Risk-informed control systems for improved operational performance and decision-making

Christoph A. Thieme[1]*, Børge Rokseth[1], and Ingrid B. Utne[1]

Abstract

Autonomous systems, including airborne, land-based, marine and underwater vehicles, are increasingly present in the world. One important aspect of autonomy is the capability to process information and to make independent decisions for achieving a mission goal. Information on the level of risk related to the operation may improve the decision-making process of autonomous systems. This article describes the integration of risk analysis methods with the control system of autonomous and highly automated systems that are evaluated during operation. Four main areas of implementation are identified; (i) risk models used to directly make decisions, (ii) use of the output of risk models as input to decision-making and optimization algorithms, (iii) the output of risk models may be used as a constraint in or modifying constraints of algorithms, and (iv) the output of risk models may be used to inform representations or maps of the environment to be used in path planning. A case study on a dynamic positioning controller of an offshore supply vessel exemplifies the concepts described in this article. In addition, it demonstrates how risk model output may be used within a hybrid controller.

Keywords: Autonomous systems, supervisory risk control, risk-based decision-making

1 Introduction

Autonomous systems for transportation, surveillance, and exploration become a near future reality. Autonomous ships, autonomous cars[1] and autonomous drones in the air[2] and underwater[3] are being used or are approaching the mass marked. Autonomous systems are characterized through the ability of making decision independent from an external supervisor to achieve a set goal[4]. This does not mean that no human supervisor is involved in the operation process. The difference between autonomous and highly automated systems is small. Highly automated systems are characterized through the automatic execution of

[1]Department of Marine Technology, Norwegian University of Science and Technology, Otto Nielsen Veg 10, 7495 Trondheim, Norway
*Corresponding author: christoph.thieme@ntnu.no

several functions, whereas higher level decisions for control of the system are given by an operator.

Concerns regarding reliability and the risk level of such systems need to be addressed in order to make autonomous systems a success[1,5,6]. Risk can be described as the combination of undesired scenarios (chain of events), the associated consequences and the associated uncertainty with respect to the occurrence and magnitude of the events and consequences[7], as summarized in equation (1). Risk is often associated with negative outcomes, however, may include positive outcomes. Safety of an operation or a system is established if the risk has been reduced to an acceptable and tolerable level[8]. That means that measures have been taken to reduce the frequency or probability of occurrence of consequences and to mitigate the impact of negative consequences.

$$R = \{s_i, \boldsymbol{C}_i, \boldsymbol{U}_i\}_{i=1}^{n} \qquad \text{(eq. 1)}$$

Testing, assurance and compliance with safety-related standards are important tools to ensure safe operation of systems. Risk assessment is carried out to support decision-making in socio-technological systems. Hence, information from the risk assessment process may be useful not only for design, but also as input to the decision-making capabilities of the control system of autonomous systems in operation.

However, autonomous systems' decision-making based on risk information during the operation of the system, is not much discussed in the literature, except for a few examples. Pereira et al.[9,10] present an approach to mission and path planning of underwater gliders, considering both the surface ship traffic density, bathymetry and currents. Their approach fits in the risk definition, expressed earlier, uncertainty (of motion and position) is combined with undesired events and consequences (loss of the underwater glider).

Lefevre et al.[11] propose a risk aware path planning algorithm based on hierarchical path planning and different A* implementations for autonomous underwater vehicles. To select the best path the objective function minimizes the path cost, which includes risk as a factor.

Johansen et al.[12] present a Model Predictive Control (MPC) path planning that considers the maritime safety navigation rules (COLREGS). The path optimization uses hazard information from a ship simulator and uses the COLREG rules as constraints. Hazards, i.e., obstacles, are identified from an electronic map. Brekke et al.[13] discuss the above and other approaches that attempt to build collision avoidance systems for autonomous vessels.

Bremnes et al.[14,15] combine a risk informed Bayesian belief network (BBN), implemented as decision graph with the control system, to make decision about proximity to ice when operating an underwater vehicle. System status, environmental factors and mission related

factors are considered in the BBN. A risk index is calculated, which will limit the allowable minimal distance to the ice.

Hobbs[16] discusses the possibility to combine risk analysis methods with safety critical functions in order to ensure safe operation. This may include a control functionality in a system. DNV GL[17] recently discussed the concept of a digital twin incorporating uncertainties with respect to the systems condition and environment to allow for risk-informed decisions. A digital twin is an assembly of mathematical formulations and models that represents and abstracts a system.

Despite some efforts to include risk information in control systems, only a few attempts have been made to incorporate risk analysis models or methods directly in the control system of an autonomous system. Such approaches are needed to support online risk analysis during operation and will lead to risk-aware autonomous system behaviour[6,18]. Utne et al.[6] recently introduced the concept of supervisory risk control of ships. Supervisory risk control refers to the assessment of risk information based on operational data to support decision-making by the autonomous system. The goal is to improve the intelligence of autonomous systems and thus ensuring safe operation and successful mission execution.

The objective of this article is to assess and outline possibilities to include risk analysis methods and models in the control systems of autonomous systems. Thereby, attempting to opening the application of risk analysis methods to be implemented with the control system, for safer operation and decision-making during an operation. The goal is to bridge the fields of control engineering (cybernetics) and risk engineering, to create a mutual understanding of topics and concepts.

Different risk analysis methods are discussed and mapped on different elements of a generic control systems. For this purpose, an example control system for autonomous systems is described. The article provides input for improving the control systems' decision-making processes for autonomous and highly automated systems in the future. Given the scope, this article does not focus on assurance, pure hazard identification methods, testing efforts, or processes related to standards for safety-related systems.

The next section describes the generic control system and frequently employed control approaches and techniques. Section 3 describes the necessary background in risk and safety terminology, together with commonly employed methods in the field. Section 4 describes the interfaces and possible approaches to connect the fields of cybernetics and risk engineering. An example based on a ship control system exemplifies the concepts throughout the article. Section 5 summarizes and concludes the article.

2 Control system design and methods

2.1 Control system architecture

In order to determine which type of risk information may be relevant for and used by the different parts of a control system, understanding both the constituent elements of a control system of an autonomous system (Figure 1) and the control system architecture is important (Figure 2). Figure 1 shows a generic component architecture of an autonomous system. The office systems as a supervising and managing entity have been added to emphasize that the autonomous system will receive goals and maintenance from humans. There is a difference between the operational and organizational environment of the entity owning and using the system.

- **System hardware:** The physical components that make up the autonomous system and interact with the control system. External system interaction refers to the interaction or communication with other systems or operators through a dedicated communication system and the necessary hardware. External communication may include sharing of maps for mutual localization, system communication of information additional to position information, or systems' coordination of actions.
- **Realtime network:** BUS/ Real time ethernet network inside the system for communication of data and commands between the system's internal components: sensors, actuators, control system, and external communication system.
- **Office network:** Network used for operating organizations internal communication between computer systems and servers for data exchange. The office network enables communication with the lower levels of the control network.

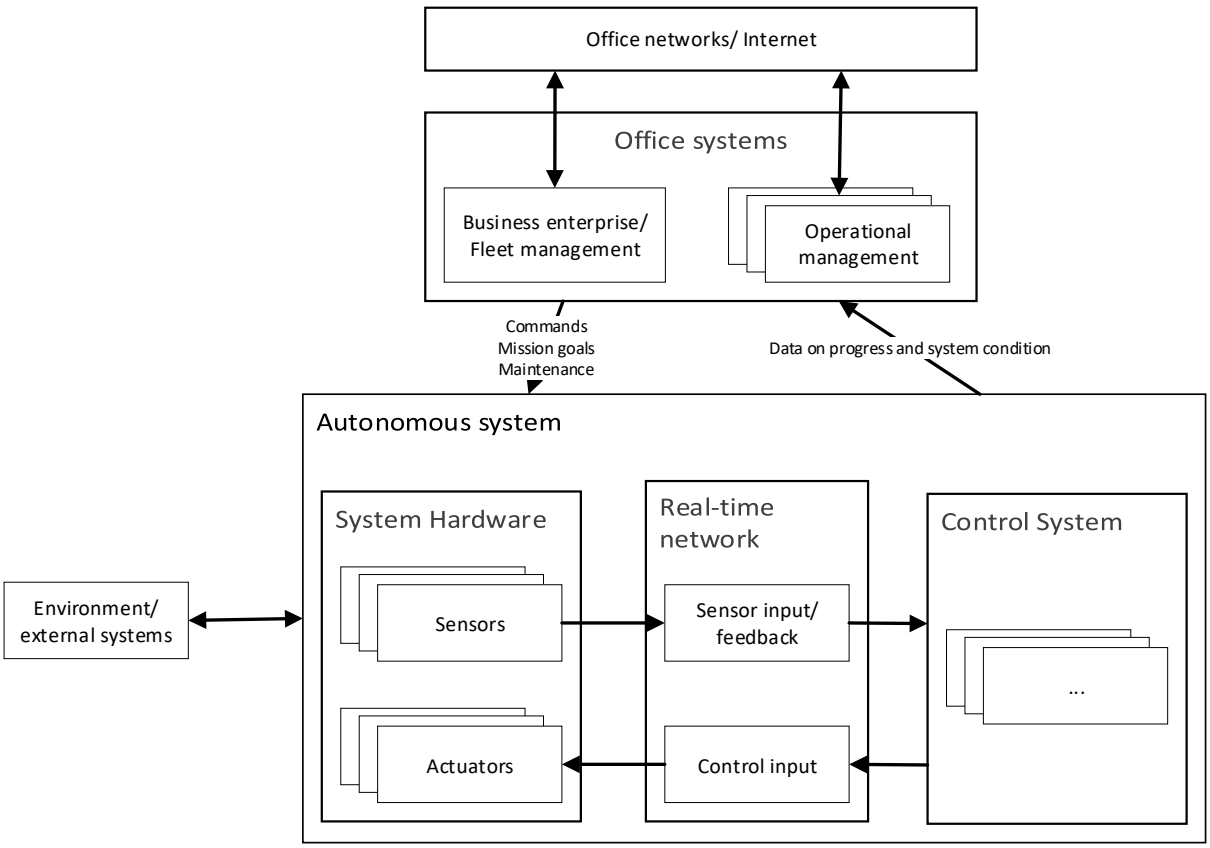

Figure 1 Generic component architecture for autonomous systems, developed from (Sørensen[19], Pendleton et al.[20])

The following systems are found in the component architecture:

- **Office systems:** The operating organizations management system to ensure the operation of systems. This includes business enterprise, (vehicle/ vessel) fleet management, operational management systems. Operational management addresses the decisions to ensure the operability of a system, that includes maintenance, condition monitoring and operational risk assessments to optimize safety and operability.

The generic control system for autonomous systems is depicted in Figure 2. The control system is based on Sørensen[19] and Pendleton[20]. The focus of the control system architecture is on planning and re-planning and the plant control elements. These are most relevant for online risk monitoring and incorporation of risk models. Some elements in the figures are depicted through several layers of boxes, for example, sensors, or behavioural planning. This means that several of these elements may be present in the system. Behavioural planning, hence, may encompass several algorithms to assess the optimal behaviour with respect to different circumstances.

The following elements (or main functions) are found in the generic control system:

Perceive: The system's ability to collect information and extract relevant knowledge from the environmental and internal measurements. This knowledge may be used to understand the environment and its context (signs, signals, obstacles, etc.) or to localize the system's position in the environment. In addition, this knowledge may be used by the system to determine its state. Different algorithms and systems have been developed for detection of traffic, roads, objects, and machine vision. No methods will be discussed in detail in this article. Generally, filters, machine learning and artificial intelligence methods may be employed for the perceive function. Risk information may make use of the information from the perceive function and vice versa. This is outside the scope of this article but should be addressed in future work.

Plan and re-plan mission: The system's ability to make decisions to achieve its mission goals. There are several types of planning that address different time horizons and types of decisions.

- Mission planning: Aims at finding an overall mission plan and high-level objectives, e.g., key points of interest to be visited, routes to be taken, etc. Re-planning occurs sporadically.
- Behavioural planning: Aims at decision making to ensure that the system follows rule restrictions and behaves in a conventional and safe manner. This may include collision avoidance behaviour, respecting safe distances, etc. behavioural planning and re-planning occurs for short and medium time horizons varying from seconds to several minutes. It sets local objectives to be achieved[20]. This also includes the decision of changing the level of autonomy, e.g., when the operator needs to take control from an autopilot.
- Motion planning/ local optimization: It aims at finding the set of actions that are needed to reach a local goal/ waypoint. Planning and re-planning considers a time horizon from milliseconds to minutes. This should be efficient, complete (time finite) and safe, i.e. avoid collision.

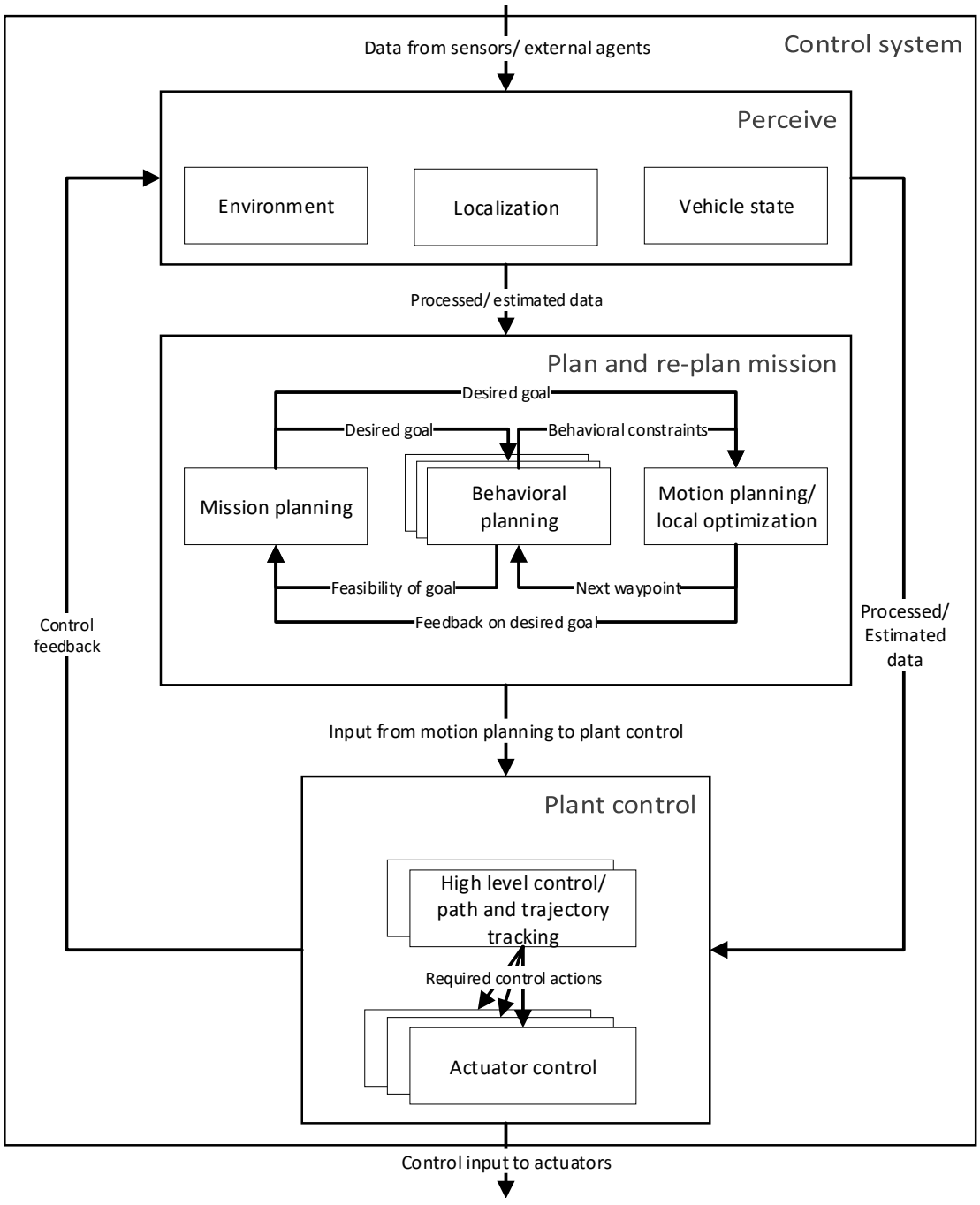

Figure 2 Generic control system structure for autonomous systems, developed from Sørensen[19] and Pendleton et al.[20].

Plant control: Comprises the ability to execute the plans and actions that have been developed in the planning layer. The time horizon of this layer is short term and varies between milliseconds and few seconds. Plant control is often based on feedback from observers that estimate position/ velocity, estimate bias, and filter noise from the measurement signals in the perception element of the control systems.

- High level control and path and trajectory tracking: Is based on a simplified mathematical model of the system and controls/ guides the system's behaviour on a system level, i.e., the position to be reached in the next time steps.
- Actuator control: The desired control actions produced in the high-level controllers need to be translated to outputs of the actuators, e.g. rotation angles and rotational output frequency. The actuator control produces the signals that trigger these desired responses in the actuators.

The different elements and their differentiation are in practice often not as clear. The plan and re-plan mission-elements are often not as nuanced. Several elements may be considered together, e.g., motion and behavioural planning. Similarly, there is often an overlap of the elements, e.g., motion planning and high-level plant control are often difficult to differentiate. However, a clear statement of the different functionalities, as described in this article, may assist in identifying control system requirements and capabilities.

One aspect of control systems that has received a high level of attention is fault tolerant control. This concept is related to detecting failures in a control system and the associated sensor/ actuation hardware[21]. Fault tolerant control should prevent a failure event on the system level arising from a fault in the component. Three types of component faults can be distinguished: sensor fault, plant fault, and actuator fault. However, risk management and related risk analysis processes cover a wider scope than fault tolerant control and contingency handling, such as collision avoidance[18].

2.2 Development process of a control system

A control system is typically developed in six steps of which some are carried out iteratively[22]. These steps and the proposed interaction with risk assessment are shown in Figure 3. The detailed steps and definitions for risk assessments will be further explained in Section 3.

Figure 3 extends the normal development process of a control system. The normal steps (1-6) have been adopted from Šabanović and Ohnishi[22]. Currently, risk assessments are only used to little extent to give input to the design of the controller e.g., to identify control system elements that need special attention. This input is mainly to the performance and requirement specification. For safety-relevant systems, processes and requirements for the design and development of control systems that are relevant for safety are laid out in, e.g., the generic industry standard IEC 61508[23], the automotive standard ISO 26262[24], or the railway standard EN 50128[25].

Risk assessments also can give input to risk models that are built to be used for decision support in the control system.[6] That means that they are part of the control architecture and

the controller design. This is the focus of this article, the implementation and connection of risk analysis methods with the control system. These risk models and their output may be used by the control systems to make decisions or adapt the systems behaviour or performance to the current risk level. This is a novel application not yet required nor implemented.

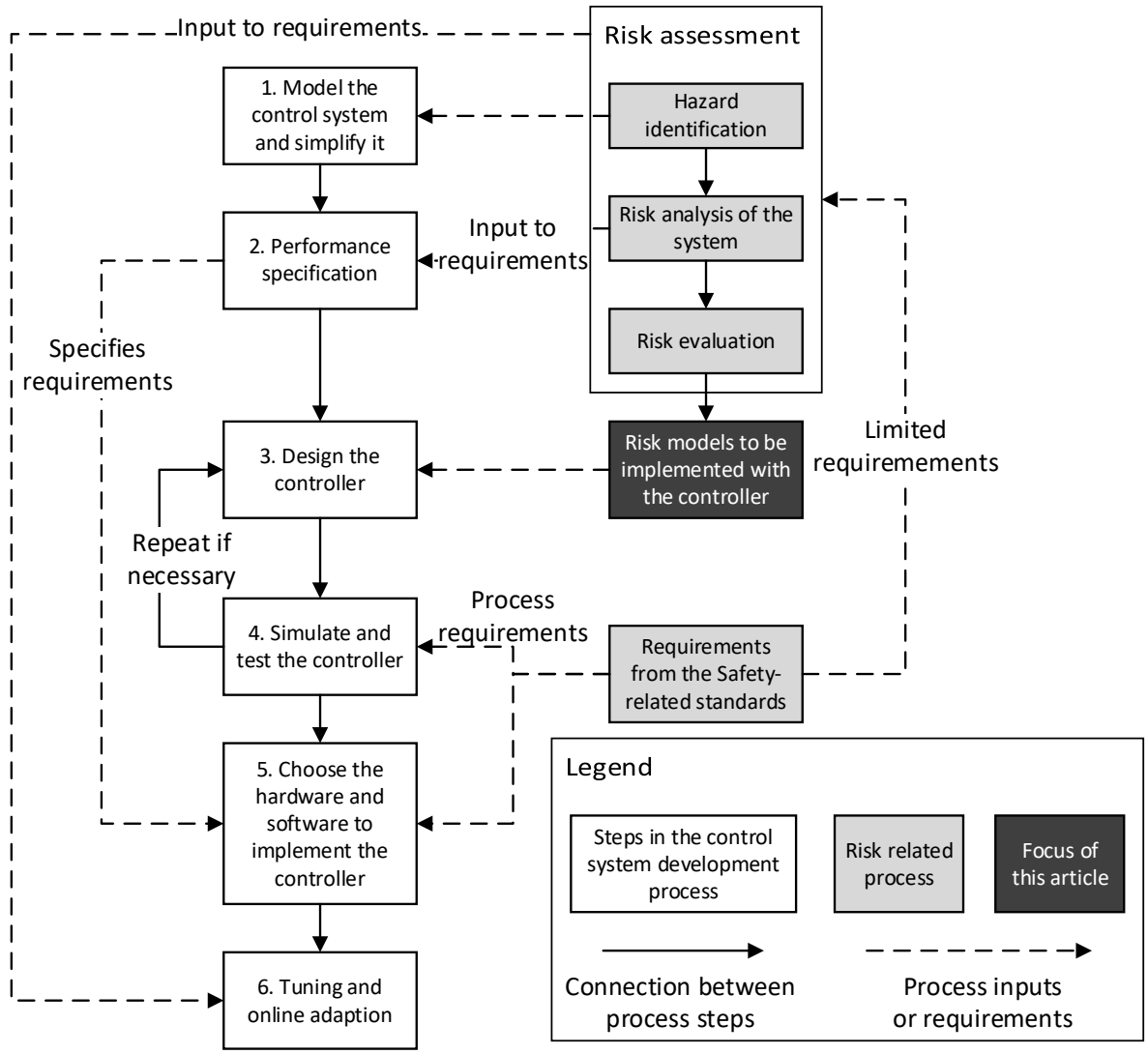

Figure 3 Development process for a control system of an autonomous system and how information from risk models can contribute to the control system.

2.3 Control approaches and algorithms

For the different elements of a control system (Figure 2), there are commonly used control techniques and approaches. These are important to understand when determining how risk assessment and models can provide useful risk information. Table 1 summarizes control techniques that are commonly used in the different elements of the control system presented in Section 2.1. These control techniques are presented in detail in the following sub-sections.

The characteristics of the techniques are summarized as computation time, and requirements and element of the control system they are mostly applied in. The requirements summarize the information needed for this technique to design the control systems adequately. The computation time describes the time needed to evaluate the algorithm. This is categorized into low, medium, high, where low refers to a speed in the area of few milliseconds and high can be in the magnitude of several 100 milliseconds or seconds. This information is useful when evaluating the application of the risk analysis models in the control techniques, since a risk model that will take long to evaluate may slow down a control algorithm. The column for requirements summarizes data or knowledge needed for implementing the method.

2.3.1 Mission planning and re-planning techniques

Mission planning identifies sub-goals to be achieved for reaching the mission goal. The most common mission planning approach for autonomous/ self-driving cars, is *graph search*[20]. Additionally, *finite state machines* (FSM) and *Markov decision processes* are often used to determine achievable mission goals and plan the next sub-goal. *Model Predictive Control* (MPC) may include behavioural constraints. However, the focus of *MPC* is motion planning, hence, *MPC* will be described there.

Graph search is mainly concerned with identifying the length of the shortest path and/ or the waypoints to follow this shortest path. Well known methods include, *Dijkstra algorithm*, and *A** algorithm and their variations and extensions[26]. Sometimes the shortest path search is referred to as visibility graph[27].

The *Dijkstra's* algorithm is a standard approach to the shortest path problem in a network or a discretized map. To find the shortest path the algorithm compares the distance from its initial node to connected nodes and then moves on to the shortest connected node. With each step the shortest distance between the initial node and additional nodes that can be reached is updated with the shortest distance. In this way, the shortest path to the end node can be identified from the vector of shortest distances[26].

*A** is a goal oriented shortest path algorithm that uses Dijkstra's algorithm and adds a potential function, favouring nodes to be checked that are closer to the target node[26]. Several improvements and combinations with other graph search techniques have been suggested to the *Dijkstra* and *A** algorithms to improve their performance[26].

Table 1 Summary and evaluation of control engineering techniques described and explored in this article

Control technique	Control system element	Computation time	Requirements	References
Artificial potential fields	Motion planning	Medium – High	Identification of obstacles and state/ characteristics of the system.	
Finites state machines	Behavioural planning, Mission planning, Motion planning	Low	All mission states need to be predefined and their transfer conditions.	
Graph search algorithms	Mission planning, Motion planning	Medium – High, depending on the algorithm and the scope of the problem. The original Dijkstra and A* have comparably high time.	Environmental map needs to be available in enough detail, with feasible regions.	26
Hybrid control techniques	Actuator control, Behavioural planning, Mission planning, Motion planning, Plant control	Low – High	All possible/ foreseeable system states need to be captured by the model.	28,29
Mixed/ Partial observable Markov decision processes	Mission planning, Behavioural planning	Medium – High	Model needs to incorporate system states and conditions. Parameters may be learned during operation.	
Model predictive control	Actuator control, Behavioural Planning, Motion planning, Plant control	High	Definition of relevant system states that reflect the control problem.	
Probabilistic road maps	Motion planning	Medium, dependent on number of sampling points	Environmental map needs to be available, with feasible regions.	30
Proportional/ Integrative/ differential control	Actuator control Plant control	Low	Plant inputs and characteristics of the actuators.	
Rapid-exploring random trees	Motion planning	Medium, dependent on number of sampling points	Environmental map needs to be available, with feasible regions.	27,30
Reachability guidance	Motion planning	Medium – High	Set of states and possible actions in the next time step.	
Signal/ Linear temporal logic	Behavioural planning, Motion planning	Low – Medium	All signals and their mapping to actuator output need to be known.	31,32
Velocity obstacles	Motion Planning	Low- Medium	Trajectory and speed of the obstacles, own systems state	33
Voronoi diagrams	Motion planning	Medium	Obstacles need to be known or identified.	

FSMs consist of states and the transitions between these states. The transitions describe the triggering condition and the action that will be triggered and lead to the next state. A state machine may only be in one state at a time. In a graphical representation the states are represented by circles. Transitions connect the states and are labelled with the conditions and actions[34]. In the context of autonomous systems and their control *FSM* may be used to decide on the next goal to be achieved or the overall mission planning approach given the current situation.

Markov models are very similar to *FSM*. Their use is described below for control and in Section 3.2 for their use in risk analysis. *Markov decision processes* are used to reason and make decisions, when uncertainty is to be considered. *Partial observable Markov decision process* or *mixed Markov decision processes* may be employed. In *Markov decision processes*, the values of actions can be directly computed, and decisions can be taken immediately by calculating the reward of each action. A domain expert is needed to build the model and assess the parameters of such a model. Efforts are undertaken to grow the *Partial Markov decision processes* and parametrize them during the system operation, through machine learning approaches[35].

Mixed Markov decision models have similar characteristics. Not all relationship and their parameters need to be available when the model is defined. Instead variables are defined in the model that can modify the nodes in the Markov model[36]. *Markov decision processes* may be used for deciding on and optimizing mission sub-goals. In addition, they may be employed to plan the behaviour.

2.3.2 Behavioural planning and re-planning techniques

Behavioural planning employs currently two main techniques[20]. These are (i) *FSM* and (ii) *signal temporal logic* (STL) or *linear temporal logic* (LTL). *Markov decision processes* and *Hybrid control* may also be used. The latter is considered for the purpose of this article a behavioural planning technique, since it alters the system behaviour according to the circumstances. *FSM* and *Markov decision processes* have already been described in the previous section.

STL and *LTL* are formal methods, originally designed for verification of the temporal behaviour of reactive software systems. *STL* is addressing analogous and mixed signal circuits. In *STL*, formulas are defined that constraint the signal. These formulas are using negation, Boolean combinations, or temporal operators to define the requirements with a standard logical notation[32].

LTL is different from *STL* in the way that there is assumed a relationship between the sensor and the system behaviour. That means that for a given input, certain actuator output is

expected. However if an input is out of its expected bounds the system behaviour cannot be guaranteed[31]. *LTL* may be used to translate requirements defined through logical statements into hybrid controllers (controlling a mixture of discrete and continuous system behaviours). This approach is advantageous for reactive tasks, i.e., reacting to information that is collected at runtime[31].

Hybrid systems theory provides a formalism for the integration of multi-functional controllers combining discrete events and continuous control. *Hybrid control systems* are characterized through either unreliable state measurements, high sensitivity to errors in the measurement of states, unsatisfactory performance of the system with only one state feedback controller, or a combination of these characteristics[28,29]. *Hybrid control systems* may be modelled through four elements; flow set, flow map, jump set, and jump map. Discrete changes (jumps) occur when the state is an element of the jump set. A system that has a discrete behaviour, i.e., jumps, can be modelled and controlled as a hybrid system[28]. *Hybrid control* may be used to combine several other control techniques and switch between control algorithms automatically based on different parameters.

2.3.3 Motion planning and re-planning techniques

Two main motion planning techniques may be differentiated: *combinatorial planning* and *sampling-based planning*. Combinatorial planning attempts to find a complete solution over the planning space while representing the space exactly. Special case solvers then exploit convenient properties of the representations.

For lower dimensional planning problems *visibility graphs*, and *Voronoi diagrams* are used. Often discretization in space is done to apply trajectory search algorithms, such as, *trajectory search tree*, or other graph search methods (see Section 2.3.1). *Artificial potential fields*, *Linear Temporal Logic* (see Section 2.3.2), or *mixed or partial observable Markov decision processes* may also be applied over discretized cells in space. *MPC*, described later, is also relevant for motion planning. However, most applications are found in plant control (Section 2.3.3).

Sampling-based planning techniques generate a trajectory graph (also referred to as roadmap or feasibility graph) by taking samples of the space. Feasible paths are than constructed from these graphs. Examples are *probabilistic road maps*, or *rapidly exploring random trees (RRT)*. Extension to these *probabilistic roadmaps* and *RRT* methods have been proposed, e.g., dynamic versions[30]. Other methods employ *reachability guidance* to check the sample spaces faster for connectivity.

RRT is an incremental sampling and searching approach. The trees are constructed incrementally, where the resolution is improved in the process. However, no explicit

resolution parameters are set. The paths identified in this way are stored in a tree structure, where the sequence of construction is random. For each point an edge is drawn to connect a new random point with the closest branch[27]. For planning a trajectory, the tree is grown randomly. The tree growth is stopped if the last connected point is in the desired goal region[30].

The *artificial potential fields* method assigns fields that either attract (e.g., the goal) or are repellent (obstacles and no-go areas). The algorithm is mainly employed for static obstacles. The goal of the algorithm is to minimize the repellent force. This may lead to non-optimal solutions, through local minima. However, algorithms have been developed to find globally optimal solutions[37].

Probabilistic roadmaps are used for finding paths over several waypoints. They employ a two-stage approach for identifying paths. Firstly, roadmaps are constructed by connecting randomly sampled points. Secondly, a connection between the desired start and endpoint are sought from the sampled trees[30]. The maps that are collected can be described as forests. The evaluation and identification of the path from a starting to point to an endpoint may use shortest path methods as described previously.

Reachability analysis and *guidance* assess the possible set of states and positions of a system in the future, starting from known states and the initial position. This also takes into account the uncertainty connected with state measurements[38]. Sampled data is shown to be enough for this approach. *Reachability analysis* can be used to plan and assess reachable waypoints in future time intervals, predicting them over a time horizon that is a multiple of the time intervals. For this purpose, the effect of control actions on the system are assessed with respect to their results. The optimal control actions will minimize the cost of reaching the goal, while assuring that control actions will lead to a desirable state. Obstacles are be considered by defining zones that are not acceptable.

Voronoi diagrams are used to represent the equidistance from several points in a Euclidean plane or space. Lines, so called edges in the diagram, represent the points that have the same distance from these points. At a vertex at least three edges meet. A vertex is at least equidistant from three points in the plane. Different measures may be used, such as the Euclidian distance or the Manhattan distance, to find the edges. *Voronoi diagrams* can be constructed by different algorithms[39]. For planning the diagrams and their evaluation are used to find clear routes, between, e.g., obstacles, where the lines would represent the furthest distance[20].

The *velocity obstacles* approach assesses velocities that would lead to a collision given that the system would move with these velocities. The algorithm applies to moving obstacles and

the trajectory of the obstacle needs to be known or estimated[33]. Different approaches exist to solve for the optimal speed of the vehicle to avoid collision.

2.3.4 High level plant control techniques and approaches

The purpose of plant control, which is the lowest element in the control system architecture (Figure 2), is to determine appropriate inputs to the actuator control level such as the desired motion which satisfies the needs from the motion planning element[20]. This may be necessary to coordinate the efforts of several actuators. Examples can be the autopilot or guidance system of a ship which determines an appropriate heading for following a path or a set of waypoints, or the cruise controller of a car which determines an appropriate engine control throttle position for achieving the desired speed.

Appropriate control approaches depend on the application. To determine the appropriate heading of a ship, *Line of Sight* (LOS) steering laws are a possible approach. *LOS guidance* follows a straight path between two waypoints, where the desired heading of the ship points towards some point a certain distance ahead on the path ([40], pp. 256).

Proportional, Integral and Derivative (PID) controllers can be applied, where a control error is calculated from the feedback signal and the corresponding desired state (e.g., the measured speed of a car and the setpoint from the driver). The calculated control input consists of a term that is proportional to the control error, a term that is proportional to the integral of the error over time, and a term that is proportional to the time differential of the control error[20]. In certain applications, adequate performance in terms of stabilization, tracking and disturbance rejection cannot be achieved using PID control. In such cases, non-linear control methods, such as gain scheduling, sliding mode control and feedback linearization may be applicable. Common for PID control and non-linear feedback control strategies are that a control signal is calculated based mainly on measurement feedback signals.

An alternative to *PID* or the *nonlinear feedback control* methods above are model-based methods such as *MPC*. Rather than determining the control input based on the control error, *MPC* seeks to identify a set of control input to minimize a cost function over a future prediction horizon by utilizing a mathematical model of the system. In addition to the control error, the cost function may penalize other costs such as fuel consumption or time delays. For example, *MPC* applied to an adaptive cruise controller can improve the system by achieving acceptably small speed deviations while minimizing fuel consumption and passenger discomfort[41].

2.3.5 Actuator control techniques and approaches

In actuator control, the objective is to make each individual actuator behave according to the inputs from the plant control layer. To achieve this, the input from the plant control must be translated to control input appropriate for the actuator under control. In general, the same control design techniques are applicable on this level as for plant control. Control methods, such as the *PID* controller are common for applications such as, electrical motor speed controllers, which commonly use a *proportional – integral* (PI) control law[42]. In other applications, such as manipulator control with direct joint drives, nonlinearities due to joint interactions makes it necessary to use more advanced controller design, such as non-linear feedback control techniques, where the effect of the motion of other joints are be cancelled out[43].

2.4 Example – dynamic positioning control system

To illustrate the elements of a control system, a generic control structure for a dynamic positioning (DP) system adapted from Brodtkorb[44] is used. A DP control system is a control system that enables a maritime vessel to maintain its position and heading (station-keeping) by means of propellers and thrusters. In addition to station-keeping, a DP control system typically provides precise manoeuvring and trajectory following functionality[45]. The main control objective of the DP control system is to calculate setpoints for each active thruster and propeller (such as, rotational speed setpoints) such that the motion of the vessel corresponds to the desired motion of the vessel. The desired motion is nowadays typically defined by an DP operator.

In the control system presented in Figure 4, the operator provides the guidance system with a desired motion. This can be location and heading setpoints or a trajectory. The function of the guidance system is to transform the operator's input into reference states. The reference states are signals that represents the wanted values for actual controllable states of the vessel such as the north and east position, the surge and sway speed and the heading and yaw rate. The operator also determines the power management mode, which has influence on fuel consumption and power capacity reserve.

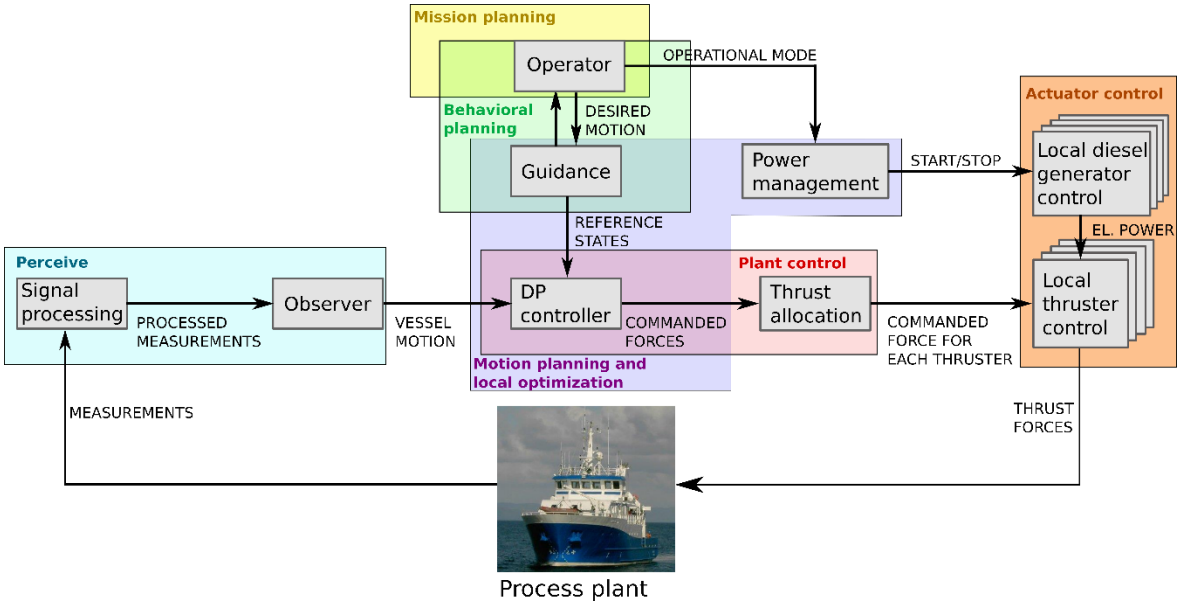

Figure 4: Block diagram of a dynamic positioning control system adapted from Brodtkorb[44]

The DP controller calculates the commanded forces in surge, sway, and yaw. The thrust allocation is responsible for allocating these forces into a commanded force for each of the active thrusters. The process plant represents the actual vessel influenced by the generated thrust forces. Different states related to the motion of the vessel, such as the north and east position and the yaw rate, are measured using for example differential global positioning system (DGPS) and gyro compass. These measurements are verified and pre-processed in the signal processing before being fed to the state observer. The state observer is responsible for filtering noise and high frequency wave motion response from the measurements as well as reconstructing unmeasured states. This is usually achieved using an Extended Kalman filter or a nonlinear passive observer[46,47]. The power management is responsible for, among other things, to start and stop diesel generator sets to ensure sufficient amounts of available power[48].

The vessel in the example will carry out sub-sea intervention with a remotely operated vehicle. For this purpose, the vessel is fixed in a position by use of DP. For this operation, the elements of the generic control systems (cf. Section 2.1) are as below. The example will be expanded later in the article related to risk analyses methods, which are presented in the following Section.

Perceive: Sensors, position reference system, state observer.

Mission planning and re-planning: For this type of operation, missions are planned at an IMR subcontractor's office and re-planned by DP operator and offshore engineers immediately before mission. Typical planning tasks address ship arrival at the field, exact positioning, and time needs and availability for conducting the mission.

Behavioural planning and re-planning: Answer questions such as: Is the necessary position/trajectory of the vessel located within the safety zone (planned in IMR subcontractor office), is the weather state suitable to conduct the operation and is it expected to continue to be suitable throughout the mission (offshore engineers and DP operators), and will the ship be oriented upwind from the installation (offshore engineers and DP operators)? Select and enter setpoints or trajectory into guidance system (DP-operator). The power management is also part of the behavioural planning.

Motion planning: Transform setpoints or trajectories into suitable reference states in guidance system. This partly overlaps with the DP controller.

Plant control: Calculate commanded forces (DP controller) and allocate commanded force for each thruster (Thrust allocation). Ensure that an appropriate amount of available power.

Actuator control: Control of the individual thrusters.

3 Risk and risk analyses methods

3.1 Definition of the concept of risk

Risk, as previously defined in equation (1), is a combination of scenarios, consequences, and the associated uncertainty. Risk and its nature are discussed in detail in, e.g., Kaplan and Garrick[49], Aven[50,51], or Rausand and Haugen[8]. In the context of this article risk is understood as a combination of events that may lead to unwanted consequences, such as damage to people, the environment, or assets. Events may be initiated by the system itself or be the consequence of the environment acting on or interacting with the autonomous system. The occurrence of the events and their consequences are expressed through a measure of uncertainty. Uncertainty is often expressed as probability. The probability can be characterized through an underlying distribution that reflects the uncertainty.

Risk assessment is the process to identify relevant risk contributors, analyse and evaluate the level of risk[7]. Hazard identification aims at recognizing and describing risks that are relevant for the operation of an organization, both positive and negative. Risk analysis is the process of risk comprehension and determination of the level of risk[7]. During risk analysis, risk models may be developed and used to reflect the relationship between the risk and the use and/or design of the system under analysis based on available information. Risk models are developed through risk analysis methods. Risk analysis should consider the sources of risk, uncertainties, likelihood, consequences, events, scenarios, and risk controls together with their effectiveness[7].

The goal of risk assessment is to obtain information that supports decisions to be made with respect to a system's design or operation. The results of risk assessments may lead to design modifications, safety requirements, or safety constraints and limits, that may be implemented in the design of a system or of an operation. Results may also be implemented by constraining, modifying, or limiting a systems behaviour in certain situations, i.e., speed limits or minimal distance to an object. Safety requirements describe a goal with respect to safety and the constraints describe how these goals can be achieved.[52]

The next section summarizes commonly used risk analysis methods. The description focuses mainly on quantitative methods since they can potentially interface with a control system and give input through numerical values. It is important to note that some qualitative methods may provide input to (more detailed) quantitative methods[6].

3.2 Methods for risk analysis and risk level monitoring

The methods described in this section are commonly used for risk analysis[8]. These methods are commonly used in the design phase, however, in this article their use and implementation in the control system is explored.

Table 2 summarizes the main features of the methods with respect to application. The required expertise refers to the required knowledge of the assessors with respect to the method and the system. Required data refers to the amount of data that is required initially to conduct the analysis. Effort required refers to the amount of resources (experts and time) for conducting the analysis. The reference column points to the reference used to find this information.

Bayesian belief networks (BBN) and *decision graphs* are a combination of a graphical representation of relationships and the quantitative relationship between the different influencing factors [53,54]. The factors are represented through nodes and the relationships through directed arcs between the nodes. *BBN* and decision graphs are acyclic. The quantification relies on the Bayesian theorem (eq. 2). The tables associated with the nodes, so-called conditional probability tables, describe the probability of a nodes state given any combination of its parents states. Decision graphs are extended *BBNs* that include nodes and logic for possible decisions and the assumed effect of these decisions[53].

$$P(A|B = \frac{P(B|A)\cdot P(A)}{P(B)} \quad (2)$$

Table 2 Summary of selected risk analysis methods.

Method	Application	Output type	Quantitative measure	Required Expertise	Required data	Effort required	Reference
Bayesian belief network/ decision graphs	Identify hazards, estimate risk, decide between options	Quant.	Probabilities of target nodes.	High	Medium	Medium/ high	[55]
Decision trees	Compare options	Quant.	Best decision according to the circumstances.	Moderate	Low/ medium	Medium	[55]
Dynamic Flow Graph Method	Identify hazards, Analyse consequences, Analyse likelihood	Qual./ Quant.	Probability of possible events/ outcomes.	High	High	High	[56,57]
Event tree analysis	Analyse consequences Analyse likelihood	Qual/ quant.	Probability of possible outcomes.	Moderate, but depends on complexity	Medium/ high for quantitative analysis	Medium/ high	[55]
Fault tree analysis	Analyse causes Analyse likelihood	Qual/ quant	Probability of the top event (e.g., system failure).	Moderate	Medium/ high for quantitative analysis	Medium	[55]
Markov models	Analyse likelihood	Quant.	Probabilities or percentage of time spent in the states.	High	Medium/ high	Medium	[55,58]
Markov Cell to Cell Mapping Technique	Analyse consequences, Analyse likelihood	Qual./ Quant.	Probabilities or percentage of time spent in the states.	High	High	High	[56,57]
Petri nets	Analyse risk states Analyse likelihood	Qual./ Quant.	Probabilities or percentage of time spent in the states.	Moderate, depends on complexity	Medium/ high	Medium/ high	[59,60]
Risk/ safety indicators[2]	Analyse/ represent risk level	Qual./ Quant	Categorial/ continuous evaluation of factors abstracting risk.	Medium	Medium/ High, dependent on complexity	High, dependent on the complexity and method	[61,62]
Simulations/ Monte Carlo analysis	Analyse likelihood	Quant	Probability of possible events/ outcomes.	High	Medium	Medium/ high	[55]

[2] Risk/ safety indicators are not a risk analysis method. However, they are a set of tools to monitor the level of risk and support operational decisions.

Decision trees are used to model decisions from an initial decision to the resulting outcomes, following the path of decisions on the way[55]. These are especially useful if sequential decisions need to be taken. Probabilities for decisions and their success can be assigned together with utilities for the final outcomes. The decision with the highest utility should then be chosen. Binary decision diagrams have been developed to model a systems reliability during different mission phases and used as part of the mission planning process.[63–65]

The *Dynamic Flowgraph Method* (DFM) is a multi-valued, discrete-time logic modelling framework to represent a cyber physical system [66]. The DFM allows for modelling of physical, functional, and dynamic characteristics of a system, with the aim to validate and analyse the design with respect to reliability and safety. The method can be used to assess the effect of a failure on the system behaviour and for backwards reasoning, i.e., inferring how a certain system behaviour may be produced [67]. The model is represented as a diagraph. The inputs, parameters within the system, and outputs to the system are represented as vectors and the relationships between these are modelled through deterministic or probabilistic relations[66].

Event tree analysis (ETA) is used to analyse the possible consequences that may arise from an adverse event, respectively. Event trees are analysing what other events may occur and what the consequences of these events will be. Corresponding diagrams facilitate the communication of risk with these methods.

Fault tree analysis (FTA) uses Boolean logic to analyse how an adverse event may occur. Fault trees are analysed from a top down perspective trying to identify possible reasons for the adverse event. Events are connected through logical gates that are used to structure the occurrence of events.

Markov models are used as a tool to analyse the state behaviour of a system. *Markov models* build on the same logic and theorems as *Markov decision processes*. The technique allows to model the operational states of a system and the transition to failed states. Failed states may be restorable to a functional state or maybe absorbing, representing a system state that is not restorable. The analysis can be time dependent or steady state, whereas the latter allows for simpler calculational methods [58]. *Hidden Markov models* are a special form of first order *Markov models*, where the states are hidden. Their occurrence is associated with a probability. *Hidden Markov models* are considered a form of a dynamic (time-dependent) *BBN*.

The Markov Cell to Cell Mapping Technique (MCCMT) separates the system's states in cells analogous to the finite element method[56]. The states can be process variables, system component condition, or system configuration. The system's behaviour is modelled through

discrete-time transitions among the cells. Transitions are modelled through a set of equations or algorithms that represent the physical and control laws the system is subjected to[56,57]. The technique has been developed for verification and validation of model-based control systems[56]. The level of computational time of the method can be controlled through truncating low probability branches.

Petri nets are bipartite graphs that also contain nodes and directional arcs. The node types are places and transitions. The directional arcs model local states and events[60]. *Petri nets* may be timed or not. *Timed Petri nets* can include information about temporal relationships, delays or dependencies. *Petri nets* are used in several industries for reliability and event modelling.[60]

Risk and safety indicators are not a risk analysis method per se. *Risk indicators* are actually derived from risk models developed for risk analysis in operation[61]. *Safety indicators* is an umbrella term for indicators that provide insights in conditions that reflect the safety performance, such as, barrier quality, scenarios, or decision-making.[68] *Risk and safety indicators* shall reflect the level of risk/safety of an operation and hence mirror the condition of the system for the current conditions to operate the system without unwanted events. Indicators may be process parameters, performance parameters of the system, or reflect organizational qualities. These indicators are system, process and company specific. Hence, they need to be developed purpose specific. Different methods and approaches have been developed to identify *Risk and Safety Indicators*, c.f., [61,68].

Simulations can have different forms and aims. One of the aims may be to predict possible outcomes and associated likelihood through mathematical models. Models may, for example, include physical system models, environmental models, or reliability models of sub-systems and components. *Monte Carlo simulation* that are run several iterations and makes use of random sampling from underlying distributions may be used as one approach[55]. *Simulations* enable analysts to analyse the system behaviour in case of a failure or accident event and the impact of subsequent corrective actions.

3.3 Example – dynamic positioning control system and risk level

With respect to the aforementioned DP system, some events may contribute to an increased risk level, such as a failure of local thrusters or diesel generators the inability of the system to produce sufficient thrust force to counteract environmental forces, erroneous position estimates from the observer causing the DP controller to drive the ship out of position, or the failure of the hardware on which the thrust allocation, DP controller, the guidance system, the observer or signal processing software is running. Ideally, the system is designed with these and other failure types in mind. Monitoring risk may give input to the optimal operation

and control of the system (with respect to risk). Considerations when building a risk model for implementation with the control system should include, among others[6]:

- Weather conditions
- Potential for damage in the event of uncontrolled motion
- Proximity to other vessels or land
- Technical condition of the machinery
- Status of the energy supply system
- Operational mode

A risk model may be used to calculate the probability of risk related events for DP operation, affected by the above factors, for example potential damage in the event of “uncontrolled motion" and failure of technical components of the machinery. For a DP system, all the above-mentioned risk analysis methods are relevant, depending on the objective and scope of the analysis. In the following, the methods are evaluated with respect to control system techniques.

4 Combining control and risk analysis techniques

4.1 Implementing risk-based information in the control system

Figure 5 summarizes the possible identified ways to implement the risk methods with the control techniques. This is an extension of Figure 2. It highlights how the risk analysis methods, risk indicators, or simulations may give input to the elements of the control system. The risk models, simulation, and indicators will use the information collected in the perception element, since this is the element that collects and prepares information.

In general, four possibilities are identified regarding how risk models, simulations and risk indicators may be used. Firstly, information may be used directly in the elements of the control system, for example, as variable in a decision or optimization algorithm. Secondly, the information may be used to modify algorithms, through adapting allowable system states or modifying the systems behavioural and safety constraints. Thirdly, the risk analysis models may be used to identify or determine the state of the system, which is then fed forward in the control algorithms. Lastly, information assessed through risk models may inform environmental maps that are used for path planning. The following sub-sections describe possible application in more detail.

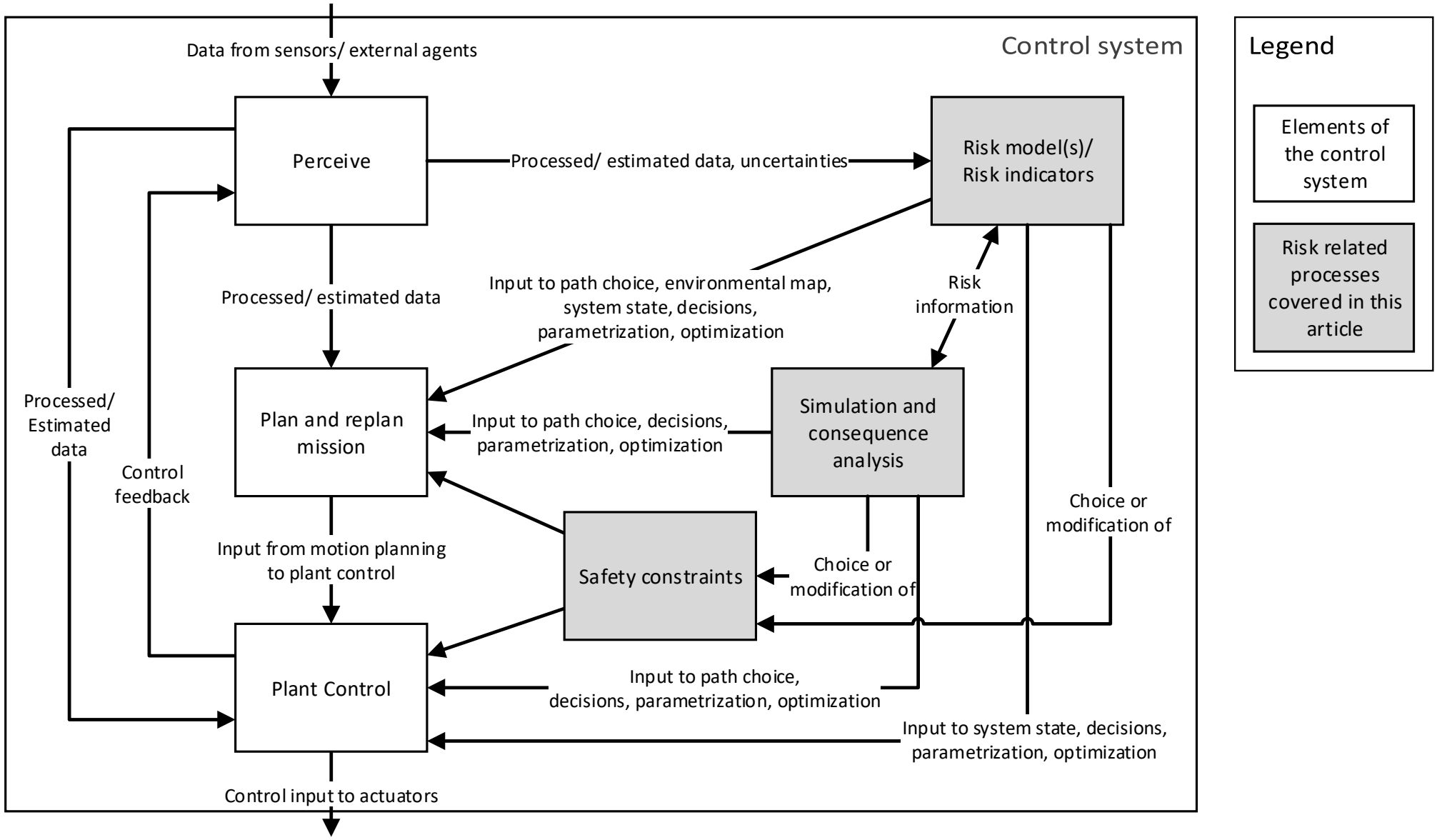

Figure 5 Overview of possible combinations of control techniques and risk analysis techniques

Table 3 Mapping risk analysis methods to the control techniques, which are mainly related to planning and re-planning. Abbreviations: BBN – Bayesian Belief Networks, DFM – Dynamic Flowgraph Methodology, FSM – Finite State Machines, MCCMT – Markov Cell to Cell Mapping Technique.

Control Technique	Risk analysis methods									
	BBN/ decision graphs	**Decision trees**	**DFM**	**Event tree analysis**	**Fault tree analysis**	**Markov models**	**MCCMT**	**Petri nets**	**Risk/ safety indicators**	**Simulations/ Monte Carlo analysis**
Artificial potential fields	Influence the potential of the fields.	No.	Influence the potential of the fields.	No.	Influence the potential of the fields.	Influence the potential of the fields.	Influence the potential of the fields.	Influence the potential of the fields.	Influence the potential of the fields.	Influence the potential of the fields, simulate trajectories.
FSM	Determine the state of the system probabilistic. Give input to decisions.	Determine the next best state to achieve. Help to determine the current state.	Determine the risk level of the possible states, determine the current state, use as decision criteria for state transition.	Determine most likely Measurement of being in the current or a future state.	Identify the current state, determine possible future (accidental) states, or state transitions probabilistically.	Very similar to FSM, may be developed in parallel. Determine future state transitions or determine the current state.	Very similar to FSM, may be developed in parallel. Determine future state transitions or determine the current state.	Very similar to FSM, may be developed in parallel. Determine future transitions of the states, determine the current state.	Determine the risk level of the possible states, determine the current state, use as decision criteria for state transition.	Simulate to choose the order of states, or to determine if a state transition is necessary
Graph search algorithms	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Determine paths to follow based on risk considerations/ decisions.	Determine the risk level/ future system states to inform path choice.	Determine the risk level/ future system states to inform path choice.	Determine the risk level to inform path choice.	Determine the risk level/ future system states to inform path choice.	Determine the risk level/ future system states to inform path choice.	Determine the risk level/ future system states to inform path choice.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Simulation of paths to assess the risk-based cost.
Hybrid control	As part of the decision process to switch behaviour or to inform the switching criterion. Fits well with the	As part of the decision process to switch behaviour.	Analyse the outcome of switching and use this information in the process of switching.	Analyse the outcome of switching and use this information in the process of switching.	As part of the decision process to switch behaviour.	As part of the decision process to switch behaviour or to inform the switching criterion.	As part of the decision process to switch behaviour or to inform the switching criterion. Fits well with the	As part of the decision process to switch behaviour or to inform the switching criterion.	Values produced by the indicators may be used as switching criterion, or to influence	Simulations of future development to inform the switching criterion or trigger switching.

Control Technique	Risk analysis methods									
	BBN/ decision graphs	**Decision trees**	**DFM**	**Event tree analysis**	**Fault tree analysis**	**Markov models**	**MCCMT**	**Petri nets**	**Risk/ safety indicators**	**Simulations/ Monte Carlo analysis**
	property of hybrid systems that state measurement are uncertain.						property of hybrid systems that state measurement are uncertain.		the switching thresholds.	
Mixed/ Partial Markov decision processes	BBN/ decision graphs have a similar structure. Inform decisions and models.	Both are decision making processes, no combination possible.	Input to the decision, based on the current risk level	Asses the measurement of choices and use for optimization.	Input to the decision, based on the current risk level	Use the states to inform the Markov decision process. Both are in the same framework implying ahigh compatibility.	Use the states to inform the Markov decision process. Both are in the same framework implying a high compatibility.	Use the states to inform the Markov decision process. Both are in the same framework implying a high compatibility.	Input to the decision, based on the current risk level	Input to parametrization or input to decision process.
Probabilistic road maps	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Influence the length or cost of a path in the connection phase of the algorithm.	Identify or predict the state of the system to connect points or choose paths.	No.	No.	Identify or predict the state of the system to connect points or choose paths.	Identify or predict the state of the system to connect points or choose paths.	Identify or predict the state of the system to connect points or choose paths.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Identify or predict the state of the system to connect points or choose paths.
Rapid-exploring random trees	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	No.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Assess the measurement of following a certain path that gives input to the length.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Influence the length or cost of a path, may give input to a risk-aware map of the environment.	Simulate several paths and select the optimal path.

Control Technique	Risk analysis methods									
	BBN/ decision graphs	**Decision trees**	**DFM**	**Event tree analysis**	**Fault tree analysis**	**Markov models**	**MCCMT**	**Petri nets**	**Risk/ safety indicators**	**Simulations/ Monte Carlo analysis**
Reachability guidance/ analysis	Determine the state of the system probabilistic. Give input to decisions.	Determine the optimal state to be achieved in the next time step. Help to determine the current state.	Determine the risk level/ future system states to inform action choice.	Determine most likely measurement of being in the current or a future state.	Identify the current state, determine possible future (accidental) states, or state transitions.	Determine the risk level/ future system states to inform action choice.	Determine the risk level/ future system states to inform action choice.	Determine the risk level/ future system states to inform action choice.	Determine the risk level/ system state to inform action choice.	Simulation of paths to assess the risk-based cost.
Signal/ Linear temporal logic	Input to constraints based on the system state. Use as input parameters, or assessment of decisions.	Choose between different constraints to be used, or input to decision making.	Input to constraints based on the system state. Use as input parameters.	Input to constraints, use of risk as Analyse different outcomes and set constraints.	Use the risk level as input parameter.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Choose constraints or predict output of the system.
Velocity obstacles	Input to the constraints based on measurement and systems state.	Input to the choice of allowable velocities.	Input to constraints based on the system state. Use as input parameters.	No.	Use the risk level as input parameter.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Input to constraints based on the system state. Use as input parameters.	Choose optimal path/ parameters.
Voronoi diagrams	Give input to a risk-aware map of the environment. Input to path choice.	Input to path choice.	Give input to a risk-aware map of the environment. Input to path choice.	No.	Input to path choice.	Give input to a risk-aware map of the environment. Input to path choice.	Give input to a risk-aware map of the environment. Input to path choice.	Give input to a risk-aware map of the environment. Input to path choice.	Give input to a risk-aware map of the environment. Input to path choice.	Simulations may be used to identify the best path of a given set of Input to path choice.

4.1.1 Plan and re-plan mission

Table 3 summarizes the potential implementation of the risk analysis methods with the control techniques that are mainly related to the planning and re-planning subsystem of the control system.

Artificial potential fields search for a minimal gradient path and for a minimal cost. Almost all risk models may be employed in combination with *Artificial Potential fields*. The calculated risk level may be used to influence the artificial potential that represents the environment and obstacles therein. Different system measurements and states may give input to the respective risk models and their evaluation. *Simulations* may be also used to test the identified paths and assist in choosing low risk path options. *ETA* and *decision trees* are not deemed suitable since no information that could inform the *artificial potential fields* can be drawn from these methods.

For *FSM* the risk models can give input to the state assessment of the system at present or in the future. Furthermore, some risk models can contribute to the decision-making process of FSM, to assess whether the state of the system should be changed. This could be through using the risk information assessed in the models or use the decision output from, e.g., the decision tree or the *BBN*.

Graph Search algorithms may use risk information in two ways. Firstly, the risk information may be incorporated in the map of the environment, increasing the cost for areas that are expected to exhibit a higher risk level. This may then in the optimization of the path be considered as a longer or more costly path. This combination is similarly presented in Pereira et al.[9,10] and Lefebvre et al.[11]. *Probabilistic road maps* could be similarly informed by the risk model information to create risk-aware road maps. The assessment of the optimal path uses graph search methods. Secondly, the risk models may be used to make decisions on which path to choose, by using the state and risk information built in the models to make decisions.

Hybrid Control may use the risk model output in two ways. Firstly, *BBN/ Decision Graphs* and *Decision Trees* may be used to identify risk-based switching of the algorithms. Secondly and in general, all the risk models may be used to assess the current or the future risk level. This information may be used to inform the switching process, e.g., as decision parameter or as modifier of decision criteria.

Mixed or *Partial Markov decision processes* are very closely related to *Markov risk models*, *MCCMT*, and *Petri nets*. These are state-based assessment methods. Hence, the latter two may be used to directly incorporate risk-based reasoning in the mixed or partial *Markov*

decision processes. In general, all the risk models may be interfaced with *Mixed* or *Partial Markov decision processes*, providing risk information to the decision process.

Like *graph search algorithms*, *RRT* may use information from the risk analysis methods to modify the length or cost of the paths. Hence the optimal, shortest path will be minimizing risk over the path length. Since decision trees provide decisions the implementation in *RRT* seems not directly possible, since no decisions are needed as input for the *RRT*.

Reachability guidance/ analysis may be combined with risk models in several ways. Firstly, risk may be considered a state of the system that needs to be controlled. Secondly, the risk may be used as optimization criteria, together with other parameters to find an optimal path. Thirdly, risk-based information may be used to define zones that the system should not enter and the size of these. Therefore, may all types of risk models be used as input. Especially state-based models, such as, *Markov models*, *MCCMT,* or *petri nets*, may be suitable for all three possibilities.

STL and *LTL* set requirements to the system to ensure operational success. Hence, risk models may be used to give input to these requirements, e.g., modify them depending on the risk level. Another combination could be to use risk as a parameter to be constrained and ensured by *LTL*. Since *STL* aims at the sensor input, risk is not a measurable sensor parameter to be constrained.

For *velocity obstacles*, risk models may give input to the determination of collision candidates, e.g., by considering the current systems ability to manoeuvre or determine its and the obstacle's position. In addition, risk may be incorporated as part of the determination of safe velocities, by modifying the margin of not allowable speeds in proportion with the risk level. *ETA* is not deemed suitable for *velocity obstacles* since its consequence assessment will not provide information on allowable speed.

Voronoi diagrams are used to determine the largest optimal distance between several obstacles. Risk information may be used to determine the minimal required distance to the objects, i.e., modify the boundaries based on risk information. If several feasible paths are identified the risk models may be used to identify or choose the path with lowest risk. Only using *ETA* as input to *Voronoi diagrams* is not deemed suitable.

4.1.2 Plant control

Table 4 summarizes the potential implementation of the risk analysis methods with the control techniques that are mainly related to the plant control subsystem of the control system. Control approaches, such as *PID* controllers, may use risk model output to be tuned. This tuning may, for example, increase the rate of acceleration in the face of high risk,

to be able to handle a certain situation. This may then be not the optimal point with respect to energy consumption or other parameters. With such an approach of dynamic tuning, it is important that the tuning is bound to the stable regions of operation of the controller. Most risk models that provide information on the state of the system or its future state may be used. *ETA* and *FTA* is deemed unfit, since the information gained from the models is a quantitative measure of failure or of the associated consequences. This does not provide information or input to possible measures to be taken.

For *MPC* all types of risk techniques may be used to give risk input. The risk information may be used in two ways. Firstly, the it may be used as a parameter in the optimization problem, minimizing or optimizing the risk with respect to other operational parameters. Secondly risk may be used as a constraint, meaning that certain risk levels are not permittable.

Table 4 Mapping of the control techniques that are mainly related to plant control to the risk analysis methods.

Risk methods	**Control method**	
	Proportional, integral derivative control or non-linear control	**Model predictive control**
BBN/ decision graphs	Based on the situation adapt and optimize the parametrization	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Decision trees	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., as a constraint.
Dynamic Flowgraph Method	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
ETA	No.	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
FTA	No.	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Markov models	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Markov Cell to Cell Mapping Technique	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Petri nets	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Risk/ safety indicators	Choose between different parametrizations of the controllers	Use risk as parameter for optimization, i.e., in the optimization function or as a constraint.
Simulations/ Monte Carlo analysis	Simulate the system and choose the optimal controller	Simulate the system and choose the optimal controller. Use risk as parameter for optimization, i.e., in the optimization function or as a constraint

4.2 Example – dynamic positioning control system with risk information

This section discusses how control techniques may use risk models and risk analysis methods in a hybrid system controller. The purpose of the example is to illustrate and discuss how one of the combinations may be realized. A full development and implementation of the case study would exceed the scope of this article and is not attempted.

The example described in Section 2.4 is used as basis for further exemplification. The DP control system is modified with a risk-based power management controller. The controller provides automatic switching between the economy mode and the high-power mode. These controllers are today based on static power reserve requirements. That means that independent of the situation a certain power reserve should be available to serve suddenly rising power demands. These limits are defined through the maximum allowable load[69]. The novel aspect of this case study is the implementation of a risk model to determine safe and efficient power reserves

For the case study it is assumed that there are two power management modes available. In the high-power reserve mode, more generators are active, to provide additional power availability for the thrusters, when needed. Generators are operated with a low load, which is not optimal with respect to fuel efficiency. The economy mode is operated with fewer generators, closer to the optimal operation point. However, in case the thrusters require more power, only a comparably lower power margin is available. Hence, in a sudden high demand scenario the ship may experience a blackout, due to overload and hence drift off position[70]. In this case study example, control of the power reserve relies on the risk level and the load level. This is described in the remainder of this section.

The case study ship is operating in the proximity of an offshore oil and gas installation, to carry out subsea maintenance. The scenario is depicted in Figure 6. In DP operation, safety is of high importance. A loss of power, due to insufficient power reserves is the main hazard. As a consequence, loss of position and collision with the offshore installations may result. Such a collision may lead to severe accidents. Hence, risk indicators are proposed for determining the allowable power reserve.

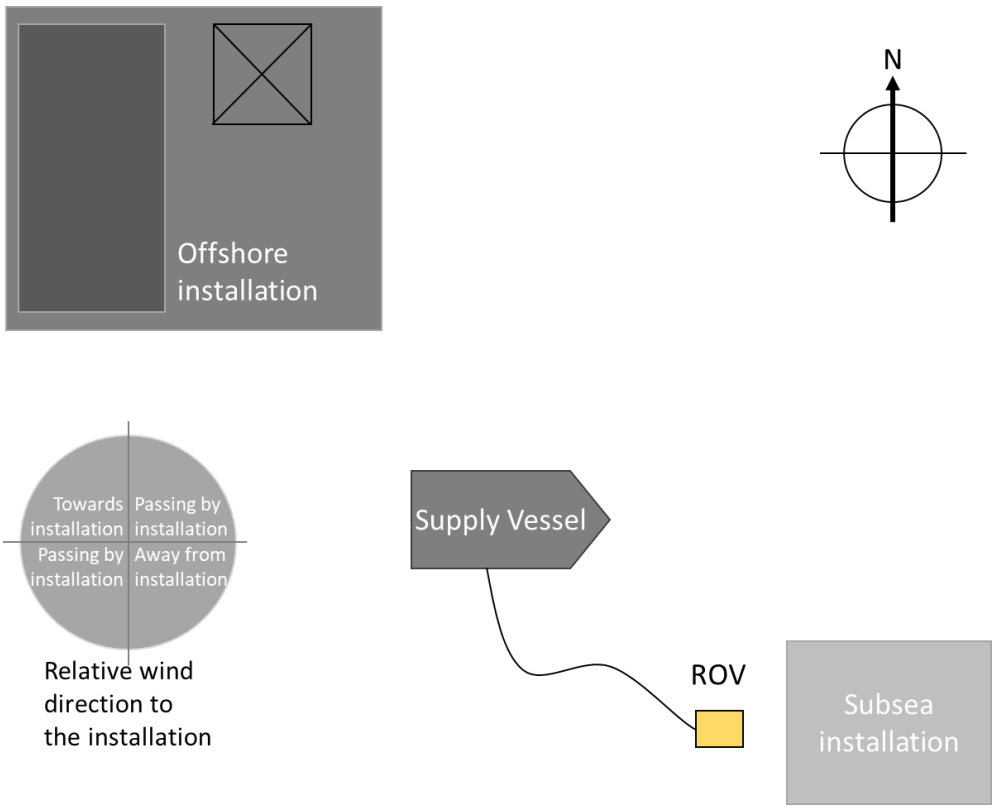

Figure 6 Scenario assumed in the case study.

Risk indicators that can be employed are listed in Table 5. The risk indicators were developed based on the information described in Section 3.3, in this section and in the articles by Rokseth et al.[48,71]. Risk indicators should be linked to a risk model, e.g., a BBN[72]. The development process of the risk indicators is not further detailed as this is outside the scope of this article. The purpose here is illustration.

The changes of the risk level with respect to time may be described using indicators 1 – 5. This may be used in the performance monitoring of the hybrid control law evaluating which algorithm to use in the power management system the high-power reserve mode, or the economy mode. Indicators 1 – 5 would form the flow set of the hybrid controller. Indicator 5 (available power) is part of the jump set of the hybrid controller since it will create a discrete response to the change of the controller. Permissible states of the flow set are defined in the flow map by the operational limitations set by the system and physical limitations.

Table 5 Indicators suggested for the case study.

ID	Indicator	Measurement source	Description
1	Distance to installation	GPS signal and digital chart	Distance to the installation, measured from the closest point of the vessel to the closest point of the installation
2	Wind direction with respect to installation	Measurements on board, estimation through estimators (e.g., [73])	The wind direction with respect to the location of ship and offshore installation, three cases may be differentiated (towards installation, passing the installation, away from the installation, c.f., Figure 6).
3	Wind speed	Wind measurements on board, estimation through estimators (e.g., [73])	The wind speed that is met by the ship.
4	Sea state	Estimation through sea state estimates, (e.g., [74])	The wave height and frequency acting on the ship.
5	Available power	Directly from the power system	Percental usage of the current available power for the DP system (c.f., [70])

Considerations regarding the jump map, indicating to change the controller based on the jump set, are described briefly. In case the forces acting on the vessel are low less generators are needed. The reserve on available power can be lower. In case, the wind direction would point towards the installation a larger reserve is needed. If the sea state and wind are strong, even more available power is needed and hence more generators are needed. More detailed considerations on the jump map are also presented in Thorat and Skjetne[73] with respect to the dynamics of the system. The presented risk indicators are already used by human operators when deciding whether to change the setpoint of the DP operation further away from the installation, or to abort mission. The suggested approach employs these indicators to automatically switch the operation mode.

Another approach to include risk information for the current example could be to determine the risk with respect to blackouts and drifting in the installation/ off position. This model could use different states and be fed by real-time measurements of the system and environment. However, the development of such a model would exceed the scope of this article. In addition, several of the proposed risk analysis methods are of limited use for assessing the risk level including all system dynamics, feedback loops and interaction hazards[75].

5 Discussion and conclusion

This article addresses the integration of risk analysis and models in the control system of autonomous and highly automated systems for risk level evaluation during operation. Four main areas of application of risk models in control systems are identified. Firstly, models can

give direct input to the mission planning and behavioural planning through risk-based decisions. Models that are mainly relevant for this purpose are BBN, and decision trees.

Secondly, risk derived from the risk models may be used as a decision or optimization criteria. All risk model types can give input to the different control techniques. Likewise, all types of control techniques may benefit from risk model input in this way. Thirdly, the risk models output may be used as constraint or modifying a constraint in the control techniques. Lastly, maps or environmental representations that incorporate risk information area useful for path planning methods, such as, graph search, artificial potential fields, probabilistic roadmaps, or Voronoi diagrams.

The few publications that take risk into account for decision-making and planning in the control system, use risk information mainly in the mission or path planning tasks of a system. Path planning often overlaps with behavioural planning and decisions. Currently, risk considerations on the plant control level are not addressed in the literature. In addition, the literature, covers the incorporation of the risk methods with the control techniques just too a little extent.

This article presents a starting point to systematically include risk analysis methods in the control system, to improve safety of operation and support decision-making. Due to the complexity of the material and the amount of available control techniques and risk models, only the most prominent methods were described. This description is also limited in detail. The identified possible relationships and considerations need to be explored further in more detail. A challenge, limiting this work, with existing risk analysis methods is that their results may be difficult to adapt to be used as input.

A challenge for use of risk models within a control system is arising from the risk assessment methods mainly being static and the nature of the model development. Hazard identification and risk analysis often include the use of experts and brainstorming assessments, including predefined taxonomies or checklists. Hence the coverage of hazardous events may be limited. Qualitative methods, such as System-Theoretic Process Analysis, or the Functional Resonance Assessment Methodology need input from brainstorming and are not automated. Hence, these have not been discussed in detail in this article.

Risk analysis often use average values, which is sufficient for decision-making in the design phase. However, for operations real-time information is needed and the dynamics of the system need to be reflected in the models. Another challenge is related to risk evaluation and acceptable risk. These need to be defined and considered in the control system design.

A case study on a power system mode controller on board of an offshore vessel demonstrates the application of the concept of implementing risk considerations in the

control system. The controller is part of the behavioural planning and re-planning layer and decides between an economy mode or a high-power mode. Such a controller may be part of an autonomous ship. A hybrid controller is outlined that switches the available number of generators based on the risk level of the current operation, leading to more energy efficient operation. The example is not fully detailed; however, it is demonstrated and discusses how risk methods and models may be used to improve decision processes and to create risk-aware control systems.

5.1 Future work

Further work includes expanding the findings presented in this article. Implementation of risk models in the control system will provide risk aware system with enhanced decision-making capabilities. Hence, the identified opportunities need to be tested to support these claims.

6 Acknowledgements

This work has been carried out as part of the projects Unlocking the potential of autonomous systems and operations through supervisory risk control (UNLOCK) and Online risk management and risk control for autonomous ships (ORCAS). The Norwegian Research Council is acknowledged as the main sponsor of project number 274441 and 280655, respectively.

The input and comments from our colleagues in the ORCAS project, the UNLOCK project and the Department of Marine Technology at NTNU are highly appreciated.

7 References

1. Wintersberger S, Azmat M, Kummer S. Are We Ready to Ride Autonomous Vehicles? A Pilot Study on Austrian Consumers' Perspective. *Logistics* 2019; 3: 20.
2. Clarke R. Understanding the drone epidemic. *Comput Law Secur Rev* 2014; 30: 230–246.
3. Yuh J, Marani G, Blidberg DR. Applications of marine robotic vehicles. *Intell Serv Robot* 2011; 4: 221–231.
4. Vagia M, Transeth AA, Fjerdingen SA. A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied Ergonomics* 2016; 53: 190–202.
5. Ramos M, Thieme CA, Utne IB, et al. Proceedings of the 1st International Workshop on Autonomous Systems Safety. In: Ramos MA, Thieme CA, Utne IB, et al. (eds) *Proceedings of the 1st International Workshop on Autonomous Systems Safety*. Trondheim, Norway: NTNU, 2019, p. 128.
6. Utne IB, Rokseth B, Sørensen AJ, et al. Towards supervisory risk control of autonomous ships. *Reliab Eng Syst Saf*; 196. Epub ahead of print 2020. DOI: 10.1016/j.ress.2019.106757.
7. ISO. ISO 31000 Risk management - Principles and guidelines. *International Organization for Standardization* 2018; ISO 31000: 34.
8. Rausand M, Haugen S. *Risk Assessment*. Wiley, 2020. Epub ahead of print March

31, 2020. DOI: 10.1002/9781119377351.
9. Pereira AA, Binney J, Jones BH, et al. Toward risk aware mission planning for autonomous underwater vehicles. In: *IEEE International Conference on Intelligent Robots and Systems*. Piscataway, NJ, USA: IEEE, 2011, pp. 3147–3153.
10. Pereira AA, Binney J, Hollinger GA, et al. Risk-aware path planning for autonomous underwater vehicles using predictive ocean models. *J F Robot* 2013; 30: 741–762.
11. Lefebvre N, Schjølberg I, Utne IB. Integration of risk in hierarchical path planning of underwater vehicles. *IFAC-PapersOnLine* 2016; 49: 226–231.
12. Johansen TA, Perez T, Cristofaro A. Ship collision avoidance and COLREGS compliance using simulation-based control behavior selection with predictive hazard assessment. *IEEE Trans Intell Transp Syst* 2016; 17: 3407–3422.
13. Brekke EF, Wilthil EF, Eriksen BOH, et al. The Autosea project: Developing closed-loop target tracking and collision avoidance systems. *J Phys Conf Ser*; 1357. Epub ahead of print 2019. DOI: 10.1088/1742-6596/1357/1/012020.
14. Bremnes JE, Norgren P, Sorensen AJ, et al. Intelligent risk-based under-ice altitude control for autonomous underwater vehicles. *OCEANS 2019 MTS/IEEE Seattle, OCEANS 2019*. Epub ahead of print 2019. DOI: 10.23919/OCEANS40490.2019.8962532.
15. Bremnes JE, Thieme CA, Sørensen AJ, et al. A Bayesian Approach to Supervisory Risk Control of AUVs Applied to Under-Ice Operations. *Mar Technol Soc J*; 54.
16. Hobbs C. *Embedded software development for safety-critical systems*. 2nd Editio. Boca Raton, Fl, USA: CRC Press, 2019. Epub ahead of print 2019. DOI: 10.1201/b18965.
17. Hafver A, Pedersen FB. PROBABILISTIC DIGITAL TWINS - Extending digital twins for risk management.
18. Utne IB, Sørensen AJ, Schjølberg I. Risk Management of Autonomous Marine Systems and Operations. In: *Proceedings of the ASME 2017 36th International Conference on Ocean, Offshore and Arctic Engineering, OMAE 2017*. Trondheim, Norway: American Society of Mechanical Engineers, 2017, pp. 1–10.
19. Sørensen AJ. Structural issues in the design and operation of marine control systems. *Annu Rev Control* 2005; 29: 125–149.
20. Pendleton SD, Andersen H, Du X, et al. Perception, planning, control, and coordination for autonomous vehicles. *Machines*; 5. Epub ahead of print 2017. DOI: 10.3390/machines5010006.
21. Blanke M, Kinnaert M, Lunze J, et al. *Diagnosis and fault-tolerant control, third edition*. 3rd Ed. Berlin Heidelberg, Germany: Springer-Verlag GmbH, 2016. Epub ahead of print 2016. DOI: 10.1007/978-3-662-47943-8.
22. Šabanović A, Ohnishi K. *Motion Control Systems*. Singapore: John Wiley and Sons (Asia), 2011. Epub ahead of print 2011. DOI: 10.1002/9780470825754.
23. International Electrotechnical Commission (IEC). IEC 61508: Functional safety of E/E/PES safety related systems.
24. International Organization for Standardization. ISO 26262:2018 - Road vehicles – Functional safety.
25. EN. EN 50128: Railway applications - Communication, signalling, and processing systems. *Software for railway control and protection systems*; EN50128:20.
26. Bast H, Delling D, Goldberg A, et al. *Route planning in transportation networks*. 2016. Epub ahead of print 2016. DOI: 10.1007/978-3-319-49487-6_2.
27. LaValle SM. *Planning algorithms*. Cambridge University Press, 2006. Epub ahead of print 2006. DOI: 10.1017/CBO9780511546877.
28. Goebel R, Sanfelice RG, Teel AR. Introduction. In: *Hybrid Dynamical Systems*. Princeton, NJ, USA: Princeton University Press, 2012, pp. 1–24.
29. Goebel R, Sanfelice RG, Teel AR. The solution concept. In: *Hybrid Dynamical Systems*. Princeton, NJ, USA: Princeton University Press, 2012. Epub ahead of print 2012. DOI: 10.23943/princeton/9780691153896.003.0002.
30. Karaman S, Frazzoli E. Sampling-based algorithms for optimal motion planning. *Int J*

Rob Res 2011; 30: 846–894.
31. Kress-Gazit H, Fainekos GE, Pappas GJ. Temporal-logic-based reactive mission and motion planning. *IEEE Trans Robot* 2009; 25: 1370–1381.
32. Kapinski J, Deshmukh J V., Jin X, et al. Simulation-Based Approaches for Verification of Embedded Control Systems: An Overview of Traditional and Advanced Modeling, Testing, and Verification Techniques. *IEEE Control Syst* 2016; 36: 45–64.
33. Wilkie D, Van Den Berg J, Manocha D. Generalized velocity obstacles. In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*. 2009, pp. 5573–5578.
34. Ben-Ari M, Mondada F, Ben-Ari M, et al. Finite State Machines. In: Ben-Ari M, Mondada F (eds) *Elements of Robotics*. Cham, Switzerland: Springer, 2018, pp. 55–61.
35. Doshi-Velez F. The infinite partially observable Markov decision process. *Adv Neural Inf Process Syst 22 - Proc 2009 Conf* 2009; 477–485.
36. Fridman A. Mixed Markov models. *Proc Natl Acad Sci U S A* 2003; 100: 8092–8096.
37. Warren CW. Global path planning using artificial potential fields. In: *Proceedings, 1989 International Conference on Robotics and Automation*. 1989, pp. 316–321.
38. Meyer PJ, Coogan S, Arcak M. Sampled-Data Reachability Analysis Using Sensitivity and Mixed-Monotonicity. *IEEE Control Syst Lett* 2018; 2: 761–766.
39. Aurenhammer F. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Comput Surv* 1991; 23: 345–405.
40. Fossen TI. Handbook of Marine Craft Hydrodynamics and Motion Control. *Handb Mar Cr Hydrodyn Motion Control*. Epub ahead of print 2011. DOI: 10.1002/9781119994138.
41. Naus GJL, Ploeg J, Van De Molengraft MJG, et al. A model predictive control approach to design a parameterized adaptive cruise control. In: del Re L, Allgöwer F, Glielmo L, et al. (eds) *Lecture Notes in Control and Information Sciences*. London: Springer London, 2010, pp. 273–284.
42. Harnefors L, Saarakkala SE, Hinkkanen M. Speed control of electrical drives using classical control methods. *IEEE Trans Ind Appl* 2013; 49: 889–898.
43. Sciavicco L, Siciliano B. *Modelling and Control of Robot Manipulators*. Second edi. London: Springer London, 2000. Epub ahead of print October 2000. DOI: 10.1007/978-1-4471-0449-0.
44. Brodtkorb AH. *Hybrid Control of Marine Vessels*. Norwegian University of Science and Technology, 2017.
45. Bray D. *Dynamic Positioning Systems*. Ledbury, UK: Oilfield Publications Ltd., 1998.
46. Balchen JG, Jenssen NA, Mathisen E, et al. Dynamic Positioning System Based on Kalman Filtering and Optimal Control. *Model Identif Control* 1980; 1: 135–163.
47. Fossen TI, Strand JP. Passive nonlinear observer design for ships using Lyapunov methods: Full-scale experiments with a supply vessel. *Automatica* 1999; 35: 3–16.
48. Rokseth B, Utne IB, Vinnem JE. A systems approach to risk analysis of maritime operations. *Proc Inst Mech Eng Part O J Risk Reliab* 2017; 231: 53–68.
49. Kaplan S, Garrick BJ. On The Quantitative Definition of Risk. *Risk Anal* 1981; 1: 11–27.
50. Aven T. On how to define, understand and describe risk. *Reliab Eng Syst Saf* 2010; 95: 623–631.
51. Aven T. The risk concept-historical and recent development trends. *Reliab Eng Syst Saf* 2012; 99: 33–44.
52. Leveson NG, Thomas JP. *STPA handbook*. 1. Cambridge, MA, USA, 2018.
53. Jensen F V, Nielsen TD. *Bayesian networks and decision graphs. Statistics for engineering and information science*. New York, NY 10013, USA: Springer Science & Business Media, 2007.
54. Fenton N, Neil M. Risk assessment and decision analysis with bayesian networks. *Risk Assessment and Decision Analysis with Bayesian Networks* 2012; 1–494.
55. ISO/IEC. Risk management — Risk assessment techniques 31010. *Iec/Fdis* 2009;

IEC/ISO310: 1–96.
56. Guarro SB, Yau MK, Ozguner U, et al. Formal framework and models for validation and verification of software-intensive aerospace systems. In: *AIAA Information Systems-AIAA Infotech at Aerospace, 2017*. Grapevine, TX, USA: American Institute of Aeronautics and Astronautics Inc, AIAA, 2017. Epub ahead of print 2017. DOI: 10.2514/6.2017-0418.
57. Aldemir T, Guarro S, Mandelli D, et al. Probabilistic risk assessment modeling of digital instrumentation and control systems using two dynamic methodologies. *Reliab Eng Syst Saf* 2010; 95: 1011–1039.
58. International Electrotechnical Commission. IEC 61165: Application of Markov techniques. *IEC Standards Online*; IEC61165:2.
59. Rausand M. *Reliability of Safety-Critical Systems: Theory and Applications*. 1st Ed. Hoboken, New Jersey, USA: Wiley & Sons Inc., 2014. Epub ahead of print 2014. DOI: 10.1002/9781118776353.
60. NEK-EN IEC. NEK-EN IEC 62551: Analysis techniques for dependability - Petri net techniques. 2012; NEK-EN IEC: 1–70.
61. Øien K. Remote operation in environmentally sensitive areas: Development of early warning indicators. *J Risk Res* 2013; 16: 323–336.
62. Thieme CA, Utne IB. A risk model for autonomous marine systems and operation focusing on human-autonomy collaboration. *Proc Inst Mech Eng Part O J Risk Reliab*; 231. Epub ahead of print July 2017. DOI: 10.1177/1748006X17709377.
63. Andrews JD, Prescott DR, Remenyte-Prescott R. A systems reliability approach to decision making in autonomous multi-platform systems operating a phased mission. In: *Proceedings - Annual Reliability and Maintainability Symposium*. Piscataway, NJ, USA: IEEE, 2008. Epub ahead of print 2008. DOI: 10.1109/RAMS.2008.4925761.
64. Prescott DR, Andrews JD, Downes CG. Multiplatform phased mission reliability modelling for mission planning. *Proc Inst Mech Eng Part O J Risk Reliab* 2009; 223: 27–39.
65. Remenyte-Prescott R, Andrews JD, Chung PWH. An efficient phased mission reliability analysis for autonomous vehicles. *Reliab Eng Syst Saf* 2010; 95: 226–235.
66. Guarro S, Yau M. Dynamic Flowgraph Methodology (DFM) Modeling of nuclear and advanced technology system risk and reliability scenarios. *Adv Concepts Nucl Energy Risk Assess Manag* 2018; 425: 353–426.
67. Guarro SB, Yau MK, Ozguner U, et al. Formal framework and models for validation and verification of software-intensive aerospace systems. *AIAA Inf Syst Infotech Aerospace, 2017* 2017; 1–10.
68. Swuste P, Theunissen J, Schmitz P, et al. Process safety indicators, a review of literature. *J Loss Prev Process Ind* 2016; 40: 162–173.
69. Kongsberg. *Kongsberg K-Pos DP (OS) Dynamic Positioning System - Operator Manual*. Kongsberg, Norway: Kongsberg Maritime, 2007.
70. Thorat L, Skjetne R. Load-dependent start-stop of gensets modeled as a hybrid dynamical system. *IFAC-PapersOnLine* 2017; 50: 9321–9328.
71. Rokseth B, Utne IB, Vinnem JE. Deriving verification objectives and scenarios for maritime systems using the systems-theoretic process analysis. *Reliab Eng Syst Saf* 2018; 169: 18–31.
72. Øien K. Risk indicators as a tool for risk control. *Reliab Eng Syst Saf* 2001; 74: 129–145.
73. Haddara MR, Guedes Soares C. Wind loads on marine structures. *Mar Struct* 1999; 12: 199–209.
74. Brodtkorb AH, Nielsen UD, Sørensen AJ. Sea state estimation using vessel response in dynamic positioning. *Appl Ocean Res* 2018; 70: 76–86.
75. Mosleh A. PRA: A Perspective on strengths, current Limitations, and possible improvements. *Nucl Eng Technol* 2014; 46: 1–10.