today’s sociotechnical systems”, and describes its features as being generated by normal processes of reconciling conflicting goals under uncertainty, incremental decisions over long time spans, and normalization of signals of danger aligning operational observations with organizational goals.
Both space shuttle accidents have been called examples of drift. Starbuck & Milliken (1988, p. 319) describe organisations such as NASA before Challenger as to "interpret past successes as evidencing their competence and the adequacy of their procedures, and so they try to lock their behaviours into existing patterns", to "evolve gradually and incrementally into unexpected states", and to "fine-tun[e] the odds". Vaughan (1996) describes NASA organisational behaviour before Challenger as "normalization of deviance".

Woods (2005), in describing lessons learned from the Columbia accident, charts the drift toward failure in this particular accident by identifying points in time where the evaluation of risks related to debris strikes shifted or could have shifted. Along with three shifts, he identifies five general patterns in the period up to the Columbia accident (Woods, 2005), which are also highly applicable to Alaska 261, illustrating the processes involved in the drift toward failure: (1) defences eroding in the face of production pressures, acute efficiency goals gradually taking precedence over chronic safety goals, (2) gaining confidence from past success instead of investing in anticipating the changing potential for failure, (3) fragmented distributed problem solving clouding the "big picture", no person had a complete and coherent view of the problem and cross-checks were missing, (4) failure to revise assessments as new evidence accumulates, not capturing and displaying indicators of safety margins, and (5) breakdown at the boundaries of organizational units, lacking effective overlap.

Woods’ (2005) analysis thus shows definite parallels to the model described by Rasmussen and Amalberti. Both models identify various pressures that drive a gradual migration/drift through a safety space (be it with clearly identifiable shifts or not). However, the migration/drift metaphors may be seen as the starting point for future research, as they do not yet solve organizations’ difficulty in setting specific safety goals as to achieve "non-events" and measuring where current activity is "positioned" relative to these goals, in terms of concrete methods of how to prevent, detect and recover from migration/drift.

**Systems-theoretic accident model and processes**

Leveson (2004) developed the systems-theoretic accident model and processes (STAMP), mostly based on work on proactive risk management (Rasmussen & Svedung, 2000), systems theory (e.g., Checkland, 1981), and cybernetics (e.g., Ashby, 1956). Some basic theory of the model is presented here because of its relevance to the subject of constraints and control, and because it is seen as one of the few systemic accident models developed to date. Systems theory is the major theoretical component of the STAMP model, which shows extensive appreciation for the concept of constraint
Leveson treats systems as hierarchically organized components that interact. She sees safety as an emergent property that arises when the system components interact, and views accidents as a control problem: "accidents occur when component failures, external disturbances, and/or dysfunctional interactions among components are not adequately handled by the control system [], when they] violate the safety constraints." (Leveson, 2004, pp. 249–250). This may occur at many levels in the organization, as well as at various points in time: accidents may involve inadequate controls at operator run-time level, at the work process management level, or at the design level and stage. Accidents result from a lack of appropriate constraints (control laws) on component interactions, and a system design must enforce the appropriate constraints to ensure safety and continue to do so as changes occur. The basic concepts in STAMP are constraints, control loops and process models, and levels of control. High levels in this organization correspond to the blunt end, lower levels to the sharp end. Interactions between components form adaptive control loops at various control levels. In order to provide adaptive control (and thus safety) communication has to occur. Information to enforce constraints is communicated downward, feedback to ensure that constraints have effectively been enforced is communicated upward. The appropriate constraints may include technical design and process constraints, management constraints, manufacturing constraints, and operational constraints (Leveson, 2004).

For example, in the Challenger accident (Vaughan, 1996), a systems theory view identifies several occurrences of safety constraints that were violated (Leveson, 2004): At a technical level, safety constraints were not enforced when the O-rings did not adequately control propellant gas release from the space between field joints of the solid rocket boosters. This information was not adequately communicated up the hierarchy with the result that higher (managerial) levels did not have adequate understanding about the violated constraints. Thus safety constraints were violated at managerial levels as well.

**Functional resonance accident model**

The functional resonance accident model (FRAM; Hollnagel, 2004) is a systemic accident model, in accordance with the principles of resilience engineering (see next Section). FRAM is based on four principles (Hollnagel, 2004; Hollnagel et al., 2008).

First, the principle of equivalence of successes and failures states that both successes and failures result from the adaptations that organizations, groups and individuals perform in order to cope with complexity. Success depends on their ability to anticipate, recognise, and manage risk. Failure is due to the absence of that ability (temporarily or permanently), rather than to the inability of a system component (human or technical) to function normally.
Second, the principle of approximate adjustments states that complex socio-technical systems are by necessity underspecified and only partly predictable. Procedures and tools are adapted to the situation, to meet multiple, possibly conflicting goals, and hence, variability of performance is both normal and necessary. The variability of one function is seldom large enough to result in an accident.

However, the third principle, of emergence, states that the variability of multiple functions may combine in unexpected ways, leading to disproportionately large consequences (non-linear effects, see also (Guastello, 2003)). Normal performance and failure are therefore emergent phenomena that cannot be explained by solely looking at the performance of system components.

Since functions can be coupled and interact, the variability of one function may be influenced by the variability of the environment of that function (other functions). A function may damp the variability of other functions, or a number of functions may be reinforced. Thus, the fourth principle, of functional resonance, states that the variability of a number of functions may reinforce one another and thereby resonate, causing the variability of some functions to exceed normal limits, the consequence of which may be an accident. FRAM as a model emphasizes the dynamics and non-linearity of this functional resonance, but also its non-randomness. FRAM as a method (see Section 2.1.5) therefore aims to support the analysis and prediction of functional resonance in order to understand and avoid accidents as well as understand performance in general and why things go right.

### 2.4.3 Approaches to safety

Various distinctions between perspectives of how to achieve safety have been proposed. Morel et al. (2008) propose a distinction between "safety by constraint" and "safety by management". Another well-known perspective is "safety by design".

**Safety by constraint** entails the attitude that safety measures should restrict the behavior of technical systems and mostly people by imposing barriers on their behavior. Typical safety recommendations include additional procedures for people to follow, or technical barriers that prevent human errors. An example is the design of the current fly-by-wire Airbus aircraft that limit the angle of attack and modify engine thrust so that it is virtually impossible to stall the aircraft, thereby constraining the aircraft performance within the safety boundaries where sufficient lift is generated. This example however also illustrates that safety by constraint may be accomplished by design. This perspective goes further than merely constraining technical systems however, in the sense that this perspective aims at constraining human action possibilities to such an extent that human error, a term that fits this perspective, is made impossible. Strict(er) adherence to
procedures and regulations, for example, is a typical result of safety by constraint view.

Hale et al. (2007) describe the principle of engineering safety into systems, that is, accomplishing safety by design. An example is given in aviation for free flight where safety is attempted to be built into the free flight airspace system through the design of systems and procedures (Hoekstra et al., 2002) or in test flying, where test pilots provide input to the early design process of new aircraft and test the design at later stages (Singer, 2001). In safety by design, all safety considerations are attempted to be covered in the design phase of a complex socio-technical system. Safety by design is therefore broader in the sense that it considers not only constraints in the limiting sense, but also the deliberate design of opportunities for action. However, reality teaches that it is virtually impossible to foresee and evaluate all future threats to complex socio-technical system performance, the Alaska Airlines and Norwegian Air Shuttle cases being examples of this. Complementary ways of achieving safety are necessary.

Safety by management acknowledges the shortcomings of the safety by constraint and design perspectives, which can provide safety only to a certain level. Designs are always merely hypothesis of how future systems will behave and will be used (Woods, 1998; Woods & Dekker, 2000), and designers cannot imagine all future situations where their designs may be used. An auto-pilot mode for example may provide safety benefits in many situations, for example, but may also have negative consequences in other situations that may not be foreseen. This is also why the safety by constraint perspective does not guarantee safety, at the specification and implementation of these constraints certain future violations of these constraints may be overlooked that may be necessary for safety to be maintained. Studies of procedures for example have shown that violating procedures sometimes is necessary to avoid an accident, and that work-to-rule strikes, for example in ATC, may result in extremely inefficient and even unsafe performance. Because of these uncertainties of future use, variability in joint human-machine system performance will occur that is not only natural because of the inherent variability that humans and technical systems display in their performance, but even necessary to overcome constraints and designs that are maladjusted to the circumstances of specific unanticipated situations. Safety should thus continuously be created, monitored, and managed, to stay alert about shortcomings that have not yet been recognized or identified.

Resilience engineering

The resilience engineering approach (Hollnagel et al., 2006) suggests that a failure is a temporal or permanent absence of improvisation and adaptation capabilities to cope with undesired events rather than a breakdown or malfunctioning of an organisation or technological component. Resilience
2.4. AVIATION SAFETY

Engineering is a new approach to improve adaptation, anticipation, and improvisation in complex socio-technical systems, broader than the search for safety. Resilience is in this paper defined as the ability to anticipate, prevent, detect, and recover from harmful events. Three interconnected essential aspects of resilience are therefore foresight, coping, and recovery.

Woods (2006) describes the essence of resilience as the ability of a system to recognize when variability in its performance is unanticipated and falls beyond the usual competence and adaptations:

Resilience ... concerns the ability to recognize and adapt to handle unanticipated perturbations that call into question the model of competence and demand a shift of processes, strategies and coordination ... The focus is on assessing the organization’s adaptive capacity relative to challenges to that capacity ... Resilience Engineering must monitor organizational decision-making to assess the risk that the organization is operating nearer to safety boundaries than it realizes. (Woods, 2006, p. 22)

Thus, knowing where system performance currently is in relation to boundaries in Woods’ drift toward failure or Rasmussen and Amalberti’s safety space is not a trivial assignment, and one of the objectives of resilience engineering. Industry representatives as well as scholars within this field of Resilience Engineering have voiced the desire to assess indicators for resilience. Whereas measurements and indicators of resilience based on system outputs can be expected to provide limited insight because they fail to capture the dynamics essential to resilient systems, process measures and model-based assessments may be powerful enough to assess the resilience of dynamic processes and organizations in order to make predictions of resilient performance or the lack thereof (Mendonca, 2008; Hollnagel et al., 2006).

First attempts at applying the new concepts of resilience engineering have been reported in several domains, among which emergency management/response (Mendonca, 2008; Petersen & Johansson, 2008; Woltjer et al., 2006) and air traffic management (Sträter, 2008). Theory and method developments in resilience engineering can however be described as being in their early stages.

2.4.4 Analysis techniques

Human reliability analysis in its first generation focuses primarily on probabilistic risk assessment of human failure. This approach has generated some criticism, for example for low not prioritizing qualitative aspects, unsatisfactory theory behind the methods, low validation, insufficient reliable data on human errors, and the need to address performance shaping factors (Dougherty Jr, 1990; Hollnagel, 1993). Subsequent modeling
methods have attempted to overcome these problems. A technique for human error analysis (ATHEANA; Cooper et al., 1996) was developed in the 1990s for the US Nuclear Regulatory Commission, as one of the first second-generation HRA techniques. The cognitive reliability and error analysis method (CREAM; Hollnagel, 1998a) was also developed as a response to first-generation HRA techniques. Similar developments have been observed in the accident and risk analysis techniques at large. Several overviews of accident analysis techniques have been published (e.g., Harms-Ringdahl, 1996, 2001; Sklet, 2004), as well as of risk analysis techniques (e.g., Cacciabue, 2000; Harms-Ringdahl, 1996, 2001; Kirwan, 1998).

As Hollnagel (2004, 2008a) argues, accident and risk analysis techniques are always associated with one or more (explicit or implicit) underlying accident model categories (see Section 2.4.2), which make the methods suitable for socio-technical systems of varying tractability and tightness in coupling. There are techniques that are not explicitly based on establishing a model of the system to be analyzed, such as HAZOP, which is an approach to revealing hazards. Suokas (1993) points out that this may cover events requiring one or two causes, but to get at events requiring three or more different causes, model-based assessment is necessary. Event trees and fault trees for example fall into the category of sequential accident model techniques. Barrier analysis methods and the Swiss Cheese model-based method TRIPOD are methods in the epidemiological model category. The methods associated to STAMP (discussed in Section 2.4.2) and FRAM (Section 2.1.5) are methods that have been classified in the systemic category by their originators. Three more methods deserve attention here: task analysis techniques, the sequentially timed events plotting method (STEP), and functional hazard analysis (FHA).

Task analysis techniques

Task analysis is one of the central methods in cognitive systems engineering, human factors, ergonomics, and human-computer interaction. It is not a method directly used in system safety, but is often part of the process of introducing new technology, and therefore closely related to safety efforts. Numerous techniques are available that have been published widely, including several antologies on the subject (Diaper & Stanton, 2004; Kirwan & Ainsworth, 1992; Schraagen et al., 2000). Within the domains of application discussed here, task analysis has been particularly effective for well-defined tasks, especially in air traffic control. ATC task analyses will briefly be reviewed here.

Numerous task analyses of air traffic control have been produced and published. This section highlights some of the better-known models, or models that are illustrative of a type of task analysis. As one of the early cognitively oriented descriptions of air traffic control work, Mattsson (1979) described cognitive functions involved in air traffic control. The
FAA (Rodgers & Drechsler, 1993) has developed a hierarchical task description of the air traffic control task, using hierarchical task analysis (Annett, 2003). Seamster et al. (1993) describe 13 tasks which have been modeled in a COGNET architecture with a (blackboard structure) mental model for sector management, conditions, and prerequisite information. The EATCHIP programme (Barbarino et al., 1996) developed a model consisting of 23 task clusters, of which seven core tasks, seven direct support tasks, and nine indirect support tasks. Each task cluster in turn contains between two to ten subtasks. The EATMP2 programme (Kallus et al., 1999) identified one control process (switching attention), five task processes and four subprocesses. The programme developed a graphical process representation (flow chart) with activities, optional actions (conditions), decisions, time buffers, and processes. Blom et al. (2001) developed a Petri-Net (a mathematical modeling language of bipartite graphs with non-deterministic execution) of 19 controller subtasks and the control modes tactical and opportunistic (from COCOM, described above). Controller subtasks were obtained by combining ten generic cognitive tasks with three en-route context specific tasks. 19 possible combinations of generic cognitive and en-route specific tasks were identified. Niessen & Eyferth (2001) developed a model of how the controller maintains a mental picture of the traffic situation. It has five modules each containing several procedures, and three information processing cycles that interconnect these modules. The SHAPE project (Voller & Low, 2004) has defined 21 cognitive functions and 18 ATM functions, and subsequently links between automation functions and cognitive processes, between cognitive processes and cognitive functions, between cognitive functions and ATM functions, and between controller functions and skill requirements.

Most of the models discussed sofar have the purpose of aiding in training and selection. Models that allow the explicit assessment of man-machine systems, i.e., controller(s)-pilot(s)-aircraft-automation systems, are rare. A notable exception is the MIDAS modeling framework (Corker, 2000), and to some extent the modeling work around the ERATO tool (Leroux, 2000).

**Sequentially timed events plotting**

The sequentially timed events plotting method (STEP; Hendrick & Benner, 1987) is a multi-linear event sequence model and method. In STEP, an accident is a special class of process where a perturbation transforms a dynamically stable activity into unintended interacting changes of states with a harmful outcome. In this multi-linear approach, an accident is viewed as several sequences of events and the system is decomposed by its structure.

STEP provides a comprehensive framework for accident investigation from the description of the accident process, through the identification of safety problems, to the development of safety recommendations. The first
key concept in STEP is the multi-linear event sequence, aimed at overcoming the limitations of the single linear description of events. This is implemented in a worksheet with a procedure to construct a flowchart to store and illustrate the accident process. The STEP worksheet is a simple matrix. The rows are labelled with the names of the actors on the left side. The columns are labelled with marks across a time line.

Secondly, the description of the accident is performed by universal events building blocks. An event is defined as one actor performing one action. To ensure that there is a clear description the events are broken down until it is possible to visualize the process and be able to understand its proper control. In addition, it is necessary to compare the actual accident events with what was expected to happen. A third concept is that the events flow logically in a process. This concept is achieved by linking arrows to show proceed/follow and logical relations between events. The result of the third concept is a cascading flow of events representing the accident process from the beginning of the first unplanned change event to the last connected harmful event on the STEP worksheet. The organization of the events is developed and visualized as a "mental motion picture". The completeness of the sequence is validated with three tests.

The row test verifies that there is a complete picture of each actor’s actions through the accident. The column test verifies that the events in the individual actor rows are placed correctly in relation to other actors’ actions. The necessary and sufficient test verifies that the early action was indeed sufficient to produce the later event, otherwise more actions are necessary. The STEP worksheet is used to obtain a link between the recommended actions and the accident. The events represented in STEP are related to normal work and help to predict future risks. The safety problems are identified by analyzing the worksheet to find events sets that constitute the safety problem. The identified safety problems are marked as triangles in the worksheet. These problems are evaluated in terms of severity. Then, they are assessed as candidates for recommendations. A STEP change analysis procedure is proposed to evaluate recommendations. Five activities constitute this procedure. The identification of countermeasures to safety problems, the ranking of the safety effects, assessment of the trade-off involved the selection of the best recommendations and a quality check.

Functional hazard analysis

Functional hazard analysis (FHA; see EUROCONTROL, n.d., for the manual for FHA discussed in this section) as defined by EUROCONTROL is part of their safety assessment methodology. FHA is the first part, which is discussed here, in the EUROCONTROL methodology followed by preliminary system safety assessment and system safety assessment. FHA identifies potential failure modes and hazards, and is ideally performed at the overall air navigation service or system level, before functions are allocated
to equipment, people, or procedures. Practical constraints, however, make that it is generally applied at sub-system level, and when some function allocation has been done due to earlier/existing system design. FHA is performed iteratively as function allocation and design progresses. FHA consists of five parts: (1) initiation, (2) safety planning, (3) safety objectives specification, (4) FHA evaluation, and (5) FHA completion. Steps 1, 3, and 4 are analysis steps, and steps 2 and 5 are of a project management nature.

The initiation develops a description of the system (change) to be assessed in terms of its functions, assumptions, earlier analyses (of hazards, incidents, lessons learned, …), operational environment and system interfaces to this environment, regulatory framework, and applicable standards. The safety objectives specification identifies all potential hazards associated with the system, the effects of hazards on operations, the severity of these effects, the safety objectives (the maximum frequency of hazards), and the organization safety target (the overall foreseen (future) risk associated with introducing the assessed system (change)). The evaluation entails verification (demonstrating that the safety objectives meet the organization safety target), validation (review and analysis of the outputs of steps 1 and 3 to ensure their traceability, completeness, correctness, credibility, and sensitivity), and assurance (a project management and internal review process).

2.4.5 Summary: Aviation safety

Principles of aviation safety and safety in general have been discussed in this section, including the concepts of safety, incidents and accidents, risk, accident models, human error, hindsight, drift and migration. The approaches of safety by constraint, safety by design, safety by management, as well as resilience engineering, have been outlined, for their relevance to constraint management and the specific cases discussed in the aviation safety studies. Two systemic models have been discussed, STAMP and FRAM, for their particular focus on constraint management (STAMP) and functional modeling (FRAM). Specific analysis techniques, namely task analysis techniques, STEP, and FHA have been described for their similarities or comparison, in the appended papers, to the modeling efforts with FRAM.
Chapter 3

Methodology

The methodology employed throughout the studies reported here is described in this chapter. The chapter starts by discussing general characteristics of research settings in cognitive systems engineering. The subsequent four sections describe four environments of discovery generally, and the specific environments and methods that were used in the six studies described in the appended papers to the thesis. The final section summarizes the methodology in an overview table.

3.1 Research settings

In cognitive systems engineering, methods that are used to analyze and describe the behavior of joint cognitive systems focus on the characteristics of observable behavior, or performance (Hollnagel & Woods, 2005). Functions and constraints take a central role here in relation to performance. Various research settings may be used to enable the analyst to uncover the functions performed by joint cognitive systems and the relevant associated constraints.

Table-top analysis (Flach, 2000), is the analysis of the published literature that describes the nature of work in a particular domain or work situation. These sources of information help identify well-established constraints (although these should be questioned), guide and ground observations in "naturalistic" settings, facilitate communication with domain experts, and can serve as a basis for further analyses (Flach, 2000).

Natural history (in situ) methods (Woods & Hollnagel, 2006) or naturalistic observation (Flach, 2000) may be used to enhance the view of the analyst on the activity to be described, studying natural settings or cases. Unclarities or contradictions in the "homework" of the table-top analysis, and aspects of the activity that experts or practitioners find difficult to describe, require observation for accurate understanding (Flach, 2000). However, observation can also be used as the fundamental process by which
participants’ actions are understood (see, e.g., Silverman, 1993, for a discussion). Natural task environments are mainly used for discovery of patterns, where the conditions of observation are not as much shaped by the analyst/observer as they are chosen from the natural environments that are available for their relevance to answering the research question at hand (Woods & Hollnagel, 2006). Capturing the “naturally” occurring constraints and people’s constraint management in various situations may thus benefit from observation in actual working situations.

Simulated task environments aim to simulate aspects of actual systems to people that engage in tasks. The kinds of aspects and systems range from one or few aspect(s) of many systems to many aspects of one system, a gradual scale that Gray (2002) calls correspondence. Spartan laboratory experiments aim for simulating the former, allowing for the control or shaping of constraints in the laboratory environment, so that parsing the phenomenon of interest becomes paramount (Flach, 2000; Woods & Hollnagel, 2006). Highly “realistic” simulations and exercises aim for simulating the latter, including many of the constraints in natural environments enabling people’s expertise in these environments to apply to the simulation or exercise. Simulated task environments moreover enable measurement of performance at many levels, consequence-free evaluation of naturally high-risk activities, and higher control of constraints in the environment than natural environments, although lower than in the laboratory (Flach, 2000).

The summarizing concept for what others have called "realism", "fidelity", or "ecological validity" has been described by Brunswik (1955) as representative design. The representativeness of the environment (Brunswik, 1955), the representativeness of the participants (as is commonly emphasized in the sampling of participants in laboratory studies), and the representativeness of their tasks, is of importance. When interested in generalization of findings, researchers need to address which characteristics of environments, participants, and tasks their research settings represent, in order to determine to which types of environments, participants, and tasks their research results may generalize. Note also that there may be (sometimes subtle) interactions between the representativeness of environments, participants, and tasks (or functions). An environment (for example, a firefighting simulation) may be very different from the environment that participants (professional fire-fighters) usually are subjected to, so that their tasks in that environment become very different from the tasks in which they have experience. Police officers, for example, may find that the communication and cooperation tasks they perform daily, may be very useful in a in a fire-fighting simulation environment, so that they may use their experience in the research/training setting to solve a new type of task.

The case study may be seen as the central approach in the studies presented here. Case studies (Hammersley & Gomm, 2000): (1) typically investigate a relatively small number of cases (here, one case in each study),
(2) gather and analyze information on relatively many features of each case (compared to very few conditions in an experiment), (3) usually focus on naturally occurring cases, or cases created by actions of researchers without the purpose of controlling conditions and measuring effects, (4) may prioritize qualitative over quantitative analyses, and (5) may or may not aim to understand the case itself, and may or may not aim to infer theory or empirical generalization. As can be concluded from these characteristics, researchers do not agree about the philosophical underpinnings of case studies. Travers (2001) identifies four traditions: positivism (striving for objective world descriptions, and representativeness), interpretivism (striving for understanding how members of society see their own actions), realism (striving for the discovery of laws or mechanisms explaining behavior, and deficiencies in people’s accounts of these), and poststructuralism (challenging the possibility for obtaining valid knowledge about the world, in contrast to the earlier three).

This thesis employs a number of methods of qualitative enquiry in Papers I through VI. Observation in natural and simulated environments, interviewing, textual analysis, and discourse analysis are the main qualitative methods (Silverman, 1993; Travers, 2001), applied to the various cases. The state space approach of Paper II may be classified as both a qualitative and quantitative method to describe performance. Indirectly, quantitative and qualitative methods were used in connection to Paper VI, where the ERASMUS experimentation as part of the concept validation process was used as an input to the modeling effort presented in this study. The remainder of this Chapter discusses the specific environments of discovery and the methods employed.

3.2 Functional exercise

Various kinds of command and control exercises are used as a training, evaluation, or demonstration tool in emergency and crisis management and military command and control (see, e.g., Trnka & Jenvald, 2006, for a review). The study reported in Paper I describes a functional exercise. A functional exercise (Perry, 2004) is a specific type of exercise that presents considerable complexity to the participants in testing planning and training, compared to a table-top exercise that tests action intentions based on an emergency narrative, but does not implement all functions in the actual context like a full-scale exercise does. Functional exercises select one or a few functions as a focus, may involve one or more emergency management agencies, are usually conducted in real-time, by operational personnel with appropriate equipment, in the field and under realistic conditions (Perry, 2004). This includes a detailed preparation and implementation of a simulated course of events, as inputs and reactions to the participants’ actions. Realism and validity of functional exercises are important issues in their design. Thus, functional exercises aim to represent the few functions cho-
3.2. FUNCTIONAL EXERCISE

sen to be included in the design, the participants that would normally perform these functions, and an environment similar to the environments that participants perform these functions in.

Drabek & Haas (1967) attempt to answer the question of what makes a simulation realistic and describe five aspects of simulations tied to realism, which determine to which situations the simulation may be generalized: (1) the group of participants history, a longer working experience together having the effect that roles and patterns of interaction have developed over time; (2) the type of task, activity, or demand, and (3) the ecological setting, which the group is subjected to in the simulation and their similarity or representativeness with respect to the history of the group, and (4) the social systems in the simulation environment, which play a role in the participating group’s functioning; and (5) whether participants are aware of their participation in a simulation or experiment or not.

Feinstein & Cannon (2002) discuss validity as the relationship between simulation development and educational processes. According to their model, internal validity relates representational and educational validity in that participants cannot insightfully engage in a simulation if it does not behave like a phenomenon from the real world that they can recognize or understand. External validity relates to (1) how well the simulation represents real world phenomena (external representational validity) and (2) that the simulation has the desired learning effect (external educational validity).

These concepts will hereafter be applied to the specific exercise that was studied and is described in Paper I.

3.2.1 Electricity network resource management exercise

The exercise was performed in real-time and simulated the effects of a devastating storm. The scenario combined hypothetical events with events similar to the January 2005 storm’s effects on power and communication infrastructures as described in Section 1.2.1. Parts of the exercise had been planned long before the January 2005 storm. The exercise was initially meant to train radio correspondents’ crisis communication skills. The repercussions of the January storm increased interest in the exercise. Because the scenario was only partially based on an actual emergency, the participants were likely to experience it as highly realistic without being overly constrained by their experience with the actual emergency.

The scenario was played in real-time at a training centre and in parallel at various remote sites on 19 April 2005. The exercise took approximately ten hours. It started with the situation that on 14 April, a big storm forcefully accelerated through the Benelux, Denmark and southern Norway through the southwest of Sweden towards Stockholm on the 15th. The violent advance of the storm resulted in power outages and downed telephone lines and masts, and associated severe infrastructure problems. The
utilities became interested in participating to be able to re-enact the consequences of the storm and learn to deal with them more efficiently.

Electricity and telecommunications utilities, as well as many other infrastructure organisations and companies, sent a total of approximately 60 representatives to the exercise. Participants were decision-makers from most of the companies involved in the real storm who were responsible for reacting to the emergency and for coordinating activities to return service to normal. Paper I reports on the observation of a team of seven persons responsible for coordinating cooperation between electric utilities during emergency operations.

The team that the first two authors (Woltjer and Lindgren) of Paper I observed was assembled by the electric companies and agencies and was the same team that would have responded to a real-life crisis. One of the team members was the designated chairperson. All other team members had flexible roles. The participants had extensive domain knowledge in their area of expertise. They normally worked in different locations and communicated by telecom links. During the exercise, they were collocated and therefore used a different set of artefacts than they were used to. In order to discuss the differences in actual and exercise settings in detail, the notions of realism and validity need to be addressed.

The aspects of realism translate to the exercise described in Paper I as follows: (1) The participants had performed in exercises before, and most of them had a history of cooperation together. The purpose of the exercise was also to strengthen patterns of interaction and gain working experience in cooperation. (2) The types of tasks were relatively clearly defined, and for those with experience these matched their previous experience and cooperation history. (3) The ecological setting, or context, was however different, in that the training setting did not match the history of the group in actual emergency situations. (4) The social systems in the simulation environment were new to the observed team, because the team that was their equivalent in the telecommunications domain was assembled for the first time, and moreover was within walking distance so that meetings could easily be arranged. (5) The participants were well aware of their participation in a simulation.

Regarding validity of the exercise setting, internal validity was not problematic as team members could quickly orient themselves in their task and environment and recognized and understood the exercise to a satisfactory extent. The representativeness of the environment and of the task, expressed in terms of artifacts used for team functions is used to discuss external representational validity and its relation to external educational validity. The study thereby analyzes the case of a specific electricity network resource management team in a specific functional exercise.

The methods used included direct observations of the team in its workrooms by the two observers, video recording of the team’s workplace, recording traces of artefact use (telephone, white-board, internet), and a
semi-structured retrospective interview with a subset of the observed participants. The data thus obtained was used to create a functional account (describing observed performance, rather than underlying mechanism, see Section 2.1.2) of how the emergency management team worked in its collocated setting using a classification of functions adapted from the Fleishman & Zaccaro (1992) team performance taxonomy (TPT; see Section 2.3.3). This specific taxonomy was used as a starting point, but modified during the coding of the functions. After establishment of a more developed set of functions, notes from observations by both observers were used to independently rate the frequency of function performance using each of the artifacts available, and discussing coding differences to resolve and agree upon the categories. Paper I relates the team functions to the artefacts used during the exercise and discuss the characteristics of the artefacts in relation to work in collocated and distributed settings. The idea for this specific orientation emerged during and directly after the exercise, as it was until quite late in the exercise unclear which parts of the exercise were accessible for study.

3.3 Microworlds

Microworlds are simulated task environments that aim to strike a middle balance on the scale of correspondence, simulating several aspects of a class of systems. The origin of microworld research is the proposal of Brehmer & Dörner (1993) that the investigation of intelligence should be conducted using simulated complex problem solving scenarios in the laboratory (Funke, 2001). They proposed the use of microworlds, environments that (a) provide a task that can be made more complex, challenging, and realistic than traditional laboratory studies but that (b) generalize to interesting parts of real world problem solving while remaining (c) controllable and more easily analyzed than field study.

Microworlds (Brehmer & Dörner, 1993; Brehmer, 2005b; Dörner & Schaub, 1992; Funke, 1993, 2001; Gray, 2002; Rigas et al., 2002) can be characterized along the dimensions of complexity, dynamics, polytely, and intransparency. Complexity refers to the number of variables and causal relations in the task. Dynamics refers to the degree of autonomous change of the controlled system, polytely to the degree of necessity of making trade-offs between multiple potentially conflicting goals, and intransparency to the fact that not everything in the system is directly observable. Tractability of microworlds (Gray, 2002) may be seen as the manageability of the microworld for the experimenter/researcher. By enabling the experimenter to manipulate these four dimensions, microworlds provide a much more tractable, reproducible, and flexibly designable research environment than a field study. Experimental participants take well-designed microworlds seriously and become engaged in the task, so that their behavior becomes valuable to the researcher (Funke, 2001; Gray, 2002). These considerations
make microworlds appropriate test beds for studying the impact of constraints on actions in command and control in complex dynamic environments. The microworlds that have been employed in this study are C³Fire and DKE. Although the microworld experiments that were analyzed here were partly performed to test hypotheses (see (Lindgren, 2007) for the C³Fire study of cultural effects on emergency management and (Kuylentsierna et al., 2006) for the DKE study on role- versus need-based information in distributed decision making), here the experimental data is used, as is also common for experimental games, to "explore decisionmaking processes and to generate hypotheses rather than to test specific hypotheses" (Shubik, 1972, p. 30). Classically, outcome measures are used to test hypotheses, but recent research has also indicated the need for process measures to explore dynamic decision making processes (Johansson, 2005).

3.3.1 C³Fire

The fire-fighting microworld C³Fire (Granlund, 2002, 2003) was used for the initial investigations into constraint modeling in civilian (emergency management) command and control.

C³Fire is a fire-fighting microworld in which a group of people work together as a team to gain control of a computer-simulated fire. Their task is to collaborate in an experimentally controlled configuration for command and control interaction, under observation of a researcher who manages the experiment. Paper II presents an overview of the C³Fire microworld. Its elements include a map showing vegetation, buildings, vehicles, and the fire, and the computer network for controlling the vehicles and communicating by e-mail.

In the C³Fire setting used in this experiment, various mobile units and stationary units are located on the map. Mobile units (trucks) need fuel to move across the map. Stationary units have a fixed position on the map and cannot move. Fire trucks are mobile units that can fight fires. To do so, they need water. Water providing units can provide other units with water. This class is formed by water trucks and water stations. Similarly, fuel providing units are units that can supply all moving units with fuel. These can be divided into fuel trucks and fuel stations. The class of stationary units consists, besides of water and fuel stations, of vegetation and buildings. The units that are part of the microworld thus create interdependencies between the units, and thereby some units constrain some other units’ function performance possibilities. When control of the vehicles is distributed across the team of people engaged in the microworld, the team members’ actions place constraints on each other, and these constraints need to be managed.

Constraints such as the rate of burning and spreading of the fire can be set by the researcher, and is typically made dependent on vegetation, terrain, the presence of buildings, and wind direction and speed. Thus,
these properties of the environment constrain the development of the fire. There were a total of eight scenarios that were played once each. Two maps were used to introduce a restricted variation in difficulty. These maps were rotated in order to keep the participants from getting familiar with the specifics of the maps. The initial size of the fire, which was started a few seconds after the start of the session, was also manipulated at two levels. The last four scenarios were made more difficult by manipulating additional variables in C3Fire. The manipulations were a higher fire spread rate, a lower vehicle speed, a longer fighting time, and a shorter burn-out time. The purpose of these manipulations of difficulty was to keep the scenarios challenging and the participants engaged.

Participants direct the units where to drive by interacting with an interface showing a map with the dynamic simulation, an e-mail tool, and some data describing the state and characteristics of the trucks.

Paper II describes constraint management through a method for generating ecological representations of spatial and temporal resource constraints in network-based command and control, and illustrates its application in C3Fire.

The method to recognize constraints, as it is applied here, has four steps. The first step in the method is a goals-means task analysis (GMTA, see Section 2.1.5) of the joint cognitive system’s process control task. The method initially used only goals-means task analysis (GMTA; see Section 2.1.5) to extract the essential variables that describe the behavior of the C3Fire participants in relation to constraints. This study has been reported earlier in the author’s Licentiate thesis (Woltjer, 2005). The method was subsequently enhanced by including the functional resonance analysis method (FRAM; mainly steps 1 and 3, see Section 2.1.5) as a basis for representation design. Analyses were performed by the first author (Woltjer). This first step facilitates the second step, that is, finding functions in the task that joint cognitive systems perform. GMTA may be used to obtain these functions. The means identified in the goals-means analysis were used in the initial definition of functions used in the second step of the method, together with descriptions of events across time from log files.

The second step in the method is a functional analysis of the joint cognitive system’s process control task, following FRAM (see Section 2.1.5). First, the functions that the joint cognitive systems perform are listed and their aspects – Input, Output, Preconditions, Resources, Time, and Control – are described (FRAM step 1). Couplings between functions may then be found by matching all functions’ aspect contents with each other. For example, the Input of one function may be the Output of another, or the Output of one function may be the Precondition of another. Then, graphical representations (with the hexagonal representation) may be created to illustrate some of the couplings found. These representations should be seen as illustrations of parts of the functional model and potential couplings between functions. The model itself is described in the tables of functions and their
aspects.

The third step in the method is to identify the essential variables, the subset of process variables that the joint cognitive system can both observe and (indirectly) affect. This step in the method identifies the variables that change in the performance of the tasks identified in the FRAM analysis. In the process control domains that are of interest to cognitive systems engineering, the units of observation are typically states, defined by the current instantiation of a large number of variables that can change rapidly over time. The term essential variables is used here for the subset of process variables that the joint cognitive system can both observe and (indirectly) affect, as discussed in Section 2.1.1. These variables may be described for all of the aspects of the FRAM functions, that is, the constraints associated with time, control, preconditions, and resources can be identified by identifying variables for each of these aspects of each function. It is implied that essential variables must be observable or known for the joint cognitive system to function adequately.

The fourth step generates recommendations for how the joint cognitive systems can be made more aware of the constraints on their actions, more able to adapt to the constraints on their actions, and more aware of their abilities to adapt and affect these constraints. This may result in 1) implications for redesign of the task, such as changes in procedures, task allocations, etc., or 2) changes in the representation design, implemented on displays and interfaces that are used by the people in the joint cognitive system. The second alternative was investigated further here, using state spaces (see Section 2.2.10). Generating state spaces entails three steps: (a) sampling the values of the essential variables (b) juxtaposing these variables to form a set of multidimensional state spaces, and (c) comparing the location of the data in the state spaces to the thresholds between regions specifying differing opportunities for action. State spaces juxtapose variables associated with related tasks in a manner that makes constraints or possibilities for action explicit by specifying regions where the constraints are (not) met. State spaces of actual trials of experimental data were generated, which were subsequently analyzed for their usefulness, and linked back to the functional description of the task to document the steps in the method, from functional analysis to representation design highlighting constraints.

3.3.2 DKE

DKE (Dynamisk Krigsspel för Experiment, Dynamic Wargame for Experiments) is a wargame microworld, which means that there are two adversaries (the blue team and the red team) battling against each other in a simulated environment. The DKE research team at Swedish National Defence College (FHS) has recently conducted a study with DKE in which two teams of three people (one commander and two subordinates) played
against each other. The wargame, and this particular FHS-study’s scenario and data are described and analyzed in Paper III.

The subordinates in each team maneuvered the units, while the commanders could only monitor the game and give orders. All players had access to a computer interface. A screenshot of DKE is presented in Paper III. The simulated battle environment contained a map that is horizontally divided with both teams’ headquarters in the leftmost and rightmost upper corners. The map contained a river flowing roughly up-down in the middle of the map with three bridgeheads on either side and roads between the headquarters and the bridgeheads across the river. The objective of the game was to conquer (meaning to advance beyond) one bridgehead and secure a road to the own headquarters from the conquered bridgehead. Constraints on own forces and adversary forces need to be managed in order to complete the game successfully.

The data collected consists of replay-able log files of movements and actions of units, and audio recordings of each of the team members individually. Paper III describes constraint management in military activity and command and control through a study of adversary war-gaming in DKE. Observations of the replay of a part of the experimental data and a preliminary discourse analysis of a communication transcript (by the first (Woltjer) and (mainly) the second author (Prytz)) were the input to a description of agile command and control at a conceptual level using the functional resonance analysis method (FRAM; mainly steps 1 and 3, see Section 2.1.5) (led by the first author (Woltjer)). FRAM as a method for military C² performance analysis was developed further through this study.

The analysis started by describing the functions at the tactical level of military activity (e.g., moving, fighting) in the FRAM functional description of input, output, preconditions, resources, time, and control. The functions and their aspects were identified and described mainly through analysis of the rules of the game, and noted in a semi-formal notation to avoid ambiguity. The function descriptions were then verified and refined through replay of one 30-minute trial and transcription in minute detail of the timing and specific aspects of the functions being performed, resulting in more than 1200 entries combined for both sides (red and blue). All instantiations of each tactical function were thus documented for this trial.

The analysis was then continued to the higher-level operational functions, which could be named and described through the replay of the experimental trial which used for the tactical functions. The function descriptions were verified and refined based on comparison of function descriptions with communication transcripts, and further coding of operational functions in two other trials, resulting in about one hundred entries per side (blue and red) per trial.

Thereafter, potential couplings between tactical functions, between operational functions, and the relationship between tactical and operational functions were described, as part of step 3 of FRAM. A preliminary analy-
sis of the blue side communication transcripts in the trial that was fully transcribed in terms of function instantiations at the tactical and operational levels was used to describe the command and control functions of the DOODA loop, and to investigate the relationship between between command and control, operational, and tactical functions, and ultimately, agility in command and control.

3.4 Case histories

Papers IV and V describe and model the Norwegian Air Shuttle incident and the Alaska Airlines accident (described in the Introduction) based on the case histories as described in investigation reports (AIBN, 2004; NTSB, 2003).

Hindsight, or knowing of the outcome of an event (Fischhoff, 1975), as well as pre-assumptions about how accidents typically happen (accident models), affect perception and judgment of contributing factors in accident and incidents (see Section 2.4.2). Researchers, as well as any reader of investigation reports, thus need to take caution when interpreting incident and accident data from these reports. Leveson (2001) discusses the limitations of studying accident cases through investigation reports. People’s accounts of accident causal factors are affected by their model of accidents, hierarchical status, position in an organization, degree of involvement, various types of pressures, and job satisfaction. Also, the data collection methods and reporting systems that are used, whether the event is an incident or an accident may affect which causes are filtered out or emphasized, possibly over-simplifying the event description (Leveson, 2001). This corresponds to what Hollnagel has termed the WYLFIWYF-principle (What You Look For Is What You Find), which Lundberg et al. (in print) argue describes the fact that investigators find the contributing factors that their underlying accident model constrains them to look for, based on investigation manuals and their application.

Thus, the focus is on the analysis of the incident and accident reports, and on the analysis of the incident and accident data as presented in these reports, keeping in mind that incident and accident data is necessarily filtered during investigation and report compilation. However, the analysis of the investigation reports gives information about which information is (not) included in the report, and thus which information is required and elicited about the event by each of the analysis methods employed.

3.4.1 Norwegian Air Shuttle flight 541

The Norwegian Air Shuttle flight 541 case (see Section 1.2.2) is particularly interesting because the factors contributing to the accident are not obvious and easily established. Possibly for this reason, the Norwegian investigation board’s (AIBN) suggested the authors of Paper IV to look into this case.
The main method employed was textual analysis of the investigation board report (AIBN, 2004), complemented by other methods of data collection.

The sequentially timed events plotting method (STEP; see Section 2.4.4) was applied to the description of the incident in the incident report and data collected using other methods, as well as the functional resonance analysis method (all steps to some extent, see Section 2.1.5), and the results were compared. The STEP analysis was performed mainly by the first author (Herrera), whereas the FRAM analysis was the main effort and performed jointly by both co-authors, as well as the comparison. FRAM as a method for incident and accident analysis was developed further through this study.

The FRAM analysis started by identifying functions that were mentioned or implied in the accident report through textual analysis. The data that was explicitly or implicitly described with regard to the function aspects was entered into the model. The FRAM Visualizer\(^1\) was used at this stage to keep track of all functions and to assure consistent naming of function aspects. The analytical task of finding of potential couplings between functions, part of step 3 of FRAM, was performed simultaneously with step 1 as the FRAM Visualizer continuously updates potential couplings while function aspects are being entered. These were continuously checked with descriptions in the investigation report. Step 2 of FRAM, defining variability of normal performance, was done through textual analysis, using the common performance conditions that were also administrated in the FRAM Visualizer. Gaps that emerged in the function descriptions in step 1, the common performance conditions in step 2, were noted for further inquiry.

After that steps 1 and 2, and the first part of step 3 (finding potential couplings) were completed based on textual analysis resulting in a FRAM Visualizer model, the visualization of the functions with their potential couplings was transferred to a diagram in Microsoft Visio. This software provided more flexibility in the marking of which actors performed which functions and spreading of variability through function couplings. After several iterations in the development of the model, recommendations of the investigation report were mapped onto the description of the system as it had emerged through FRAM, and compared to barriers and monitoring that step 4 of FRAM could generate. Since this step was the least well-developed step in the method as it had been described thusfar (Hollnagel, 2004), more interpretation and improvization of this step of the analysis was necessary in comparison to the other steps, which was not pursued to great detail, as the focus was here not to develop this step of FRAM in a more systematic manner. The STEP analysis was developed at the same time as the first stages and iterations of the FRAM analysis, and could now be compared to the result in FRAM.

Questions and results on the individual models and the comparison between methods were collected. The models were complemented by information gathered through a visit to the Oslo Gardermoen control tower including a short observation and unstructured interview, and other unstructured and semi-structured interviews with operational personnel. A workshop with the AIBN was organized and performed. In this manner, the analysis with STEP and FRAM was an iterative process between the authors and operative personnel, involving a total of about 50 people, including air traffic controllers, pilots, and accident investigators. Feedback on the FRAM model was also obtained during the 2nd FRAM workshop. After each time feedback on questions or modeling results were obtained, the models and comparison were iterated into new versions.

### 3.4.2 Alaska Airlines flight 261

The Alaska Airlines flight 261 case (see Section 1.2.2) is particularly interesting for the long time period that needs to be described to get an adequate understanding for the development of the contributing factors to the accident. Moreover, the NTSB (2003) report documents a wide range of contributing factors in the Human, Technology, and Organization (MTO) categories, which in principle enable a more systemic investigation than other reports that may oversimplify or filter out the organizational factors. The development of the FRAM analysis method (see Section 2.1.5), as it is used here for accident analysis and understanding (touching on all steps of the method), was largely performed using this case, and as such the description and model that were developed was iterated upon at times during the course of a few years, updating and revising the model while the theoretical and practical aspects of FRAM became more clear and detailed, based on their initial description (Hollnagel, 2004), simultaneously developing the method. The study presented in Paper V thus had the dual purpose of providing insight in a complex accident development, as well as developing FRAM further as an accident model and an analysis method. The analysis was performed mainly by the first author (Woltjer).

The steps of FRAM were applied for Paper V in a similar manner as the description of how the study of Paper IV was performed, but over a longer time period, with a different focus, and using different administrative tools. The FRAM analysis started after an initial analysis of barrier systems and barrier functions that played a critical role during the course of events up to and after the accident. As FRAM was selected for the analysis of C3Fire in 2006, after its initial description with GMTA, to more explicitly capture constraints on functions and essential variables (see Section 3.3.1 above), the FRAM analysis on Alaska 261 was started in order to have a better understanding of the accident development.

---

2This analysis was performed as part of a doctoral course on System Safety and Control of Risk, lead by Erik Hollnagel, Linköping University, Graduate School for Human-Machine Interaction, 2004.
understanding of FRAM through its application for a purpose that was closer to the method’s initial purpose, the analysis of system safety.

As not many FRAM analyses had been performed at that time, the model developed over several iterations, but as a visual model initially. Gradually the insight that the visual representation of the model should not be the end or possibly not even the means of FRAM as a method became stronger, and the focus was shifted to the description of functions in tables in step 1 of FRAM, continuing to the first part of step 3, finding potential couplings, to verify and refine the function description of step 1. Steps 1 to 3 were the focus of analysis after the initial barrier analysis, all performed through textual analysis of the NTSB investigation report. After some time the FRAM Visualizer was used to administrate the model, combining this with Microsoft Visio for more flexible visualization. Steps 1 through 4 were iteratively applied, verified against the data in the accident report, and further developed from their initial specification. During the analysis process, preliminary versions of the model were presented and discussed at the first and second FRAM workshops, at the International Symposium on Aviation Psychology (Wolter & Hollnagel, 2007), and during workshops on FRAM at that symposium and at the 2008 International Symposium of the Australian Aviation Psychology Association.

3.5 Natural task environments and simulations

Whereas the other studies report on either natural or simulated environments of discovery, the ERASMUS project combined both of these kinds of environments. This was important because of the continuous iteration of the concepts of the automation that was being developed during the project, and constant calibration of these concepts while being exposed to examination in operational reality and experimental studies. Thus, the project is an example of the continuous development of an early prototype of a system with continuous input from studies in natural and simulated environments.

3.5.1 ERASMUS

The methods used in the development of the ERASMUS automation tool for air traffic control (see Section 1.2.2) started with table-top analysis of the nature of work in air traffic control, specific models that have been published as a result of earlier task analyses (see Section 2.4.4), and documentation of the ERASMUS concept in project deliverables such as the Concept of Operations (ConOps).

The ERASMUS project proceeded with experimentation in “human-in-the-loop” simulations (see, e.g., Vicenzi et al., 2008, for an overview) of the functioning of ERASMUS in cooperation with controllers, and pilots, performed in the laboratories of the Direction Générale de l’Aviation Civile
(DGAC), Direction des Services de la Navigation Aérienne (DSNA), Direction de la Technique et de l’Innovation (DTI) Research & Development at Rangueil, Toulouse, France, and the South-East Area Control Center (ACC) of France located in Aix-en-Provence. The experimental conditions at both sites used operational data (e.g., actual traffic samples, sectorization) from the Aix ACC. A recent project deliverable (Drogoul et al., 2009) discusses four experiments (numbered 1, 2, 3, and X) where controller perspectives were evaluated, and one (experiment 4) where pilot perspectives were addressed. Experiments 1 and 2 employed low-fidelity laboratory studies. Experiments 3 and X employed high-fidelity air traffic control positions with executive and planner positions and pseudo-pilots and in the case of experiment 4 an adjacent-sectors planning position. Experiment 4 employed a high-fidelity Airbus A320 cockpit simulator. Together, these experiments form the real-time simulation human-in-the-loop evaluation and validation effort of ERASMUS in a baseline, year 2007 environment. Observations, interviews, and quantitative experimental methods for subjective evaluation of workload, situation awareness, task complexity, difficulty, safety, etc., were collected and analyzed (see (Drogoul et al., 2009) for details). Operationally active controllers from Aix ACC, and pilots, were the participants in all of these experiments. The first author (Woltjer) participated in the observation of several trials of experiments 3, X, and 4. Detailed experimental and scenario design and quantitative experimental analyses were mainly performed by other partners in the project, but both co-authors of Paper VI participated experimental design meetings, and in the compilation, write-up, and summary of the experimental findings in reports.

A demonstration of what the ERASMUS concept(s) could look like in a year 2020 SESAR-inspired (Single European Sky ATM Research) environment with various operational prototype concepts was also performed, labeled experiment 5, where participants were asked for their feedback on the demonstrations of ERASMUS concepts and tools in a workshop environment. Controllers and pilots, as well as developers of technical systems and managers, from several EUROCONTROL-member countries, participated and gave their feedback (see Bonini & Drogoul, 2009, for details on the tools and the results). The author participated in one day of the workshop. The 2020 tools were however not included in the modeling effort of Paper VI.

The natural task environment part of the study included visits to three European area control centers which operationally worked with medium-term conflict detection (MTCD) tools or had done test trials with such technology. The author participated in two of these one-week visits. The visits started with observation at controller positions rotating through a number of diverse sectors, for the first one or two days, after which the remaining days were filled with interchanged observation and interviews. The observations focused on separation management, and specifically on the
use of tools at the position as well as communications between controllers (executive-planner within a sector, and between sectors) for this purpose. Semi-structured interviews (with Dr Deirdre Bonini as co-interviewer) focused on issues related to current practise and tools for separation management and the potential role that ERASMUS could play in separation management of the future, using general questions and specific scenarios. Interview questions and scenarios were used as a starting point and different issues were elaborated on in varying depth depending on the answers of the controllers.

The resulting understanding of the joint cognitive controllers-pilots-aircraft-ERASMUS system resulted in a model using FRAM as a method (see Section 2.1.5, mainly steps 1 to 3). It was used as a means to indicate and evaluate the effects and impact on controller and pilot work resulting from ERASMUS automation. FRAM as a method for risk assessment was developed further through this study.

Step 1 of FRAM, the description of controller functions, is based on the observations of activities of controllers in the natural and simulated task environments that were observed, as well as on the table-top analysis of air traffic control task analyses. The ERASMUS functions and pilot and aircraft system functions were derived from project deliverables and discussions.

Step 2 of FRAM, the evaluation of the potential for variability for each of the functions, was performed using the same data and information. Common performance conditions were used to evaluate potential variability, as well as the newly introduced variability phenotypes, similar to failure modes as known from CREAM. In cases where data is available about actual performance of the functions, the CPCs may be used as a checklist to assess functions as they are actually performed in practice. In the present case of evaluating an envisioned system, a CPC assessment in this way is more difficult because data on actual use were collected during simulations during the project and are necessarily very limited due to the setup of the experimental conditions. Moreover, the CPCs are meant to describe the actual context and situational conditions of the use of a system, and many of these (procedures, organizational support, working conditions) in experimental settings are not (yet) comparable to how the system would actually be used, even when assuming that all other systems and circumstances do not change. Thus, instead, the potential dependence and sensitivity of the identified functions to the CPCs was assessed. Step 2 also added the evaluation of variability phenotypes, derived from the failure modes described in Section 2.1.5 as part of the variability evaluation. To do this, each of the function aspects was evaluated for in which form variability would manifest itself. This is an addition to FRAM as previously described.

Step 3 was performed by matching function aspects in order to discover potential couplings between functions. The evaluation of functional resonance aimed to combine the functions’ sensitivity for the CPCs in displaying variability, as well as the variability phenotypes. Step 4 of FRAM was
not further developed, but identified risks using the method were listed and taken into account in further project discussions.

3.6 Summary and progression

When looking back and analyzing the philosophical standpoints taken in the six studies, the reader may note a progression from a mainly positivistic to a mainly interpretivist way of describing complex socio-technical systems. The command and control studies were much concerned with representativeness and constructing an "accurate" model of the functions that were investigated, focusing on facts and not considering differing world views that the participants may have had. The aviation safety studies, mostly being performed later than the command and control studies, develop a more interpretivist approach where it is important to go back "into the tunnel" and describe the view "from the inside" (Dekker, 2002) how people at the time may have interpreted the world around them. The descriptions of functions in these studies is more interpretivist, a stance that seems necessary to describe and explain the variability in performance that was observed and described in the case histories (Papers IV and V) and scenarios of system use (Paper VI).

Table 3.1 gives a summary of the studies that are described in the appended papers, and their purpose, the domain, the main task that was studied, the research setting, and the methods and modeling employed.

It should be noted that when FRAM is indicated, FRAM as a method was developed further during the course of the studies, so that the method "FRAM" that was applied was different in each study.
Table 3.1: Overview of studies described in appended papers, and their purpose, domain, main task studied, research setting, methods, and modeling employed.

<table>
<thead>
<tr>
<th>Paper</th>
<th>Purpose</th>
<th>Domain</th>
<th>Main Task Studied</th>
<th>Research Setting</th>
<th>Methods</th>
<th>Modeling</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.6. SUMMARY AND PROGRESSION
Chapter 4

Results and analysis

The results of the studies are discussed here, in two parts, first for the command and control studies, then for the aviation safety studies.

4.1 Command and control

This section describes a model of constraints in command and control. It is based on the theoretical framework outlined in Chapter 2.

4.1.1 Electricity network resource management exercise

Paper I presents a functional account of how the emergency management team worked in its collocated setting using a classification of functions adapted from a taxonomy of team performance (Fleishman & Zaccaro, 1992, see Section 2.3.3). Paper I relates the team functions to the artefacts used during the exercise and discuss the characteristics of the artefacts in relation to work in collocated and distributed settings. The seven categories proposed by Fleishman & Zaccaro (1992) (presented in the leftmost column of Table 1 of Paper I (p. 342)) aim to cover the full range of managerial activities one expects to find in team work, meaning the functions that are used for getting work done as a team, rather than the actual functions that are performed as part of the task for which the team is assembled.

The team’s task was to work as a coordination team, facilitating the cooperation between the companies. To do this, the team had to:

- collect information,
- assess the situation before and during the power failure,
- distribute material, equipment, personnel and other resources from other parts of the country to the affected area,
• maintain continuous contact with the agency responsible for the power grid, and
• find the people and competence needed to cope with the given situation.

Combining this task description with the taxonomy by Fleishman & Zaccaro (1992), the task that the team set out to perform overlapped with the functions necessary for the group to function as a team. For this reason the function categories could combine team coordination functions and team task functions. Where necessary these were split up into internally and externally directed functions, as for the functions of "information exchange" and "information conflict resolution".

The team usually works in a distributed manner. They typically stay at their home offices and coordinate their work through telephone conferences during which they discuss and present updates on the developing situation. The team has a meeting location for highly critical emergencies, but so far the members have never used it. During the exercise, the team was collocated and had two rooms at its disposal. Thus, the exercise setting did not match the actual work practice.

Since the team worked in a collocated manner, they had shared access to some of the artefacts in the rooms to perform these functions. Paper I reports on the estimated relative frequency with which the team performed each function using the various artefacts. Face-to-face communication was included in this analysis to highlight its salient role in collocated work.

The whiteboard and face-to-face communication were the most frequently used media for all categories of team functions. The whiteboard and face-to-face communication can be used only when the team is collocated. When comparing the communication media (including ICT artefacts), it becomes clear that the remaining artefacts lack certain important characteristics for the utility team’s communication. Paper I argues that these characteristics are essential in team performance, based on the literature on collocated work, and collocation enables a shared local context, informal interactions between team members, gestures, coreference to objects and implicit cues about peripheral activity. These characteristics clearly seem to be beneficial given the frequent use of face-to-face and whiteboard communication during the exercise.

The paper provides the rationale for three claims:

1. The team should benefit from being able to communicate face-to-face and develop trust and common ground.
2. They can develop patterns of interaction that may transfer to distributed work and, hence, enhance operations.
3. However, the communication, task allocation and artefact (ICT) use during training may not readily transfer to their actual distributed work setting.
The observational study sheds light on the first and the last of the three claims. There was no data to document any transfer of benefit to the team’s actual distributed work setting, because observations during actual operations were not available and in general practically difficult to perform.

The frequent face-to-face communication and whiteboard use strongly suggest that the team benefited from the collocated setting. The continuous shared referencing to the whiteboard and the Power Grid Cooperation System reveals that the team developed a consistent pattern of interaction that guided virtually all functions.

Moreover, the emergence of fluid subgroups and social (non-professional) conversations is consistent with the development of mutual trust and common ground. None of these social aspects were mediated by ICT. The observations during the study suggest that ICT can be useful for the practice of emergency management. However, it is difficult to imagine that ICT would be used to mediate social interaction during distributed emergency management. Thus, the availability of ICT would be a constraint, both enabling coordination and cooperation, partly able to replace meeting face-to-face, and limiting coordination and cooperation, partly disabling the advantages that collocation supply. Functions are thus constrained, enabled and limited, by ICT and collocation.

The observation that the collocation-enabled modes of communication by the emergency management team are most frequent suggests that there may be a mismatch between what the team learned during training and what they need on the job. Constraints in the training environment did not match constraints in an actual work environment. In addition, because tasks were assigned given the collocated work conditions, function allocation in training may not correspond to or be appropriate for actual distributed work. Moreover, collocated interactions are instrumental for the development of trust and common ground, distributed teams may be at a serious disadvantage that technology may not be able to overcome.

Even though the observed exercise differed markedly from how the team usually works, the benefits of the exercise for future emergency management are clear. The experience of working in the same physical location most likely helps team members get to know each other better, which in turn can lead to the development of trust, common ground and more efficient communication, which were part of the goals of the exercise. Exercises like the one observed may also enable emergency management workers to discover new ways of working that can be incorporated in their organization of work. A training programme including carefully designed exercises in both collocated and distributed settings is thus advisable.

### 4.1.2 Fire-fighting emergency management

Joint cognitive systems are systems of people and artifacts that together perform some function or have a common goal. In the case of command
4.1. Command and Control

and control, many joint cognitive systems (JCSs) with different boundaries may be defined, depending on the (often dynamic) allocation of functions across people-artifact systems. This description takes the C\textsuperscript{3}Fire setting as an example of emergency management and models the JCSs, functions, and constraints in this environment.

The model thus focuses on a fire-fighting task in which four commanders have six fire trucks, three water trucks, and three fuel trucks at their disposal to control a fire. Fire trucks need water to fight fires and all trucks need fuel to move around. Water is provided by water trucks and fuel by fuel trucks. This command and control situation is in principle simpler than military command and control because the fire does not intentionally try to weaken the command and control system. In such a fire-fighting task, one may for example have devised functions so that there are four joint cognitive systems (JCSs) each consisting of one person (the commander) and three vehicles, two in charge of fire fighting, one of water supply, and one of fuel supply. Each of these may have specific functions and goals: for example two fire-fighting JCSs that fight fires in geographically different locations, a water-providing JCS, and a fuel-providing JCS.

Note that from there on one may define other JCSs: These four JCSs then can be grouped to a larger JCS which may be called a fire-fighting command cell, which in turn may be part of a larger network of emergency management command cells. On the other hand, a vehicle within our three-vehicle-commander-JCS may be operated by a driver and an operator which together form a vehicle-driver-operator-JCS, etc. The definition of JCSs is defined by goals and functions, and it is important to determine what may be controlled by which JCS and what not. If parts of the environment of a JCS can be controlled, this environment and JCS may be part of a larger JCS. An example is when (unlike in C\textsuperscript{3}Fire) ‘combat engineering’ troops are available and the terrain or infrastructure may be adjusted to the command cell’s needs. The infrastructure then becomes part of a larger JCS including command cells and combat engineering troops. These JCS boundaries are often dynamic, for the combat engineering troops may for example be reallocated (even during function performance) to another part of the command and control system and serve the purpose of another command cell.

Regarding the wildfire as the process to be controlled, joint cognitive systems with various boundaries at different levels take action on this process and aim to meet the variety of the process to be controlled (as in the Law of Requisite Variety, see Section 2.2.2). Since controlling means steering the behavior, or meeting the variability and reducing it to fit the goals of the JCS, one can also say that the JCS aims to constrain the behavior of the process to be controlled. Figure 4.1 illustrates this concept.

The various units in C\textsuperscript{3}Fire are mutually enabling and constraining because of interdependencies in the consumption and provision of water and fuel. Interdependencies among decision makers arise whenever different
classes of units are assigned to different participants in the simulated command and control system. For example, if fire trucks need water and fuel, water trucks and fuel trucks constrain the actions of fire trucks. If different people have control over these different units, their actions are mutually constraining, both enabling and limiting each other’s actions.

The method to recognize constraints, as it is developed and applied here, has four steps, which were discussed in Section 3.3.1.

The first step in the method is a goals-means task analysis (GMTA) of the joint cognitive system’s process control task. The analysis is conducted by (or in consultation with) a domain expert. This method has been described in Section 2.1.5. The goals-means task analysis of C²Fire is shown in Figure 4.2.

This first step facilitates the second step, that is, finding functions in the task that joint cognitive systems perform. GMTA may be used to obtain these functions. The means identified in the goals-means analysis may be the initial functions used in the second step of the method. However, if there are other descriptions of work available, such as communication protocols, descriptions of events across time, or task analyses of other kinds, these may also be used as a starting point for the function descriptions.

The second step in the method is a functional analysis of the joint cognitive system’s process control task, following FRAM, as described in Section 2.1.5. This analysis is also conducted by (or in consultation with) a domain expert.
4.1. Command and Control

**LEGEND**

- **GOAL**: achieved by
- **TASK**: may require
- **PRE-CONDITION**:

Fire trucks have enough water to fight fire

- Move fire trucks to fire location
- Refill fire trucks' water tank

Fire trucks at fire location

- Fire trucks have enough water to fight fire

Fire closed out

- Fire trucks fight fire

Move unit U to location L

- Unit U has enough fuel to move to location L
- Refill unit U's fuel tank

Unit U has enough fuel to move to location L

- Move unit U to location L
- Refill unit U's fuel tank

Figure 4.2: Goals-means task analysis of the C3Fire fire-fighting task.
expert.

Three operational functions that trucks may perform were defined, \textit{truck T fights fire}, \textit{truck T moves to location L}, and \textit{truck T refills resource R from source S}. Based on these functions, functions may be instantiated with the elements of the simulation, such as \textit{fire truck fights fire}, \textit{fire truck moves to fire}, \textit{water truck moves to fire truck}, \textit{water truck refills fuel from fuel truck}, and \textit{fuel truck refills fuel from fuel station}, describing the functions that may be performed to accomplish the goals that the GMTA revealed. With even more specifics used for instantiation of functions, one may specify identifi-

able elements and time-stamped events in the simulation, such as \textit{fire truck 1 fights fire at (4,15)}, \textit{water truck 7 moves to fire truck 3}, and \textit{fuel truck 11 refills fuel from fuel station at (17,23)}.

Part of a functional modeling analysis of C$^3$Fire is shown in Figure 4.3, giving an example of each of the three functions. The analysis reveals how functions are linked and thereby how units are mutually enabling and con-

straining. To take the example of the ‘fire trucks fight fire’ function, the input of this function is burning vegetation and buildings, which may be saved from or lost to the fire (the output of the function). The preconditions of fighting a fire are (1) that the fire truck has water (which is the output of the water refilling function) and (2) that the fire truck is collocated with that fire (which is the output of the moving-to-fire function). The time and pace of fire trucks fighting fire is influenced by the fire spread rate and fire fighting rate for example. The main resource used for fire fighting is water.

![Diagram of FRAM critical function analysis of C$^3$Fire fire-fighting task.](image)

Figure 4.3: FRAM critical function analysis of the C$^3$Fire fire-fighting task.

The third step in the method, is to identify the essential variables, the
subset of process variables that the joint cognitive system can both observe and (indirectly) affect. This step in the method identifies the variables that change in the performance of the tasks identified in the FRAM analysis. Paper II develops how essential variables may be identified based on the descriptions of functions with FRAM.

The fourth step generates recommendations for how the joint cognitive systems can be made more aware of the constraints on their actions, to overcome problems that would result from a difference between actual and perceived constraints. These representation designs should facilitate the discovery of constraints and highlight the dynamic character of constraints, as discussed in Sections 2.2.9 and 2.3.6. The specific technique for highlighting constraints developed in Paper II is the state space (see Section 2.2.10). Paper II presents several detailed examples of state spaces that are the outcome of this step. Partial descriptions of the joint system form states (coordinates) in the state space, and the trajectory connecting these states describes part of the behavior of the joint system over time in relation to the constraints.

The documented method in four steps from an analysis of goals, to functions, to essential variables, to constraint-based state space displays may thus be seen as a result of the study presented in Paper II.

4.1.3 Wargame DKE

The concept of joint cognitive systems that aim to control a process does not change significantly when stepping from emergency management command and control to military command and control.

Figure 4.4 sketches a description of the idea of how joint cognitive systems attempt to constrain their adversary’s functions in an adversarial command and control setting. JCSs may be defined with various boundaries, depending on the functions they perform. For example, an artillery unit may perform tactical functions such as moving and firing artillery, and is composed of cognitive systems (a commander and subordinates) and artifacts (the physical artillery machinery). At a higher level, several such artillery units and other ground vehicles may form a JCS, performing operational functions such as defending a city or taking a bridgehead. In an adversarial command and control situation, agile command and control means then to develop and adapt strategies in an efficient manner, and translate these into effective performance of operational and tactical functions that constrain adversary function performance, so that own objectives are reached. In FRAM this translates to affecting the adversary’s function aspects (indicated conceptually by the thunderbolts in Figure 4.4, using the opponent’s colors for indicating constraint on adversary function aspects) so that their outputs cannot be realized or only outputs beneficial for own objectives are made possible.
RESULTS AND ANALYSIS

Figure 4.4: Joint cognitive systems in an adversarial C² task aim to constrain or control the opponent.
Building on the levels of military activity, and on the DOODA loop (see Section 2.3), three categories of functions can be identified. (1) Tactical functions (e.g., move, fight) are performed by the units in the simulation, and are observable when data is available. (2) Operational functions (e.g., take bridgehead, secure road) are inferred and ascribed to groups of units that aim to reach operational objectives and consist of various performances of tactical functions. (3) C² functions (data collection, sensemaking, and planning, according to the DOODA loop) are general functions that enable the exercise of command and control. C² functions are not directly observable but may be inferred, by studying groups of units and the behavior of the commanders that comprise the C² system (e.g., retrieving data, giving orders, discussing tactics).

Detailed examples of functions in each of these categories are given in Paper III, of how functions in the same category within one force and between opposing forces may be coupled in enabling or limiting relationships, as well as how functions in the three categories may be interrelated. These descriptions are related to experimental data from the DKE microworld.

Paper III documents a modeling approach to the understanding of functional interdependency and constraint management, that addresses the challenges of (1) the adoption of a detailed description of interdependency and associated understanding of interdependent functions (Brehmer, 2007) and (2) the application of that description to both own and opponent forces’ opportunities and vulnerabilities to provide for agility and resilience. FRAM is shown to be able to describe the C² functions of the DOODA loop and the tactical and operational functions of military activity. FRAM has been applied to modeling functions of own and opponent forces in DKE to reveal and characterize both agile and unsuccessful C² practice.

More specifically, the paper shows that:

- FRAM’s recursive way of functional modeling is suitable for modeling functions at various levels, such as the tactical, operational, and C² (strategic) levels,
- The DOODA loop may be developed into detailed specifications of functions through FRAM, enhancing the understanding of military activity and C² functions, and the interdependencies and relations between them,
- FRAM has the potential to describe and analyze functions involved in adversarial C², and enables the analyst to specify the constraints on own and adversary functions, in order to identify strengths and weaknesses in function performance on both sides,
• The FRAM methodology has been extended to allow for the description of military activity and its links to command and control functions,

• Data collected during an experimental simulation may be used to develop a functional model, and can be organized in a functional manner following the FRAM function description.

4.1.4 Summary and progression

The modeling effort presented here represented functions and constraints on performance in command and control, and the emergency management or military activity that command and control functions\(^1\) aim to direct. The modeling progressed throughout the three studies on command and control.

The utility restoration exercise study (Paper I) investigated command and control functions in a systematic manner, and highlighted constraints on these functions in terms of the properties of the artifacts that were used (in training compared to actual operations) to mediate function performance.

The C\(^3\)Fire study (Paper II) resulted in a method for modeling operational functions in emergency management (the operational level being the lowest level of activity in emergency management terms) and representing constraints on these functions in state space representations. Although command and control functions were not explicitly modeled, the model of operational functions in emergency management aimed to provide structure in the data collection function that is part of command and control. The representations of data and constraint, based on this model, aimed to aid the design of command and control systems to support sensemaking and planning. Links between GMTA and FRAM representations were exemplified.

The DKE study (Paper III) resulted in a functional modeling method that covers a broader range of tactical and operational, and command and control functions in a military context. The modeling effort attempted to establish a minimal model of command and control, sufficient for the explanation of what makes command and control and military activity agile, and fragile. The modeling efforts thereby aimed to contribute to the feasibility and usefulness of functional modeling in the analysis of command and control performance.

Consistency of the FRAM models within the studies has however been difficult to establish, has required considerable work, and may be regarded as results of the studies, as the original specification of the method was not detailed enough to unambiguously and consistently classify features (constraints) of functions to FRAM function aspects. Consistency of the FRAM

\(^1\)For this summary we will assume command and control to consist of three functions: data collection, sensemaking, and planning, following Brehmer (2007), see Section 2.3.5.
models between Papers II and III was not the focus of the modeling effort, so that the method in Paper III could progress with a more developed semantics for defining functions.

4.2 Aviation safety

The results of the three studies in aviation safety, presented in Papers IV to VI, are discussed in this section.

4.2.1 Norwegian Air Shuttle flight 541

The STEP method (see Section 2.4.4) was applied to the incident case description, as well as FRAM (see Section 2.1.5). A total of 19 functions essential to the description of the NAX541 incident were identified and grouped in accordance to the area of operation. Various descriptions of instantiations were created based on time intervals and the performance of functions, and the couplings between functions that became actual under these time periods. All of the steps 1 to 4 of FRAM were applied to the data of the incident report.

The results of the comparison of both methods was as follows. STEP is relatively simple to understand and provides a clear picture of the course of the events. However, STEP only asks the question of which events happened in the specific sequence of events under analysis. This means that events mapped in STEP are separated from descriptions of the normal functioning of socio-technical systems and their contexts. STEP only looks for failures and safety problems, and highlights sequence and interaction between events.

FRAM refrains from looking for errors and safety problems alone but tries to understand why the incident happened. Since FRAM addresses both variability of normal performance and the specifics of an adverse event, FRAM broadens data collection of the analysis compared to a STEP-driven analysis: Thus the development of the incident is contextualized in a normal socio-technical environment. Through asking questions based on the common performance conditions and linking functions in instantiations, FRAM identified additional factors and the context of why performance varied becomes apparent.

STEP provides a "mental motion picture" illustrating sequences of events and interactions between processes, indicating what happened when. FRAM instead sketches a "functional slide show", of which the instantiation presented in the paper is the first documented example in the history of the method, with its illustrations of functions, aspects, and emerging links between them in instances, indicating the what and when, and common performance conditions, variability, and functional resonance, indicating why and how. FRAM’s qualitative descriptions of variability provide more gra-
Another important finding is that it was possible to identify additional factors with FRAM. STEP interpretations and analysis depends on investigator experience, FRAM introduces questions for systemic factors and enables the explicit identification of other relevant aspects of the accident. The example also illustrates how unwanted variability propagates. However, several incidents in different contexts would need to be analysed to generalize these findings across cases.

4.2.2 Alaska Airlines flight 261

For the study in Paper V a simplified model of how systems’ behaviour developed according to the NTSB (2003) accident report was established, describing around 50 functions and around 100 couplings of functions and various instantiations of the model over time. These functions describe system behaviour as designed/prescribed and as actually performed, over a long time period (from the design of the MD-83-predecessor, the DC-9 in the 1960s, to the crash in 2000).

FRAM is shown in this paper to be able to address the behaviour of complex socio-technical systems in a systemic way through its flexibility and the use of the function as the unit of analysis. Functions describe technical, human, and organizational factors as integrated systems, focusing on their normal performance and their variability. FRAM is shown to provide the possibility to capture the dynamics of socio-technical systems over longer periods of time than previous snap-shot context-free assessments of performance to understand the full development of a complex accident. This is done by differentiating between potential couplings, associated with descriptions of functions as in how they should and/or may be performed, and actual couplings between functions, which arise when instantiations of functions are modelled to interrogate and manipulate states of the system over long periods of time.

FRAM focuses its analysis on the explicit description of function instantiations over time, enabling the description and modelling of day-to-day organizational aspects of work, such as policy, standards, work organization, and normal operations, as they develop over time. Descriptions of variability of functions aim to capture these aspects of normal work and their effects, as the analysis shows. The various pressures, such as regulatory, economic and personal pressures, may be illustrated as the relationships and couplings between functions become explicit in FRAM. Common performance conditions as sources of variability address these issues, as well as their effects may be traced through the functional description.

All four steps of FRAM are applied, and explained at a more detailed level than any other documented or published account of the method FRAM. As part of step 1, function descriptions are shown in detail a techni-
4.2. AVIATION SAFETY

cal and a human-machine oriented function, and organizational functions are discussed. The description of functions in terms of states, and the description of function performance as the interrogation of states, is an attempt at enabling matching between aspects in a consistent manner. As part of step 2, variability assessment using common performance conditions is demonstrated on human-machine oriented functions, and exemplified with data from the accident investigation report. As part of step 3, variability and its spreading to other functions is illustrated, for a number of functions. As part of step 4, the recommendations are related to the FRAM model, and the general description of how recommendations in terms of barrier systems and barrier functions may be linked to FRAM models is outlined as a new development of the method. Strengths and weaknesses of FRAM as the method is currently specified are presented.

The main strength of FRAM is that it manages to step away from linear cause-effect models of thinking about safety, which recently have been argued not to be sufficient to understand and explain all of the complexity of present-day socio-technical systems (see Section 2.4.2). It is thus one of the first such methods of its kind that steps away from linear descriptions of accidents in terms of structural system descriptions or events, and instead provides a functional systemic approach.

Another strength of FRAM is its ability to apply the description of a function to the description of performance of a complex socio-technical system. This means that although the method in itself is not complex, or even complicated, it has the ability to create complicated representations of complex systems and thereby meeting the complexity that other approaches may have oversimplified in the past.

FRAM is flexible in the sense that functions of a system may be described with different granularity. This allows the analyst to focus on the functional aspects of a socio-technical system that need to be addressed and describing them in more concreteness and detail, while maintaining more abstract coarse-grained descriptions of the functional parts of the system that are not of interest with respect to the modelling purpose.

Each of the steps in the method still poses FRAM analysts with a number of challenges. These challenges are identified here as a pointer to future research and continued work on the development of FRAM as a theoretical accident model and a concrete analysis method.

Step 1, the descriptions of functions, leaves with its current definitions some ambiguity to the function aspect definitions. Many aspects and states may be recurrent or overlapping as inputs, resources, preconditions, etc., making it a challenge for the analyst to establish a consistent description of functions. Because modelling in FRAM entails finding potential couplings between functions, more guidance also needs to be developed of how function aspects may be linked. As function aspects may overlap, the “network” of functions easily gets very complicated when a large number of functions is modeled.
Step 2, the variability assessment of functions would also need more specification. The common performance conditions were found adequate in describing the variability of the functions performed by human-machine systems. The list of common performance conditions applies to a lesser extent to strictly technical and organizational conditions, and a more general description of potential sources for variability therefore seems necessary, building upon the CPCs as they are currently defined.

Step 3, the identification of functional resonance for now serves more as a metaphor for the effects of damping and amplification and spreading of variability through a network of potentially coupled functions. The exact definition of how variability is amplified and under which conditions it may spread is subject of further research, because it demands a more explicit description of this step of the analysis than is currently specified.

Step 4, the identification of recommendations in terms of barriers and performance monitoring in order to establish variability management, is also underspecified. Given the accident report, the FRAM classification of barrier systems and barrier functions and their relation to the functional model could be made explicit. However, the method has yet to provide the analyst (or investigator) with guidance on which barrier systems and functions are applicable and appropriate given the potential for unwanted variability. Also, FRAMs potential facility of the distinction between unwanted variability which may lead to unwanted outcomes and necessary variability for outcomes that require adaptation and creativity, is an advantage of the model and the method. However, the method has yet to specify how these effects of variability may precisely be assessed, modelled, and predicted.

For all steps in the analysis it holds that the administration of functions, states, potential couplings, instantiations, and actual couplings is particularly laborious, partly because it is currently not adequately nor completely supported by administrative tools such as analysis software. A prototype analysis tool exists, named the FRAM Visualizer, but it currently does not adequately support the complicated models that step 2, 3, and 4 generate. As the method develops, matures, and becomes more explicit, a software tool needs to develop alongside the method to support the analysis in the administration of FRAM models.

4.2.3 ERASMUS ATM automation

Paper VI identifies a set of functions in en-route air traffic control and cruise flight from previously published task analysis efforts, observations in ACCs, and ACC position and cockpit simulators, and used as an input to the establishment of a FRAM model. The functions of the envisioned ERASMUS applications are included, based on project report deliverables, resulting in a model of the functions of the joint controllers-pilots-aircraft-ERASMUS system. The functions are evaluated by assessing common per-
formance conditions, variability phenotypes, and the potential for and effects of propagation of variability among functions.

Besides the CPCs that may be used to assess an entire function for potential variability, variability phenotypes may be applied to the function aspects. Step 2 of FRAM, the evaluation of variability for each function, is thereby extended with the use of variability phenotypes (described in Section 2.1.5 as failure modes). For the Monitoring function, for example, performed by the (pair of) air traffic controller(s), the variability phenotype timing of when aircraft trajectory data (input) is available varies. The accuracy of the specific data also vary, for example requested flight levels may be inaccurate, flight plans may be incorrectly filed or outdated, and radar may be inaccurate in mountainous areas. As an example of the output variability, the understanding of the position and timing of expected conflicts between aircraft varies. As an example of variability in the time aspect, the duration of the time available for monitoring is highly variable. Another example may be provided by the ERASMUS server function of Issuing CTOs, as it is mostly variable in its output in that CTOs vary in duration and distance (the time/distance during which aircraft fly under CTO varies), and the speed that the CTO entails for the aircraft varies.

Step 3 of FRAM is extended in that five ways of coupling and spreading of variability are illustrated. First, the consequences of variability may combine if the outputs of several functions are used by another function (as either input, precondition, resource, time, or control). For example, the aspect of (ground) speed of aircraft is an input to the Monitoring function, and is variable depending on wind direction and force, among other factors. At the same time, accepted CTOs entail speed changes for aircraft, which add to the variability of aircraft speed. From a FRAM analysis one can imagine that, from a controller (monitoring) perspective compared to an unaffected “normal” speed, a speed change by ERASMUS combined with a head or tail wind may lead to an increased variability of the Monitoring function.

Second, CPCs co-determine the variability of a function, hence, the function’s quality of the output. For example, the variability of pilot-accepted and implemented speed changes of ERASMUS CTOs may affect the input aircraft trajectory data (ATD) of the Monitoring function. This may however go unnoticed if the controller has little time available for monitoring, if many conflicting goals need to be attended to, or if the HMI prohibits effective assessment of aircraft speed.

Third, variability of the quality of a function’s output may be either detected and/or corrected, or used as is by another function. For example, from a controller perspective the variability in assessments of feasibility of the CTO by the pilot-aircraft system is assumed to be mostly undetected and simply used by the Monitoring function as described above. When the variability introduced by ERASMUS does get detected, it mostly does not need damping because it helps rather than disturbs conflict resolution.
Fourth, the variability in the quality of the output of one function may result in a change in the CPCs of another function. For example, the variability of Monitoring an especially complex traffic situation may influence the available time conditions for other functions such as Planning; an unexpectedly high number of clearances to be issued in a short period of time may influence the performance conditions for Coordination, etc.

Fifth, variability may arise due to inputs not being available, thus disabling couplings between functions. This could happen if ERASMUS works in a subliminal mode and pilots wish to negotiate with controllers about the various solutions to conflicts available to them, including the ERASMUS solution, an issue that has come up during experimentation in the project (Drogoul et al., 2009).

4.2.4 Summary and progression

Regarding constraint management, as in the previous studies on command and control, the four function aspects of preconditions, resources, time, and control form the constraints on function performance. In connection to FRAM, constraints within and between joint cognitive systems, defined by their functions may be established through functional modeling. Constraints may be elicited using FRAM for the description of work as it is designed or envisioned to be performed, through as-designed function descriptions and intended couplings between functions. Functions have also been modeled as they are performed during normal operations, representing how constraints are adhered to and/or negotiated in daily practice. Moreover, these constraints are the same constraints, in the FRAM way of thinking, that contribute to the success of functions, in achieving their goals.

The term constraint management adequately describes what happens with the constraints on functions as function performance changes over time, and accordingly constraints change over time. This is why the concept and philosophy of FRAM stresses constraint management, rather than constraint enforcement (see Sections 2.4.2 and 2.4.3), as variability is not only seen as something negative but also as the source for successful adaptation. Functional modeling, in the way it is done here, is essential for the understanding of constraint management because it enables the description of constraints on goal-directed behavior, as well as how performance influences constraints. In FRAM, the four function aspects other than input and output (resources, preconditions, time, and control) describe the constraints on performance. These aspects also highlight the simultaneous limiting and enabling character of constraints. A certain amount of a resource limits behavior to a set of possible actions, which at the same time are enabled or facilitated. The potential and actual couplings of the output of one function to these four function aspects describe how performance can or
does affect constraints. The potential and actual couplings between functions describe how performance potentially or actually shapes constraints.

The functional modeling efforts in aviation safety presented here have attempted to increase the consistency in FRAM as a method, which is best documented in Paper V. In step 1, the description of functions in terms of states, and the description of function performance as the interrogation of states, is an improvement of the method. The idea was coined by Hollnagel in 2008 but as yet had not been applied, and the current description of this model seems to be a step forward in consistency as now all aspects are states and all aspects can thus be linked to one another in a more straightforward manner. In previous studies the matching between function aspects was problematic and had to be forced through explicit specifications of variations of the same aspect to get a match between one function’s output and another aspect of another function. This development builds on the concept of instantiations, coined by Hollnagel in 2008 and first applied in Paper IV, which exemplifies the difference between potential couplings and actual couplings at specific points in time.

The articulation of the problem with the assessment of variability in step 2 is articulated in Paper V, which describes which aspects in the variability of technologically, human-machine, and organizationally oriented functions are (not) captured by the common performance conditions. The specification of variability in Paper VI with variability phenotypes (formerly known as failure modes) also contributes to the degree of elaboration of this step of FRAM.

The combination of common performance conditions and variability spreading through couplings between functions has also become clearer, mainly in Paper V but also in Papers IV and VI, as this aspect of step 3 of the method was unclear initially. The specification in more detail of step 3, the functional resonance, has aimed to contribute to the theoretical articulateness and plausibility of FRAM, as the functional resonance metaphor has been related to empirical data.

Regarding step 4, the usefulness of FRAM in generating countermeasures could not be developed in detail, since this step too is as yet too underspecified, but Papers IV and V address the issue of how to model countermeasures identified in investigation reports, and Paper V does so in considerable detail, thereby contributing to the FRAM method.
Chapter 5

Discussion

This chapter discusses the conclusions that may be drawn from this research, the contributions of the thesis, and possibilities for continuation of this research.

5.1 Conclusions

Three hypotheses were postulated in the Introduction of this thesis, which will here serve as the frame of reference for the discussion of conclusions.

1. The performance of complex socio-technical systems is shaped by constraints, and the actions of complex socio-technical systems shape constraints in order to manage constraints.

2. Constraint management provides a basis for the analysis and improvement of (a) the safety of socio-technical systems (specifically, in aviation), (b) joint system performance and the design of support systems (specifically, in command and control), as well as (c) the training of teams in complex socio-technical systems (specifically, in emergency management).

3. Functional modeling of constraint management provides an adequate means to address constraint management in complex socio-technical systems for the purposes and domains outlined in (2).

An extensive literature study has shown that theories and studies related to cognitive systems engineering emphasize how constraints shape actions, that less attention has been directed to how actions shape constraints, and that the discussion of a combination of these perspectives, the reciprocal relationship between constraints and actions, is rare. The studies presented here have aimed to address this relationship explicitly.
The study of the electricity network resource management exercise has shown that the performance of functions of socio-technical systems is shaped by constraints, as the performance of the electricity network resource management team was affected by the constraints in their environment, consisting of the information and communication technology available in their collocated setting. The study argues that the team’s function performance would have been different in an actual setting with radically different constraints on communication and cooperation. Actions of the team shaped constraints for example as they chose the deliberate organization of their work and their work environment, and performed the functions that determine how cooperation and collaboration should be performed. By performing team coordination functions, they shaped the constraints for future actions. The study thus provides a functional account of constraint management.

The studies of the microworlds have shown that the environment of a system constrains that system, and that some constraints stem from the internal functioning of the system, and exemplify what this means for emergency management and military command and control. The focus was on joint cognitive systems, meaning human-machine systems that are defined by their functions and thus have relative boundaries. The C^3Fire study documents a viable approach to generating representations of constraints, and hypothesizes and argues that these displays may aid in the constraint management of joint cognitive systems. The DKE study documents a viable approach to generating function representations, and hypothesizes and argues that these representations may aid in the constraint management of joint cognitive systems. The distinction in the DKE study between the tactical and operational levels of activity, as well as their relation to command and control functions, have shown the relevance and feasibility of modeling constraints and their management at different levels of granularity.

The three aviation safety studies have provided examples of joint cognitive systems acting under constraints to keep them within safe performance boundaries, and that they have to adhere to certain constraints in order for them to perform safely and efficiently. Especially the Alaska 261 study has demonstrated however that the perspective of constraints on action does not explain complex socio-technical system behavior satisfactorily, as the case has numerous examples of how constraints are negotiated and adjusted over time. The case is also a clear example of the fact that in actual practise, constraints need to be constantly adapted, because of competing goals of a system in relation with its environment (e.g., production vs. safety), or under- or overspecification (e.g., unclear, missing, or too strict procedures or oversight) of constraints on actions. The perspective of actions on constraints, is thus needed in complement to the perspective of constraints on actions, to understand actual system performance.

The study of ERASMUS risk assessment also exemplifies the need for the perspective of a reciprocal relationship between actions and constraints,
as the evaluation of system use suggests that constraints that are enforced by automation (ERASMUS) and by people (air traffic controllers) may not be unifiable with each other, nor with the constraints of the process that is to be controlled (the aircraft flying). Actions that shape constraints, allowing for adaptation in the face of competing goals, are thus a necessary perspective in the analysis of risk. This perspective may be summarized under the term constraint management.

The six studies presented here have thus illustrated the need for addressing constraint management, and the reciprocal relationship between constraints on joint cognitive systems’ actions. In order to fully address this relationship, it is necessary to describe performance of joint cognitive systems and constraints in an intertwined manner. This has been accomplished here by addressing functions and constraints in models of performance of socio-technical systems, using a model presenting a taxonomy of functions performed by teams, and, chiefly, a model and associated method called FRAM. These functional models provide insight in the management of constraints, which leads to the third hypothesis that this thesis has addressed:

Combining constraints with the descriptions of functions, the four function aspects of preconditions, resources, time, and control form the constraints on function performance in FRAM. Actions are the result of the actual performance of functions, transforming the inputs of a function into its outputs. Constraints on actions are thus described in terms of function aspects. Actions on constraints are described in terms of couplings between the output of a function and the preconditions, resources, time, and control aspects of other functions. FRAM distinguishes between potential couplings and actual couplings. The management of constraints in FRAM is thus manifested in the performance of functions, producing outputs given constraints, so that actual couplings emerge linking outputs to constraints of other functions, given potential couplings between functions. Five of the studies have shown that for the purposes of understanding team training, supporting command and control, and retrospective and prospective analysis of safety in system performance, these constraints on functions form useful units of analysis with respect to function performance. As such, the studies show that functional modeling of constraint management is suitable for the description of both desirable and undesirable outcomes, in harmony with the premises of resilience engineering.

5.2 Contributions

The thesis has generated the following contributions:

- The evaluation of team performance using a functional team performance taxonomy has been applied to an emergency management team in a functional exercise, and has been shown to be useful in the
description of constraints of information and communication technology and the likely impact of these constraints in situations of training as well as actual practise. The thesis has thus contributed to the conceptual tools available to designers and evaluators of emergency management training environments.

- The perspective of constraint management to emergency management and command and control is a contribution in the understanding of these activities, providing guidance in the design of support systems and representations of tasks in order to enhance control and agility in command and control. The significant difference with earlier approaches is that it explicitly combines the perspectives of how constraints shape actions, and how actions shape constraints. The representations (support displays and conceptual models) aim to highlight constraints in order to facilitate the perception of constraints, so that actions may be chosen to meet or adapt constraints.

- The perspective of constraint management contributes to aviation safety and the science of safety in general by providing an integrative perspective combining safety by constraint, safety by design, and safety by management, in one central concept.

- The method to generate state spaces that represent constraints and performance in combination has been developed to the extent that it for microworld settings can generate functional models and constraint representations based on these models. The state space itself, although not new, is a contribution to the tools of analysts of network-based command and control, and logistics problems in general, providing a concept for process measures rather than outcome measures for the evaluation of performance. The method of generation of the state space may be regarded a contribution of this research.

- The relationships between constraint, constraint management, and functional representations, have been interpreted in a new way through the functional resonance analysis thinking of FRAM. It thus provides an alternative to the modeling of constraints and functions that overcomes certain limitations of established modeling methods.

- The functional resonance accident model and associated analysis method (FRAM) has been applied in various ways, evaluated, and developed further, regarding all of the steps in the method, during the course of this research, be it in slightly different directions in the different studies. This thesis contributes to exploration of the possible interpretations of FRAM, and offers some insight towards making it more well-defined and consistent at least within the domains of application, if not across.
5.3 Continuations

Research may be continued from the work presented here in at least four related directions:

- The analysis of both functions and constraints in command and control tasks has been developed into considerable detail and may be used for field study of natural command and control environments. Combination of the analysis of functions and constraints for example with tools for the analysis of actual exercises and operations of command and control (see, e.g., Morin et al., 2000) as well as with software modules of support systems for command and control (see, e.g., Kolfschoten et al., 2006) seems a promising way to go. Also for more realistic evaluation of its potential it would be necessary to take the analysis method to face the messy details of actual command and control work, such as the research setting studied in Paper I.

- The functional resonance accident model is appealing in its potential for addressing complex socio-technical systems in new ways, in attempts to make ultra-safe systems even safer, because of its fundamentally different causal assumptions. It is however in its early stages of development and would require application, conceptually and with the associated method, to cases where the description of variability and the exemplification of negative as well as positive variability can be supported with empirical data. Although this thesis has attempted to contribute in this direction, theoretical articulate-ness would need to be extended in this respect.

- The functional resonance analysis method, thus also in its early development, needs to be developed further to supply guidance to analysts in order to generate consistent results and be generally more understandable for a wider audience, as each of the steps in the method is currently underspecified to some extent. It has not been the purpose of this thesis to write detailed guidance on the application of the method, but such detailed specification with many more examples than have been shown here in a handbook would be desirable. Application and evaluation by collaborating analysts and developers in a wide range of complex domains could further test and develop the method, as more empirical support becomes available. In a similar manner as in this thesis, iterative description and testing of the method seems the way to develop a general description of such a method. When specified in more detail, the range of applicability of the method seems very large, as functional analysis is particularly applicable to human-machine systems around us, from the everyday and (seemingly) trivial to the complex and intractable.
Once more consistent results can be expected from the functional resonance analysis method, predictive accuracy, usefulness in generating countermeasures, resource efficiency, usability, and auditability may be evaluated and increased. Software tools are likely to be a useful support, when developed in parallel to method development. This approach seems to be the way to get closer to the ultimate goal of methods for safety, risk, and resilience assessment: the application in industry.
Bibliography


Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge: Command... control... in the information age*. CCRP.


Grant, T., & Kooter, B. (2005). Comparing OODA & other models as operational view C2 architecture. In *Proceedings of the 10th international command and control research and technology symposium (ICCRTS).* CCRP.


BIBLIOGRAPHY


BIBLIOGRAPHY


Sun Tzu. (n.d.). *The art of war*.


BIBLIOGRAPHY


BIBLIOGRAPHY


Index

accident models, 58
adaptation, 22
air traffic control, 8
ALARP, 56
area control center, 8
behavior, goal-directed, 20
behavior, purposeful, 20
C³Fire, 79
CAEC loop, 48
case study, 74
Challenger, 36, 64, 65
clearance, 9
cognition, 21
cognitive systems engineering, 21
cognitive work analysis, 27
Columbia, 64
complex interactions, 20
complexity, 20, 78
constraint, 33
contextual control model, 47
control, 20
controlled time over, 10
correspondence, 74
coupling, 20
cybernetics, 20
decision making, 44
DKE, 81
DOODA loop, 50
drift toward failure, 62
dynamic decision making, 46
dynamics, 78
ecological approach, 18
ecological interface, 28
emergency management, 40
environment, of system, 18
ERASMUS, 8, 86
extended control model, 48
failure modes, 32
feedback, 20
fire-fighting, 79
FRAM, 29, 65
function, 21–23
functional account, 22
functional exercise, 75
functional explanation, 22
GMTA, 25
GPS, 26
human error, 59
IDEFO, 26
intransparency, 78
joint cognitive system, 21
line of behavior, 18
maintenance error, 59
manageability, 20, 78
method, 25
method criteria, 25
MFM, 25
microworld, 78
migration toward boundaries, 62
minimal modeling manifesto, 23
model, 23
naturalistic decision making, 46
normative decision theory, 52
Occam’s razor, 23
OODA loop, 50

phase plane, 37
phase space, 37
plan recognition, 42
polytely, 78
prediction, 20
process, 20
process control, 25

rationality, 61
recognition-primed decision making, 46
recursion, 26
representation, 23, 37
representation design, 37
representative design, 74
representativeness, 74
resilience engineering, 67

SADT, 26
safety, 55
safety by constraint, 66
SESAR, 87
simulated task environments, 74
STAMP, 64
state, 18
state space, 37
system, 17
system, behavior of, 18
system, dynamic, 18
system, open and closed, 18

table-top analysis, 73
taskwork, 43
team, 43
team performance, 43
teamwork, 43
TOTE, 26
tractability, 20, 78

variable, 17

WYLFIWYF, 83