PART IV: CONCLUSIONS AND FUTURE WORK

In this thesis methods for agent-based analysis and support of human functioning in complex socio-technical systems have been addressed. The methods are in accordance with a three-dimensional framework for classification of agent-based models, and enable the modeler to develop models at different abstraction levels, and to establish interlevel relations between those models. Models of human functioning were created and analyzed in safety and healthcare domains with the aim to provide support and recommendations both at the level of the whole system and at an individual level. More specifically, Part II of the thesis is dedicated to the analysis of multiple components of a safety critical socio-technical system and Part III is focused on the analysis and support of humans in socio-technical systems related to the healthcare domain.
The main outcomes of the work are summarized below:

- Eight executable agent-based domain models were developed in order to understand the behavior of complex systems illustrating the functioning of interlevel relations: four models for safety-critical domains and four models for healthcare
- Three support models were developed: one for the safety domain and two for healthcare. One of the healthcare support models resulted in a real-life application - the eMate system based on the COMBI agent-based model of human behavior change
- Three qualitative models according to other informal modelling approaches in the safety domain were developed and contrasted to an agent-based model of the same case study
- Two integrated computational models were created. One is composed of three computational (sub)models of human functioning and another of two (sub)models. One integrated model was applied for safety analysis of a real life incident
- The mobile ambient support system eMate for therapy adherence and behavior change was developed based on the computational domain model of behavior change

The main contribution of the research described in the present thesis is that more insight was obtained in complex systems’ behavior and behavior and roles of humans in joint cognitive systems and human-technology interactions. This insight can help to build more resilient and user-friendly socio-technical systems. The main uitcomes of the current research and its concrete contributions are discussed in the sections below.

Discussion of Results

In this section the results of the current research will be presented in the light of the research questions described in Part I of the thesis. The main goal of the current research is to analyze global emergent behavior of a whole system based on the local behavior of the system's constituents with the focus on human functioning using agent-based modelling techniques and to provide an adequate support and advice based on this analysis, either direct at the level of an individual or indirect in the form of general recommendations concerning system's design. An overview of the contributions of each chapter of the thesis related to the above listed research sub-questions is given in Table 1.

Table 1. Answered research questions per chapter.

Research questions	Research subquestions	Chapters
I. How can agent-based modelling contribute to the analysis of complex socio-technical systems?	a) How can a global behavior of a socio-technical system be explained by the behavior and interactions of its local components (top-down system analysis)?	2, 4, 10
	b) How does a change in the behavior of a local component result in global behavior of the system (bottom-up system analysis)?	1, 2, 4, 5, 7, 8, 9, 10
	c) How can different modelling approaches and techniques to study socio-technical systems be compared and contrasted with respect to analysis possibilities and complementarity?	3, 6
II. How can agent-based analysis of complex socio-technical systems help in providing an adequate support for humans in abnormal contexts?	d) How can computational modelling and analysis at different levels contribute to the development of ambient support systems in eHealth and safety critical domains?	1, 7, 8, 9, 10
	e) How can computational modelling and analysis at different levels contribute to the development of useful recommendations in eHealth and safety critical domains?	2, 3, 4, 5, 6, 9

Further, the detailed results will be presented according to the initial research questions. The main research objective can be subdivided in the following two questions:

Research question I

This question addresses the contribution of agent-based modelling techniques to the analysis of complex socio-technical systems, such as the air traffic management system, or the e-Health system and is formulated as follows:

How can agent-based modelling contribute to the analysis of human functioning in complex socio-technical systems?

This question is split into three sub-questions that are addressed below.

a) *How can a global behavior of a socio-technical system be explained by the behavior and interactions of its local components (top-down system analysis)?*

In Chapter 2 it is shown how an agent-based approach can be applied to model and to analyse a real life incident in aviation. The incident formed a case-study for the agent-based analysis. The incident description was obtained during an interview with a retired pilot of a European civil aviation company. It describes a case when due to a misinterpretation of a command of air traffic controllers, pilots of one aircraft initiated a take-off procedure without take-off clearance (permission). At the same time another aircraft received take-off clearance from the tower air traffic controllers and initiated take-off on the crossing runway. As a result two aircraft were taking off simultaneously on the intersecting runways. Fortunately the tower controllers recognized the conflict on time, gave a stop command to the pilots of one aircraft, which was still able to abort the take-off procedure.
The given case-study was formalised both in the form of a trace and in the form of a dynamic agent-based model. The trace reconstructs the events that took place during the incident while the dynamic model allows to simulate both the real scenario and hypothetical scenario's that might have happened with different parameter configurations of the related agents. Development of these two agent-based formal models at the local level with respect to time, process and agent clustering dimensions serves as a means for backward (top-down) analysis of the incident. In other words, it is retrospective incident analysis when an occurrence of an incident is a given historical fact that represents global behavior of a system. The incident can be regarded as an emergent undesired property of the system since the whole air traffic management system was not designed with a purpose to exhibit this type of behavior. 'Non-incident' is also a type of system's behavior, though this type of behavior is a desired property of the system that was developed with the purpose to provide control

of air traffic in a safe manner. It is demonstrated in this chapter how emergent, both desired and undesired properties of the system can be explained at different levels of aggregation. By means of the formal definition of lower level properties of the system that are related to the high-level global properties and automatic checking of these properties upon local traces, more insight was obtained into emergent properties. For instance, in the context of the case-study described in Chapter 2, two types of system behaviors at a lower level of process abstraction and time dimensions can result in the emergent behavior 'no-collision' of the system: either the intermediate property 'absence of simultaneous take-offs on crossing runways' holds or the property 'correction of simultaneous take-offs on crossing runways' holds. In both cases it is obvious that no collision incident would happen if the airport did not have such a characteristics as 'crossing runways'. Investigating the causal interlevel relations between properties from more global to local levels across a process abstraction dimension brought us to the cause of the problem: a wrong action was performed by the pilots due to a specific mental model of the aircraft pilot(s) and belief about take-off clearance that was based on this mental model.

In Chapter 4 an existing dynamic risk agent-based model of a runway incursion incident [14] was analyzed by means of identification of conflict recognition and conflict resolution events and recoding them in Monte Carlo simulations of the dynamic risk model to obtain more insight in the behavior of agents during the simulations and their role in conflict-related events in different visibility conditions. The agent-based collision risk assessment model under analysis represents a future taxiing into position and hold (TIPH) airport operation. This operation aims at placing an aircraft on the runway soon after the landed aircraft, ready for immediate take-off.
Relevant events for the model's analysis were selected according to their possible connection to emerging conflicts. The selected events were related both to observable behavioral events, such as aircraft maneuvers in air during the final approach and taxing aircraft maneuvers on taxiways activation of runway incursion alert of the ATC technical system, and cognitive non-observable events at a lower process abstraction level, such as pilots' observations, decisions and intentions. Here, 'aircraft collision' and 'no-collision' emergent properties of the modeled system are contingent upon such system's properties as visibility conditions in the airport, ability of pilots of both aircraft to detect a conflict by own observations, performing successful taxiing off the runway by the pilots of the taxiing aircraft, initiation and execution of a missed approach as a result of a conflict detection by the pilots of the landing aircraft. It became clear from this analysis that the 'collision' property may emerge as a result of too late conflict detection since the initiated actions of agents in order to resolve the conflict do not help to prevent a collision. However, in some cases a collision is prevented if earlier conflict detection by human agents occurs. It was not trivial to trace back the causes of system's global behavior in this model due to stochastic relations between numerous components of the model.

In Chapter 10 an agent-based ambient support system for therapy adherence and health-related behavior change is presented. The system is based on the domain model of a human agent and on the analysis and support models integrated in an ambient agent according to the integrative modelling approach presented in [1]. The core of the system is a multi-level human agent model of cognitive processes, from local to more global cognitive levels, that result in a certain health-related behavior. Here the system's global property corresponds to the activation of health behavior while intermediate and local cognitive properties are represented by an activation of multiple cognitive modules.

The model was designed in such a way that it can allow for automatic top-down reasoning about the causes of undesired behavior exhibited by a human agent. The reasoning about these causes, updating the values of cognitive modules and an execution of appropriate interventions based on model-based backward (top-down) reasoning are performed by an ambient agent. The model consists of three layers of cognitive modules and one behavioral layer that are interconnected with each other by means of deterministic causal relations. The values of each layer are calculated from the values of the related lower-level layer. For instance, the value of the individual's attitude to healthy diet, exercising and correct medication intake in the intermediate layer is calculated taking into consideration the values of beliefs about the advantages of healthy lifestyle and medication intake, emotional connotations associated with healthy lifestyle and medication intake and social norms about the relevant type of health-related behavior. Reasoning is executed by an ambient agent using model-based diagnosis. The reasoning algorithm about the causes of unhealthy behavior first investigates the most proximate causes of unhealthy behavior in the nearest intermediate layer of the search tree, if the values of the related variables in this layer are below a threshold, then the search continues further up to the lowest-level behavior determinants. Model-based top-down analysis is effective in this case due to the causal relations between the elements of the model.

b) *How does a change in the behavior of a local component result in global behavior of the system (bottom-up system analysis)?*

In Chapter 1 model-based bottom-up reasoning is performed by an ambient agent about the influence of some low-level aspects of information presentation in a display on operator's performance. In this chapter a human performance domain model that focuses on the way information is presented to an operator in a display is integrated with an earlier developed operator's functional state model [1]. Within the analysis component of the ambient agent, by means of reasoning forward in time based on the domain model, predictions are made about future states of the human agent and the environment. The information presentation effects domain model of a human agent has a hierarchical structure where the functioning fit of an operator can be regarded as a global property of the human information processing cognitive system. The functioning fit property represents an optimal balance between processing demands and alertness resources. This global property is dependent on lower level cognitive

intermediate properties and each subsequent level of elements is dependent in its turn on the values of elements that are situated in the lower levels in the hierarchy at a cognitive level. The relations between the system's elements are causal. Forward analysis about the influence of information presentation aspects, such as changing the background color of the display or luminance, grouping of objects in the display, demonstrates that functioning fit of an operator is determined by combinations of multiple aspects and not purely by separate low-level elements. In some cases increasing the value of one low-level element in the cognitive hierarchy, for instance luminance, may not influence performance value in a negative way, in other cases it influences operator's performance in a positive way dependent on the values of other factors, such as experienced pressure and operation's time criticality. This bottom-up model-based analysis demonstrates that it is difficult to predict the global system's behavior based on the values of its separate low-level elements.
The same conclusion can be derived after analyzing the behavior of the multi-agent model of a runway incursion scenario described in Chapter 2. If one of the agents exhibits non-desired behavior (initiation of a take-off operation without an issued clearance in this context), it may still result in the desired global system behavior 'no collision' if the other agents are able to detect this deviated behavior and to perform effective corrective actions.

In Chapter 4 additional analysis of the agent-based model was performed by taking several agents out of the monitoring or control loop at a local level of agents clustering and process abstraction dimensions and the behavior of the whole system at a global level was observed based on these local changes. The results of this experiment demonstrated that it is difficult to predict global emergent behavior of the system based on the changes in its local elements, especially if the model is based on stochastic relations between its elements.
Chapter 5 describes a complex agent-based model of operator's performance that integrates three cognitive agent-based models: operator's functional state model, situation awareness model and decision making model. The joint model was applied to a real life aviation incident reported in the NASA Callback newsletter [10]. Simulation experiments with the integrated model demonstrated how observable behavior of human operators could emerge from changes in their task and stress levels. The situation awareness submodel incorporates a stochastic element at the point of translation of operator's observations into beliefs. Due to this element the prediction of system's global behavior, such as unsafe actions of pilots, is not trivial for any particular simulation, though the estimation of relative frequency of erroneous actions is possible.

In Chapter 7 an intelligent support system for diabetic type 1 patients uses model-based bottom-up analysis in order to predict blood sugar level for a diabetic patient and to provide an automatic support to the patient with respect to his insulin intake. The analysis component of the ambient agent system in this work uses a domain model of blood sugar level. The domain model utilized in this work is at a physiological level:, relations between the variables of the domain model are expressed with the help of differential equations. In reality it is not so easy to predict blood glucose and insulin level for each individual since many other factors play a role except of physiological parameters. It is quite easy to predict the emergent property of the system under consideration, the glucose level, if all initial parameters and values are known and are correct, the challenge in these types of systems is that not all initial values and parameters of the system at each time point are known. If food is not taken according to the pattern that is known for the system, the system will make wrong predictions. Variability and dynamics of contributing factors complicates the forward analysis based on this model. In this case complexity is confined in uncertainty rather than in the number of system's elements and their interrelations.

In Chapter 8 a habit formation model is described where an interplay between goals, feelings, intentions within one human agent along with external cues determines individual's behavior and formation or no formation of a habit. The model was inspired by elements from neurological literature including the concept of Hebbian learning. In this context the main predictors of global behavior of the whole system at a low process abstraction level are the strength of counteracting goals and the strength of the existing habit. However, individual parameters such as learning speed and internal bias also play an important role. Whether habitual behavior will be formed is determined by the frequency of simultaneous goals and cues activations, as well as the strengths of the internal connections; so here not only the systems' elements themselves play a role, but also their frequent activations. Relations between the elements of this model are not linear and are shaped by threshold functions with unique parameters for each aspect. Prediction of the model's behavior based on the values of its elements and their relations are not straightforward at all since all connection values should be taken into consideration and their effects.

In Chapter 9 habit contagion in agent-based networks has been designed and analyzed both at an individual and at the agent clustering level. Each agent in these networks possesses his own internal dynamics of cognitive processes that involves attitudes, goals, feelings and intentions with respect to certain behavior, and agents are connected to each other in different ways. Model simulations at a group level demonstrated that a change in some local components of the agents' cognitive architecture in a small fraction of the population might drastically change the global system's behavior in small world networks. Although most nodes in a small world network have a low number of connections, these types of networks also contain a substantial number of hubs – nodes with a high number of connections. For instance, if hubs in a small world network with multiple direct connections adopt particular attitudes towards healthy behavior, this attitude may spread very fast to other individuals and a larger portion of the network would adopt this attitude and the corresponding behavior. If we consider another type of network, changing attitudes and behaviors of some individuals in the network may not play a great role at the global level. This study demonstrates that emergent behavior of a system depends drastically upon the system's structure and interconnections between its elements because networks with the same initial values, but different connections configurations may develop different behavior patterns.

c) *How can different modelling approaches and techniques to study socio-technical systems be compared and contrasted with respect to analysis possibilities and complementarity?*

In Chapter 3 the agent-based modelling approach LEADSTO [5] is compared to three other popular approaches in safety critical domains: Event Trees, STAMP [9] and FRAM [8]. This comparison was performed by means of an application of each of the approaches to one case study of a runway incursion incident described in Chapter 2. Traditionally, air traffic management safety analyses are performed based on the Event Trees approach. Event trees are based on graphical representations of Boolean logic relations between success and failure types of events. It is a bottom-up approach that starts with an initiating event and ends with its consequences. Event Trees can be quantified by associating with each branch a conditional probability, given the successes/failures associated with all branches leading up to it. The approach is still widely used, although there is an increasing awareness that it has some limitations, especially when it comes to analysing dynamic systems with time-dependent interactions.
The Systems-Theoretic Accident Model and Processes (*STAMP*) methodology [9] uses system and control theory to describe socio-technical organisations. In this methodology, an accident is not understood in terms of a series of events, but rather as the result of a lack of control or the constraints imposed on the system design and operations. As part of additional analysis, STAMP uses system dynamics to describe interactions and dynamics between organisational processes and their effect on safety. The variables in these types of models are typically at an aggregated organisational level, rather than at the level of individuals in the organisation.
The Functional Resonance Analysis Method (*FRAM)* [8] provides a framework for systemically describing and evaluating functions and performance variability within ATM systems. It is more a task-oriented than an agent-oriented approach. FRAM characterises socio-technical systems by the functions they perform rather than by how they are structured. For each function that is identified, six aspects are described, namely input, output, resources, control, precondition, and time. Time dynamics is captured by modelling non-linear dependencies and performance variability of system functions. Based on this, the modeller can find combinations of variability of the functions that may lead to 'functional resonance', i.e. situations where the system loses its capability to safely manage variability.
The agent-based approach *LEADSTO* is a formal language and software environment for modelling and simulation of dynamic processes in terms of both qualitative and quantitative concepts. En extensive description of this modelling language is given in

the Introduction of the present thesis. In this context one should note that the LEADSTO approach is not synonymous with the various other agent-based approaches that exist. However, LEADSTO shares the standard assumptions of the agent paradigm, such as the idea to conceptualize the ATM system as a multitude of autonomous entities, and to analyze the system's overall dynamics as emerging from the individual agent processes and their interactions.
Based on the literature and expert opinions, sixteen criteria for comparison were identified. As far as the practical usage of the models concerns, it appeared that an agent-based approach provided much insight into time dynamics at a local level that was crucial in understanding the emergence of the given runway incursion incident. Moreover, by changing of parameters of the model, one can simulate and analyze hypothetical scenarios when the behavior of local agents would differ from their real behavior according to the scenario. Such simulations contribute to better understanding of the processes and their interrelations that contributed to the occurrence of the incident. STAMP analysis appeared to be very useful in providing insights into the whole ATM system behavior from both a global agents clustering perspective, including managerial and organizational factors, and a local perspective, including safety requirements and constraints imposed to each actor in an organization. Event trees was the most user-friendly and less time consuming modelling approach, though it is obviously too simplistic with respect to its expressive power. It has limitations regarding modelling complex and non-linear processes. However, identification of events allowed for a clear and structured event picture that served a good starting point for developing more sophisticated models according to the other approaches. FRAM might be useful in the analysis of system's functions variability, though its application to this case study did not provide any substantial insights into the emergence of the incident.
An important conclusion of this study is that it is not possible to perform a comprehensive comparison of different modelling approaches based on only one case study. More cases are needed in order to explore modelling potential of different approaches in safety analysis. Nevertheless, the main capabilities and features of the approaches have become clear after their application to this case study.

In Chapter 6 a more narrow comparison was made between two formal computational methods: a general agent-based approach was compared to a population-based agents clustering approach that is frequently used for modelling epidemics in a population. The two approaches were applied to a scenario of Situation Awareness spread in a group of eight agents. The comparison was made with respect to the predictions that both models can make concerning the dynamics of situation awareness and its equilibria states. Simulation results of two types of models, agent-based and population-based, were compared and analyzed. This study demonstrated that the general tendency of spread of Situation Awareness in a relatively small population could be approximated by averaging agent-based simulations, though any individual agent-based simulation may provide results that differ much from the results obtained in a population-based model. Agent-based modelling appeared to be useful for getting more insight into global behavior of a system by zooming into the pattern at an individual agents' level. This comparison demonstrated that it is difficult to predict system's typical global behavior pattern based on one individual simulation due to the dynamic connections between the agents.

Research Question II

The second general research question is formulated as follows:

How can agent-based analysis of complex socio-technical systems help in providing an adequate support for humans in abnormal contexts?

This question is split into two sub-questions that are addressed below.

d) How can computational modelling and analysis at different levels contribute to the development of ambient support systems in healthcare and safety domains?

Analysis of complex systems can provide useful support and recommendations in air traffic and healthcare domains. For instance, an analysis component is a crucial part of an ambient support system that may support humans in their everyday work or health related activities.

In Chapter 1 an information presentation support system for an operator in any safety critical domain is developed based on model-based backward and forward reasoning. An ambient agent is able to observe and analyze the role of an operator's functional state and information presentation effects in a display on an operators' performance fit. Based on this analysis, appropriate support is provided by means of adjusting different aspects of a display, such as background color, luminance or objects grouping. This support can be extended by providing direct feedback on the stress level of an operator or by means of allocating part of the work to other human agents that experience less workload at the moment.

In Chapter 7 an autonomous intelligent support system for diabetic patients was developed based on forward reasoning-based analysis of a domain model of blood glucose level performed by an ambient agent. The system gives advice to a patient on the time and amount of insulin that should be taken depending on daily activities of a patient and his food intake. The ambient agent can give advice on insulin intake based on the input from an electronic agenda of a patient and an electronic blood meter and predictions that are made during the simulations of domain model of insulin intake for this particular patient. The simulation results demonstrated that in the scenario with the patient's physical activities the intelligent support system helps better to

maintain the appropriate blood glucose level in comparison to the regular insulin prescription. The system has a potential to release a bit of the burden of diabetic patients as it can predict the effect of the upcoming activities more precisely than humans.
The computational model of habit learning described in Chapter 8 can enable ambient support for lifestyle change. Though the ambient support component has not been developed yet for this system, the domain model of habit learning provides a good foundation for developing a support system for lifestyle change. For instance, an ambient agent that performs formal analysis of the model is able to identify the internal cognitive states and environmental cues that shape person's habits and based on this analysis, it is able to provide adequate support and coaching to the person with respect to his goals, intentions and his responses to external cues. The challenge for the artificial ambient agent is to derive a person's internal states and his behavior in order to provide adequate support. These observations can be possible with the embedded sensors or direct text interactions with a human.

In Chapter 9 contagion of habitual behavior in social networks and its mechanisms is modeled and simulated. The introduced model may also be embedded in an ambient support agent that may provide coaching with respect of social influences on adopting a healthy lifestyle. If such a system is aware of a social network of the person and his goals and intentions, it can perform model-based forward reasoning and make predictions of social influences and cues on habitual behavior, it may try to adapt the environment in such a way that the habit will be changed, for example by drawing one's attention to the long term goals, or by suggesting changes in the (virtual) social network around persons in order to stimulate contagion of states that will have a positive influence on the habit.

It is demonstrated in Chapter 10 how a mobile ambient support system for therapy adherence and health-related behavior change can be designed and implemented. The ambient support system eMate described in this Chapter uses model-based diagnostic reasoning in order to derive the possible bottlenecks of healthy behavior. The system was designed to support chronic patients with type 2 diabetes, HIV and cardio-vascular disease, yet it can be applied to many health and lifestyle domains. Model-based analysis of human cognitive health-related modules forms the core of the support system and provides a good foundation for tailored interventions. Subdivision of cognitive modules of behavior determinants into hierarchical layers allows for efficient reasoning and interventions when the most proximate cause of non-adherence at a more global level is targeted and not all lower level behavior determinants.

e) *How can computational modelling and analysis at different levels contribute to the development of recommendations in healthcare and safety critical domains?*

In Chapter 2 the agent-based analysis of an air traffic management system at local and global temporal and process abstraction levels provides much insight into the behavior of this system. In Chapter 3, the comparison of different accident analysis approaches, including the agent-based approach described in Chapter 2, demonstrated an advantage of the agent-based approach in the sense that it provides more insight into the temporal dynamics of the runway incursion incident.
The main recommendations that can evolve from this extensive analysis of one case study are the following:

- careful design of airport runways configuration;
- creating new procedures that allows for safer time ranges between take-offs on intersecting runways depending on weights of aircrafts, amount of traffic, weather conditions etc.;
- development of new procedures that prescribe the variability of air traffic operations for the same aircraft to prevent learning effect and increasing of awareness about learning effects on performance;
- promotion of better safety culture within aviation organizations.

Chapter 4 extends the analysis of an agent-based dynamic risk model of a future TIPH (Taxi into Position and Hold) airport operation [14]. The analysis addresses accident risk related events and roles of agents in conflict resolution and recognition during the runway incursion scenario. The purpose of this additional analysis was to get more insight into the model's behavior over millions of Monte Carlo simulations. The analysis consisted of two steps: 1) a conditional risk for some relevant events was calculated; 2) several selected agents were taken out of monitoring or control loop and the consequences of agents' degraded performance were analyzed in the light of its effect on a collision risk. The findings of this computational analysis give rise to some recommendations regarding this operation. For instance, one of the findings of the study was the result that, in contrast to another similar study [13], the ATC system plays an important role in conflict recognition by the pilots of the taxiing and the landing aircraft in all visibility conditions. It gives some implication for the operational designers with respect to the role and functional reliability of the ATC system. Further, it was found that for this particular scenario, the pilots of the landing aircraft play a somewhat larger role in conflict detection and resolution in comparison to the pilots of the taxiing aircraft. This finding is not surprising, as the pilots of the taxing aircraft have a more limited visual observation area in comparison to the pilots of the landing aircraft. These observational limitations of the pilots of the taxiing aircraft should be taken into consideration while designing this new operation.

Simulation experiments with the integrated agent-based model in Chapter 5 revealed drastically increasing probability of erroneous actions given high workload of human operators that leads to incorrect situation awareness. The integrated computational model that incorporates multiple factors, such as an interplay of personal characteristics of operators, their mental processes and environment provides a useful tool in understanding the whole coherent picture of incident related factors, their interdependences and thus creates a more correct incident risk estimation for safety experts and for the direct participants of the operating process (e.g. pilots, managers, support staff etc.).

Chapter 6 of the thesis presents the comparison of two computational models of spread of Situation Awareness in a group at different levels of abstraction across agents clustering dimension: the population-based agents cluster model and an agent-based individual model. This comparison demonstrated the complementarity of two approaches in the sense that the population-based approach is useful in predicting the general pattern of the dynamics of group beliefs at an aggregated level while agent-based modelling provides more insights in the emergence of this global pattern across time and agents clustering dimensions. These complementary dynamic models might be useful for better understanding such social phenomena as 'group think', norms, conformity and might contribute to useful recommendations for managers and policy makers.

Finally, the agent-based habit contagion model introduced in Chapter 9 may provide useful insights for health care policy makers regarding the spread of habits formation in a population. The model allows for predictions of behavioral patterns of healthy habits after the introduction of a campaign that aims at spreading of healthy habits in a population.

Contribution

The main objective of the research was to explore agent-based modelling capabilities for the analysis of human functioning in complex socio-technical systems according to the three-dimensional abstraction framework described in [3]. An overview of the position of each chapter of the thesis within the three-dimensional framework is given in Figure 1.

As can be seen in this figure, most models that were described in the present thesis are situated at an intermediate level across *process abstraction* dimension (cognitive process abstraction level), at temporally local level of *time* dimension and at a local level of *agent cluster* dimension of the cube. It concerns the models described in Chapters 1, 2, 3, 4, 5, 6, 8, 9, and 10. The grain size of a model is chosen by a designer based on the goals that the model is supposed to serve and modelling techniques that the designer intends to apply based on the goals. The main scope and goal of the present research was human cognitive functioning in complex systems, thus it is not surprising that the majority of the models occupied intermediate positions across the *process abstraction* dimension. The main methodology applied in this research was agent-based modelling that per definition implies the local level across the agent cluster dimension. It should be noted that though the three-dimensional framework was developed for formal models, the current research revealed the possibility to classify informal models in safety domain according to this framework as well. For instance, safety analysis models according to FRAM, STAMP and Event Trees modelling approaches can be also positioned within the cube.

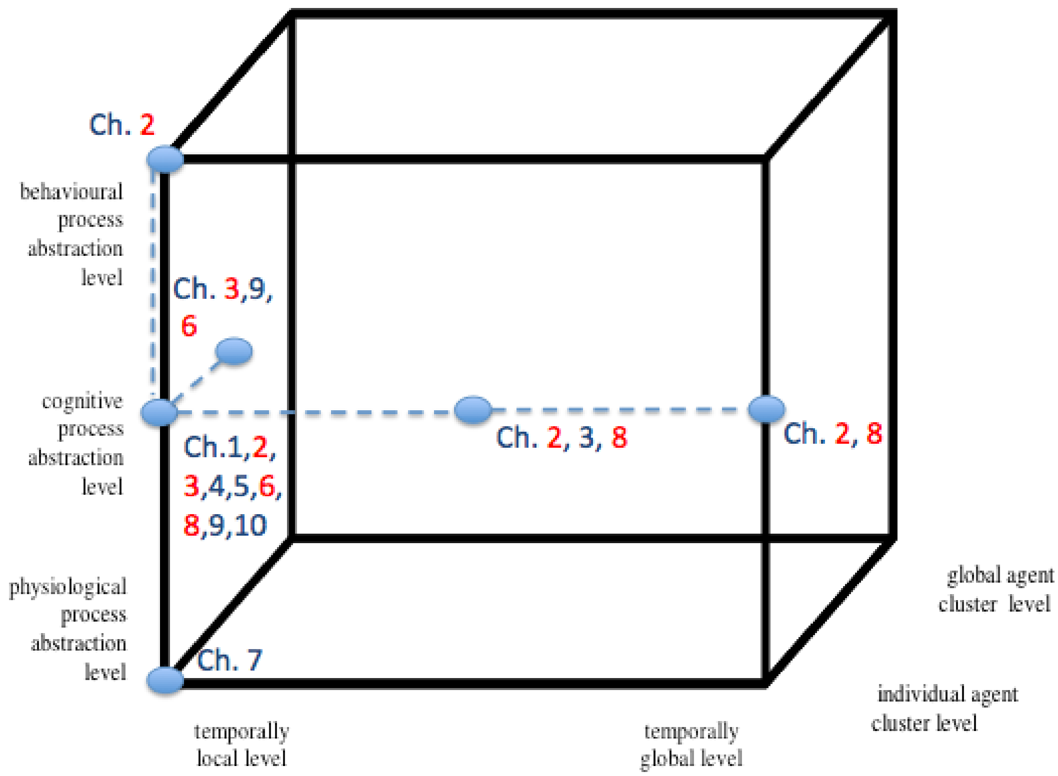

Figure 1. Position of the models described in the thesis within the three-dimensional models classification framework (indicated by the relevant thesis chapters).

Regarding the issue of establishing formal and informal interlevel relations between models at different levels of abstraction, several models were analysed from this perspective as well. It concerns the models from Chapters 2, 3, 6 and 8 (they are depicted in red in Figure 1). The interlevel relations defined between these models both formally and informally are represented by the dashed lines within the cube in Figure 1.

The main conclusion of the current research is that agent-based modelling proved to be a good approach for the analysis of systems in safety and healthcare domains where either complex dynamic relations are present (e.g. temporal dynamics in the runway incursion model in Chapter 3) or where cognitively complex human agents interact with each other or with technical systems in a complex manner (e.g. habits contagion model in Chapter 9). This type of modelling allows for the analysis of a system's behavior at different levels of abstraction, both at a global and at a local level of the three dimensions (see Fig. 1) and provides useful descriptions and insights into the emergent behavior of a system. This conclusion is in line with the statements about agent-based modelling found in literature [6], [7]. Agent-based modelling is not widely used in the safety domain yet; though the advantage of this type of modelling is well understood by many safety experts, see e.g. [11], [13]. Concerning the healthcare domain, agent-based modelling is quite a novel approach for dealing with complexity. However, it is admitted that it is a very promising approach for modelling and understanding large-scale processes and complex interactions [7]. In addition to this broadly recognized capability of the agent-based approach, the present thesis provides evidence for the appropriateness of the agent-based paradigm for the development of intelligent healthcare applications that interact with human users in order to provide ambient support for lifestyle and behavior change.

The present agent-based analysis was performed in the context of the agent-based models classification framework described in [3] across three dimensions (or modelling aspects): processes that are being modeled, representation of time and agents clustering. Performing the analysis within one classification framework allows for a structured way of modelling and for better overview of models' characteristics and capabilities. Moreover, this framework emphasizes the gradation between the levels of detail in different models and the possibility of mapping the models with different levels of representation, or grain size. This interlevel mapping is the essential factor in understanding behavior of complex systems. Understanding human functioning in complex systems better will help people not only to interact with the system, but will also help to design more resilient and effective socio-technical systems.

Among domain specific contributions of the present research, one can also mention the demonstration of advantages and disadvantages of formal incident analysis in safety critical domains by means of contrasting formal and informal incident analysis approaches and the exploration of the possibility of artificial agents to react dynamically to a human operator's functional state. For the healthcare domain, the present thesis explores the possibility of automatic support and coaching of chronic patients with respect to their therapy adherence or healthy individuals with respect to adopting and maintaining a healthy lifestyle. The formal model of habit learning described in Chapter 8 can be embedded in a software support agent capable of good habits coaching. The model-based healthcare application eMate that is described in Chapter 10 is now in the process of evaluation and validation. A preliminary pilot study performed with this system demonstrated its capability to interact with users in an effective manner and to persuade them to eat healthy food and to exercise more often. If the main validation study of the system provides promising results, the system can be brought to the market.

Research limitations

The main limitation of the present research is that validation (in the strong sense) of some of the domain models is missing; however the present thesis was concerned with the analysis of system behavior at different levels with the focus on the role of human agents in these systems. It is possible to claim that some type of validation was performed as the models either were constructed to represent a real life case study and they are able to reconstruct the behavior of a real-life multi-agent system, or the models represent the general realistic patterns that are described in the literature. Further, no validation of support models has been performed during this research. It may form the grounds for follow-up research within the agent-based ambient support paradigm, as this research is only a basic step towards the exploration of agent-based modelling techniques applied in safety and healthcare domains.

Specific limitations related to different parts of the research are also addressed in individual chapters of the thesis. For instance, in Chapter 3 comparing of an agent-based modelling approach to other influential modelling approaches in the aviation domain was performed only on one case study that substantially limits the possibilities of each of the approaches. The main limitation of the behavior change model and the support system where the model is incorporated is that it does not explicitly take social influences into consideration while they are considered to be an important determinant of human behavior. The model does not consider individual differences either and this can be regarded as a substantial limitation.

Future Work

In this thesis the mapping between models at different levels was established qualitatively, though these interlevel relations can be formalised in order to enable automatic mapping and translation between different types of models and to get more insights into the system's behavior. An example of the methodology for formal mapping based on ontology mapping between the models at different levels of abstraction can be found in [12]. The present work can be extended by performing this type of formal mapping between the models that contains ontology mapping.
For the domain of cognitive modelling and analysis, future research will focus on more complex rich cognitive models and try to perform their analysis in order to understand human cognition better. The possibilities of development and design of ambient support systems based on the analysis of human cognition can be further explored and the existing systems evaluated. Evaluation and agent-based analysis of the eMate system described in Chapter 8 is ongoing. An experiment with chronic patients will be performed with control and experimental groups. The control group

will be provided only with the limited functionalities of the eMate system: these subjects will not receive any motivational messages and direct feedback on their behavior while the experimental group will actively interact with the system. It will be evaluated afterwards whether the group with the full functioning eMate system will develop better therapy adherence and adopt a healthier lifestyle.

Additionally, in future the model of behavior change that forms the core of the eMate system can be initialized with certain values of external elements that are mapped into the internal perception of the same elements by an individual, for instance objective skills of a person to overcome barriers or social norms exist beyond an individual's perception of the same concepts and different individuals perceive their abilities differently and interpret cues from their environment differently. Thus each individual should have different parameters with respect to mapping of external cues into their internal mental representation. Further, the connections between different layers of this cognitive model can also be tuned depending on the individual's personality and experience. It is quite difficult to predict individual behavior without the knowledge of these individual parameters, though using average parameters one can approximate it.

Furthermore, formal support models can be developed based on habit formation domain models at individual and social levels as described in Chapters 8 and 9. These support models can in the future be embedded in ambient support agents that can help humans to learn healthy habits and to avoid unhealthy habitual behavior.

The work described in the present thesis was mainly focused on model development and model-based analysis. The issue of validation was not addressed in many of the presented models. Although some models were analyzed by means of logical verification (models in Chapters 2, 3, 8) that revealed the models' internal consistency and the model described in Chapter 10 was validated by means of a statistical analysis of empirical data, still more work is needed in order to validate the models described in Chapters 1, 7, 9 and 10 either by means of verification, sensitivity analysis, experts evaluation or empirical data. Further, in Chapter 3 comparison of four modelling approaches in aviation safety was performed only on one case study. It would be interesting to consider other case studies in the same or other domains for this type of comparison. At this moment, the work on applying the STAMP methodology to safety modelling in the pharmaceutical domain is ongoing. This methodology can be contrasted in the future to an agent-based modelling approach in pharmaceutical safety.

Concerning long term research, many models described in this thesis can be potentially developed into real life ambient support applications that provide support to humans during their functioning in abnormal contexts, e.g. in highly dynamic environments such as aviation and modern warfare or in the disease management and healthcare domain. This can be done by embedding support models that are based on relevant domain models into an ambient agent component according to the GAM framework [1], [2]. Up to now only one domain model described in Chapter 10 became part of an ambient support application in healthcare. Further, many models can be incorporated in serious games that aims at providing better insight into behavior of humans in joint cognitive systems. For instance, the domain models in Chapters 2, 5, 6, 8, 9, and 10 can become parts of serious games for safety experts, managers, policy makers, patients, healthy humans and operators both in safety and in healthcare domains.

Concluding remarks

It became clear that modelling and analysis of a human cognitive functioning in socio-technical systems has many challenges. This can be explained by the degree of complexity that it possesses. The human brain is a self-organized complex system that evolved after billions of years of evolution and it is regarded to possess the 'strong' form of emergence as the global behavior cannot be predicted by analyzing the behavior of separate local elements. We are not aware yet of all human brain functions and capabilities and about the precise structure of the human cognition system, the cognitive models presented in this thesis are only approximate simplifications of some small fraction of the brain's functioning. Analysis of human-designed complex systems, such as air traffic management that was studied in this thesis on the one hand can be characterized by a lower degree of uncertainty: at least the structure of the system is well-known. On the other hand, modern socio-technical systems are becoming more and more complex due to multiple couplings and interconnections between its elements and this tendency will bring even more challenges with respect to their analysis, understanding and making predictions of their behavior. In addition, human-designed socio-technical systems contain humans that interact with technical systems and this makes things more complicated since human functioning and behavior are parts of these systems and the human cognitive system is also a complex system in itself which behavior is characterized by strong degree of emergency and uncertainty.

Different ways of complex systems analysis are being explored nowadays. It is demonstrated in this thesis that agent-based modelling and analysis of complex systems can be a promising approach towards getting more insights into complex socio-technical systems' behavior.

References

1. Bosse, T., Both, F., van Lambalgen, R.M., and Treur, J. (2008). An Agent Model for a Human's Functional State and Performance. In: Proceedings of the Eighth IEEE/WIC/ACM International Conference on Web Intelligence

and Intelligent Agent Technology, WI-IAT'08. IEEE Computer Society Press, vol. 2, pp. 302-307.

2. Bosse, T., Hoogendoorn, M., Klein, M. C., & Treur, J. (2010). A Generic Architecture for Human-Aware Ambient Computing. Agent-Based Ubiquitous Computing. Atlantis Press, 2010. 35-61.
3. Bosse, T., Hoogendoorn, M., Klein, M.C.A., and Treur, J. (2010). A Three-Dimensional Abstraction Framework to Compare Multi-Agent System Models. In: Pan, J.-S., Chen, S.-M., and Nguyen, N.T. (eds.), Computational Collective Intelligence: Technologies and Applications, Proceedings of the Second International Conference on Computational Collective Intelligence, ICCCI'10, Part I. Lecture Notes in Artificial Intelligence, vol. 6421, pp. 306-319. Springer Verlag, 2010.
4. Bosse, T., Hoogendoorn, M., Klein, M., van Lambalgen, R., van Maanen, P. P., & Treur, J. (2011). Incorporating human aspects in ambient intelligence and smart environments. Handbook of Research on Ambient Intelligence and Smart Environments: Trends and Perspectives. IGI Global, 128-164.
5. Bosse, T., Jonker, C.M., Meij, L. van der and Treur, J. (2007). A Language and Environment for Analysis of Dynamics by SimulaTiOn. International Journal of AI Tools 16 (3) 2007, pp. 435-464.
6. Dam, K. H., Nikolic, I., & Lukszo, Z. (Eds.). (2013). Agent-based modelling of socio-technical systems (Vol. 9). Springer.
7. Denton, B. T. (Ed.). (2013). Handbook of Healthcare Operations Management: Methods and Applications. Springer.
8. Hollnagel, E. (2004). Barriers and accident prevention, Aldershot, Ashgate, 2004.
9. Leveson, N. (2011). Engineering a safer world: Systems thinking applied to safety, MIT Press, 2011.
10. NASA: Callback newsletter. (2012). http://asrs.arc.nasa.gov/publications/callback.html. July 10th 2012.
11. Shah, A.P., Pritchett, K.M. Feigh, S.A. Kalarev, A. Jadhav, K.M. Corker, D.M. Holl, R.C. Bea. Analyzing air traffic management systems using agent-based modeling and simulation. In Proceedings of the 6th USA/Europe Seminar on Air Traffic Management Research and Development , Baltimore, MD
12. Sharpanskykh, A., and Treur, J., (2010). Abstraction Relations between Internal and Behavioural Agent Models for Collective Decision Making. In: Pan, J.-S., Chen, S.-M., and Nguyen, N.T. (eds.), Computational Collective Intelligence: Technologies and Applications, Proceedings of the Second International Conference on Computational Collective Intelligence, ICCCI'10, Part I. Lecture Notes in Artificial Intelligence, vol. 6421, pp. 39-53. Springer Verlag, 2010.
13. Stroeve, S.H., Blom, H. A.P., Bakker, G.J. (2011). Contrasting Safety Assessment of a Runway Incursion Scenario by Event Sequence Analysis versus Multi-Agent Dynamic Risk Modelling, *ATM 2011*.

14. Stroeve, S.H., Doorn, van B.A., Bakker, G.J. (2012). Safety Assessment of a Future Taxi into Position and Hold Operation by Agent-Based Dynamic Risk Modelling. Journal of Aerospace Operations, Vol. 1, 2012, pp. 107-127.