Development of Safe Autonomous Mobile Service Robots using an Active Integrated Approach

Philipp Ertle, Holger Voos
ZAFH Autonome Mobile Serviceroboter at Hochschule Ravenburg-Weingarten, Germany

Dirk Söffker
Lehrstuhl Steuerung, Regelung und Systemdynamik at Universität Duisburg-Essen, Germany

Abstract

Mobile service robots are expected to provide services in various domains of life. Herein, the main challenge for the robot is the execution of complex tasks within an unstructured dynamic environment. In order to achieve the necessary highly flexible behavior, mobile service robots must have a high degree of autonomy. However, although much effort is spent on the investigation and realization of autonomy since decades, it is mostly neglected, that autonomous robots are also causing new types of safety problems. It seems impossible to solve these problems with traditional passive safety approaches. Therefore, new methods are introduced in order to check safety during operation time.

1 Introduction

Mobile service robots are expected to provide services in various domains of life. Herein, the main challenge for the robot is the execution of complex tasks within an unstructured dynamic environment while collaborating with human users in a natural and intuitive way. In order to achieve the necessary highly flexible behavior, mobile service robots must have a high degree of autonomy. However, although much effort is spent on the investigation and realization of autonomy since decades, it is mostly neglected, that autonomous robots are also causing new types of safety problems. While a safe stationary robotic manipulator can be obtained by avoiding any collisions with users or the environment (e.g. with the help of a separating safety cage), this is no longer the case for autonomous mobile robots. Here, touching human persons might be even necessary on the one side, while also the pure decisions of the robot, e.g. delivering a requested medicine, could also cause safety-critical situations on the other side.

In this contribution, the main focus is related to the development on a more comprehensive realizing a suitable and adequate level of safety. The contribution is organized as follows: First, autonomous robots are identified as safety-critical systems. Here safety is interpreted as a state of the robot-environment-interaction. Any method to check safety during operation requires suitable measures. So called active measures are introduced, whereby safety is mainly influenced by the safety of the state of operation and the safety of the behavior of the robot as described in the following part. In the experimental section, it is described how this can be practically realized on a real robot or within a simulation environment.

2 Background and Related Work

A mobile robot which is intended to move autonomously in a dynamic environment will be considered. The robot is interacting with objects and human persons over a longer period of time in order to solve given tasks. It is often neglected that a higher degree of autonomy also results in higher safety requirements, especially if these autonomous robots have to interact closely with human users.

2.1 Robots are safety-critical systems

Safety-critical systems are those systems whose failure could result in loss of life, significant property damage or damage of the environment. Therefore, a system is called safe if it can be ensured that risks are kept at an acceptable level (IEC61508). Herein, risk is the possibility of injury, loss or environment incident created by a hazard, while the significance or level of the risk is generally determined by the probability of an unwanted incident and the severity of the consequences. Mobile robots have a considerable mass and kinetic energy during operation, they share the same environment with human users and autonomous mobile robots are in addition even enabled to come to own decisions. Therefore, mobile robots and especially those with a higher degree of autonomy are clearly safety-critical systems. Some first contributions that are especially focused on the development of safe autonomous mobile robots can already be found in the literature, see e.g. [5], [6], [8]. However, most of these papers are mainly focusing on single aspects of this special safety-related problem like software verification, special redundant hardware systems, special software development processes etc.

2.2 Active Safety Measures

In this contribution, the main focus is related to the development on a more comprehensive realizing a suitable and adequate level of safety. Nevertheless, the development of safe autonomous mobile robots should first of all also include all well known development processes and measures for safety-critical systems in general. Most of these measures however are passive, i.e. they are applied during the development phase in order to achieve a safe system. This includes safety-related analysis, a suitable design of the mechanical and electrical/electronic parts, programming guidelines, verification and validation of the software, robust control etc. However, it is typical for all passive safety measures that possible hazards and failures must be foreseen and included in the safety analysis process during development. It is turned out, that safety is always related to the overall state of the robot and the environment, also including human persons, and it is nearly impossible to foresee all possible interactions between robot and environment at design time. Therefore, it seems to be more promising to add active measures to ensure safety in the case of autonomous mobile robots, see e.g. [3]. These measures are active during the operation of the mobile robot in order to ensure safety. Herein, the risk of failure is mainly effected by the knowledge/perception of the state of the robot, the safe behavior, and finally the correctness of the application. The last aspect, i.e. correct application of the robot, means that the robot is applied in a specific task as originally intended, i.e. not operated outdoor if originally intended for indoor application. However, in most cases of practical interest, this can already be guaranteed by the human user and hence the main focus of this contribution is on the two remaining aspects.

2.3 System Architecture

It is recognized that efficient robot control architectures are combining reactive control and deliberation to a hybrid deliberative/reactive architecture [3], [6], this architecture is also adopted here for robot control. In a hybrid deliberative/reactive architecture, complex and long-term planning tasks based on the world model are solved on a deliberative layer, the generated plans are executed in a reactive fashion by the activation of a set of suitable behaviors in the underlying reactive layer [3]. These layers are extended in this work in order to achieve the active safety concept. On the deliberative layer, the planning and decision making procedures also take the result of the risk assessment of the current and future predicted situations into account. The generated plans therefore must lead to situations whose risk is always kept below the accepted level. On the reactive layer, the risk assessment of the situations must also be considered during the activation of suitable behaviors. This reactive behavior is solved by the help of a model-predictive control approach [1]. Hereby, the execution of the plans is formulated as an optimization problem while the safety aspects as a result of the risk assessment are forming constraints. An on-line optimization finally leads to the optimal action while keeping the safety constraints. The overall structure of this architecture (see also [7]) is finally shown in the following **Figure 1**.

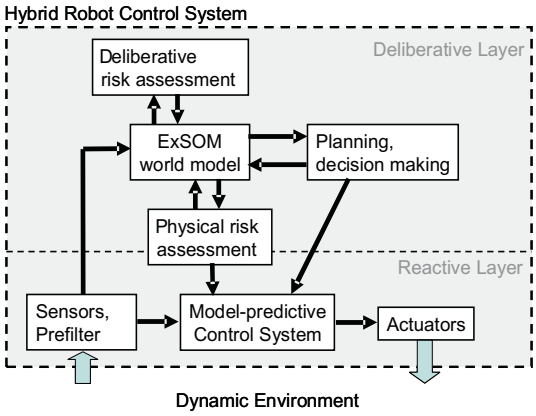

Figure 1: Overall system architecture for the proposed safety concept for autonomous mobile robots.

3 Active Measure Describing Safe Situations

The safe state of operation includes reliable functioning of the sensors, the computer systems as well as safe functioning of the actuators. Here, several measures to increase safety with regard to this point are applied. However, it is difficult to identify the safety-critical components for autonomous mobile robots because the occurring of hazards often depend strongly on environmental constellations. Many cause-effect-relationships cannot be foreseen. For example, even a wheel-driven small robot can become dangerous: It may cause harm to humans if, for instance the robot is falling down the stairs.

The definition of mechanisms to avert these dangerous situations seems to be very difficult. A more promising concept at least to increase safety is to formulate and observe operating conditions for actuating components, such as servos, arms, propellers or driving motors. The violation of the conditions initiates movement limitations or restrictions, for example the limitation of driving speed, acceleration or angle. The conditions are defined with focus on functionality of the single actuating hardware without considering any high level capabilities. These conditions shall exclude these states of operation which cannot be detected by a system itself (system crash) and which are a priori known as unwanted. This method is generally adopted from traditional safety assuring techniques: A power supply of a rotating production machine's motor is disconnected when the chassis is open. The condition 'chassis open' is realized by a contact sensor, the limitation of the actuator is realized by a disconnection of the motors power supply by relays, for instance. If the movement of actuators are those that transfers potential dangerous energy to

the system, it is worth to have an closer look at them: The actuators could provoke hazards but they are controlled by systems in order to produce a defined behavior. If there are rising hazards, these can be caused by failure of the superordinate control layer or its superordinate layers.
A simple concept to avoid undefined and therefore potential dangerous states of an actuator, is to await a certain 'liveliness' of the direct superordinate control component. The superordinate control level again awaits defined 'liveliness' from its next superordinate component(s). The surveillance of the 'liveliness' can be simply realized with the help of the so-called 'watchdog'-method, for instance. The superordinate control of a component generates a periodic signal ('heart beat' signal). A redundant safety component observes this signal and limits the corresponding actuator, respectively. The surveillance of a heart beat does not allow a detailed failure diagnosis of the superordinate component but allows at least a conclusion on its maximum execution performance. A strongly reduced execution performance of the superordinate component leads to signal irregularities or even to reduced frequency, which requires a reaction, for example the reduction of the acting speed. The breakdown of the heart beat requires the disconnection of the superordinate control component from the actuator or even taking-over of control by a redundant system. An example of such an system is illustrated in **Figure 2**.

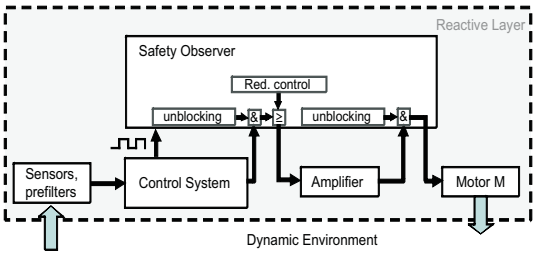

Figure 2: Example of an integrated component safety redundancy.

The *controller* has to provide an adequate 'heart beat' signal to the *safety observer*. The frequency is adopted to the kind of motor usage (Motor *M*) or may also vary in order to adapt to a changing motor usage. The safety observer can disconnect the motor directly after or prior to the amplifier. Additionally, a logic is included to realize an override of the controller signal. This basically enables the redundant functionality of the safety observer.
The safety observer can provide additional functionality when the 'heart beat' fails. A control redundancy is realized. First of all, additional hard- or software or mostly has a negative effect on the overall system availability. However, if undefined states can be recognized and suitable countermeasures are activated, the safety with respect to considered component can be increased. If dangerous states are avoided, which are awaited to occur with a certain probability, the saftey is increased. However, the failsafe functionality must be realized very carefully, especially if it cannot be realized by simply switching off the motors, for instance.
One further advantage of the proposed approach, which is transferred from other safety-critical applications, is the realization of functionality with different tool chains (programming languages, compilers etc.) for the control system and the safety system. This reduces the probability of systematic compiler errors.

4 Active Measures for a Safe Behavior

The second aspect, as already mentioned, deals with the safe behavior of the robot, meaning that the robot is fulfilling its tasks and reaching its goals as planned, while keeping the risk below an accepted level. However, this risk does not only depend on the robot itself but clearly also on the dynamic environment. In this contribution, the two types of risks - the deliberative or physical risks (see also [7]) as shown in Figure 1 - are subsumed as behavioral risks, it is focused the feasibility of the on-line risk assessment in principle. The risk assessment is a basic feature of the proposed approach and is a dynamic process that takes the overall state of the robot and the environment, as well as the interaction between robot and environment into account. A suitable model of the robot-environment-interaction, i.e. a suitable world model is required. For that purpose, an extended version of the cognitive framework called 'Situation-Operator-Model' (exSOM) [4], [7] is applied. Core of this exSOM-approach is the assumption, that the real world is modeled as a sequence of discrete-time situations and operators. Situations describe the current state of the world comprising the robot and the environment as perceived by the robot. Each situation is that extract and internal representation of the real world which is of current interest for the robot. The situation is described by a suitable set of so-called characteristics. The characteristics could be numerical, boolean or linguistic variables, but also more complex data structures. They include information which is perceived from the external world, i.e. the environment, with the help of suitable sensors and signal processing, and also information of the current physical and cognitive state of the robot, e.g. the current active goals and plans.
Changes in the real world are represented by so-called operators. The operators are linking the situations in a way that the situation snapshot at any point in time is transferred by an operator to the following situation [4]. These changes can be caused by the robot itself. With the help of a knowledge database that consists of situation-operator-situation sets, a planning process can be realized in order to change current situation to a desired goal situation. This intentional acting is also called behavior. A 'safe behavior' is to ensure that risk are lower than acceptable level. In the discrete-time description of the exSOM approach, safety is a property of any situation $s(k)$ at any time step

k, and is described by a risk value which is assigned to the situation. This risk value clearly depends on the values of the current set of characteristics $\{c_i(k)\}$ and thus also depends on both robot and environment. A risk value of a situation describes the probability of an unwanted incident in the future and the severity of the consequences, given the current set of characteristics. In this approach, a risk is only assigned to situations and not directly to operators. Therefore, operators transfer a current situation with current risk into the next situation with respective current risk. The risk information is encoded with risk values. These combine both, accident severity and probability ($Risk = S_{Acc} \cdot P_{Acc}$). Therefore, accidental events ($P_{Acc} = 1$) or hazards ($0 \geq P_{Acc} > 1$) can be described. The severity is assumed to be $0 \geq S_{Acc} \geq 1$ whereby $S_{Acc} = 1$ denotes worst case accident.

Basically, future risks of a mobile robot in a dynamic environment can hardly be foreseen and modeled completely beforehand. [9],[2]. Therefore, a framework is developed to integrate and utilize dynamic risk knowledge in order to enable a on-line risk assessment. The risk knowledge can be included manually during a very early development phase. The basic idea of this risk assessment is to transfer the safety expert knowledge into the system instead of implementing specific safety assurance mechanisms. The knowledge can also be transferred in a highly abstracted manner because the SOM approach and therefore the exSOM approach is able to deal with very high abstractions. Furthermore, the risk causes are formulated as basic principles which shall be valid for all future situations. For example, heated plastic materials generate toxic vapors or start to burn. So, the principle that the combination of intense heat and plastic materials is dangerous is not true in all cases, but not wrong in 'conservative' safety concept. If an exception is known it could be added to the knowledge, in all other cases the system would avoid to approach '*plastic*' to '*heat sources*' generally. Therefore, the principles are valid premises which can be used with (valid) observations for deductive conclusion. If the principle '*plastic object close to a heat source is hazardous*' exists and there would be the observation '*plastic object is close to heat source*', the conclusion '*hazard occurs*' would be correct.

This first part of a safety principle represents a conditional part. The goal of this conditional part is to recognize the occurring of a hazard in principle, for example by detecting a heat source and a plastic object in the same situation (temporal relation is given by the exSOM approach). The related computational part contains the instructions for the determination of the respective risk. Therefore the computational part describes which information is needed and how this information is processed to generate a quantitative risk value. In many cases, the geometrical relations between involved objects are important, for instance the relative speed and the distance are important when computing risks of kinetic energies or collision risks, the temperature is essential when computing inflammation risks. Initially, the computation parts of each safety principle have to be constructed. However, it is often difficult to formulate the dependencies of risks in a specific context. First of all, a transfer from a understandable qualitative to quantitative description can be realized with the help of fuzzy methods (see [7]). Furthermore, it is suggested to interpolate between known data, for instance it is assumed that a plastic object with no distance to a hot stove is melting and therefore producing toxic vapors. Further on, it is assumed that the same plastic object is safe in a distance of $50cm$. This information can be interpolated. The resulting risk function is a linear function depending on the distance of the to objects heat source and plastic object.
One remaining question considers the completeness of the safety-related knowledge with respect to these two parts of the risk assessment system. It is assumed that a first definition of the principles is based upon the expert knowledge of the engineers during the design phase. Since this knowledge usually will be incomplete and not cover all possible situations with a considerable risk, learning will be applied in the future realization. Here, reinforcement learning could be a suitable approach, where a human supervisor enters his own risk assessment results with the help of a suitable man-machine-interface during robot actions in simulations or experimental runs.

5 Realization Examples

5.1 Observing Safe Situation

The state of operation shall be surveyed by a very simple concept to avoid undefined and therefore potential dangerous states of an actuator. A certain 'liveliness' of the direct superordinate control components is awaited. Therefore a periodic 'on-off' signal, a so-called 'heart beat' signal is generated by the superordinate component and examined by a safety observer. To give an example, the differential drive of a 'Pioneer 3D' robot platform is regarded as a safety-critical component.
That illustrated concept comprises a simple microcontroller board (called 'safety board') that has its own power supply consisting of batteries. The safety board is connected to the main computer system and is able to disconnect the power supply of the electrical drives for motion generation, by a suitable electronic circuit. The main robot control software is extended in a way that a signal is generated in a regular fashion ('heart beat') which is connected with the interrupt input of the microcontroller. The signal generation extension itself again includes further observing mechanisms to observe further 'liveliness' conditions of the next superordinate level. In this example the response time of the collision avoidance is checked. If the response time is dropping, the frequency of the 'heart beat' signal is lowered, too. Furthermore, the frequency can also be reduced when the driving motors are not used in order to save processing power. If the main computer fails or has a deadlock, the heart beat signal will be no longer generated.

The safety board reacts to changes of the 'heart beat signal'. The reactions are shown in **Table 1**.
In case of a missing 'heart beat' signal the safety board is disconnecting the drives like an emergency stop. If the 'liveliness' condition is fulfilled again, the regular operation can take place again (unblocking of analog circuit after $< 1s$, software after $10s$). In case of exceeding the maximum frequency a 'repair mode' is activated. Such high frequency is caused by contact problems which then have to be repaired. After repairing the system, the circuit could be reset to normal operation.
The realization of safety-critical systems with programmable devices includes the risk that also software faults could occur. Therefore, a additional analog circuit is implemented. The analog circuit consists of a low pass filter and two voltage comparators. The mean value of the heart beat signal is expected to remain between an upper and a lower voltage threshold. While this threshold condition is fulfilled, the motor's power lines are kept connected (via relays). So, both analogue and digital circuit have to 'unblock' in order to activate the 'normally open contact' relays.
In addition, the safety board also includes some sensors like acceleration sensors to detect irregular and fast motion of the robot or attitude sensors to detect if the robot tends to fall. In all these cases, the safety board is switching of the drives, too.

HB signal	Reaction	Regular operation
none	disconnection of the motors	after <1s/10s
50Hz...95Hz	limit. to 10% ...70% speed	after 10s
~100Hz	100% speed, normal operation	-
>105Hz	Repair mode, limit. to 20% speed	after repair

Table 1: Reactions of the safety board under consideration of the 'heart beat' signal.

5.2 Observing Safe Behavior

The safe behavior of the robot means that the robot is fulfilling its tasks and reaching its goals as planned, while keeping the risk below an accepted level. However, this risk does strongly depend on the dynamic environment. Therefore, risk assessment is a basic feature of the proposed approach and must be a dynamic process that takes the environment as well as the interaction between robot and environment into account. In a small demonstration a possible solution for a risk assessment is focused to enable evaluation of the interaction between a robot and its environment. This demonstration takes place in a small virtual grid world. This grid world is seen as a internal representation of a simplified real world. It is assumed that object recognition techniques and technical cognitive system (exSOM) are able to generate such internal representation dynamically. Furthermore, uncertainties in modeling and object recognition are not considered. The mentioned grid world can be seen as world model or world map of the current environmental world. When the interaction between a robot and its environment is considered, the investigation of temporal and spatial relations of objects is important in order to observe the interplay of the objects handled by the robot and of other surrounding objects. Therefore the grid world contains scene objects which are placed at corresponding map position. These objects are represented by a single character at respective position in the grid world. The robot is represented by a rectangular shape and can be moved in horizontal and vertical direction. The robot can grip and transport one object. The gripped object is represented by a character inside the robot's shape. The scene objects are known objects which means that they can be recognized by the object recognition and further object information can be retrieved from an object information database. Such additional information can be attributes of objects, e.g. *'plastic'*, *'liquid container'*. This additional information is assumed to be integrated manually prior to operating time. It is expected that such information could also be obtained by learning mechanisms.
The robot can change the spatial relation of the scene objects by gripping and transporting of scene objects. The change of the spatial relation can provoke risks. These risks are an essential aspect when safe behavior should be assured. For example, the robot comes to close to a fireplace and starts burning, the robot provokes fire by putting a plastic object on a hot stove in accordance to its task '*bring dishes to kitchen*' or the robot is instructed to fetch the blind user's medicine and delivers the wrong one. For a first step, it is important to exclude such situations. Therefore, a set of safety principles is implemented. The principles are formulated in a way such that they are abstracted from specific hazard cause in order to generate generalized principles. This abstraction is achieved by formulating principles with respect to objects or their attributes, for instance, the spatial combination of '*plastic*' objects and '*extreme heat*' is dangerous. If this conditional part of a safety principle is fulfilled, the respective computational part determines the risk value. Therefore a risk function is specified. This risk function can be any function that is suitable to map the situation to a risk value. In the example step functions, linear functions or combinations of linear functions are used. Function arguments can also be information about the internal state such as the speed of the robot or the distance of the objects or any exSOM characteristic. It is suggested to interpolate known risk values (most often these are the extremes) in order to specify these functions, for example with linear, polynomial functions or even with neural networks. In the following, three examples are explained in more detail in order to illustrate this approach. The numerical risk values are exemplary.
The used scene object are abbreviated by capital characters and furnished with attribute as described in **Table 2**.

5.2.1 Decoding of the Grid World Distance

The grid world that is used to simulate the dynamics of object interplay is kept as simple as possible. The limitation to geometrical measures between objects is insufficient.

Therefore, some of the geometrical distance ranges represent other properties. The relative distance of two objects with $dist = 1$ is defined as '*these objects are combined with each other*'. A symbolic task procedure description can be used for risk assessment, for example '*give*' '*Object A*' to '*Object B*' or '*put*' '*Object A*' into/on top of '*Object B*' etc.

The remaining distance ranges are geometrical measures and represent $1 > dist \geq 2$ '*to be in contact to*', $2 > dist \geq 3$ '*to be very close to*', $3 > dist \geq 4$ '*to be close to*', $4 > dist \geq 5$ '*to be in the range of*'.

Abbr.	Object	Attributes	
H	Human		
C	Coffee-Bowl	A:plastic	A:hot liquid
M	Drug	A:chemical	
S	Stove	A:heat	

Table 2: The used scene objects, their abbreviation and their assigned attributes.

5.2.2 Modeling danger of chemical intoxication

An chemical intoxication can take place if a chemical product is given to a human. This can happen when there is a error with respect to correct recognition of the chemical itself or the dosage. Hazards can occur, if the robot has to deliver chemicals to a human user, e.g. if the robot is instructed by a human user to dissolve a medicine in a glass of water. In doing so the mishap may happen that more than the desired number of pills is dropped in the glass of water. Further on, a blind user instructs the robot to fetch essential medicine and the robot delivers the wrong medicine because the object recognition was disturbed. In order to increase safety of service robots in such context these tasks have to be specified more precisely or they are seen as being too risky in general, as described in the following example.

In the grid world the handing over of one object to another is assumed when the distance between these two objects is $dist = 1$ grid field.

Object 1 (Robot/carried)	Object 2 (in environment)	Principle	Input	Risk1	@Inp1		
A:chemical	H	StepMFunction	distance	1	1		

Table 3: Relation of of objects with the attribute chemical when applied to humans formulated with the help of a step function.

If the robot is intends to '*give*' an object with the attribute '*chemical*' in to the object '*human*' ($dist_{A:chemical,human} = 1$), a risk $risk_{A:chemical,human} = 1$ results in accordance the safety principle realized with a step function as shown in **Table 3.**

This rule is applied to a grid world example. If the robot carries a chemical object, for example a '*medicine M*' (definition in accordance to Table 2) and would give it to the object '*human H*' (distance of 1 grid world field) this will result the risk value of '1', what is show in **Figure 3** with the help of a colored risk values (color gradient from red for $risk = 1$ over orange, yellow, gray to white for $risk = 0$ are used).

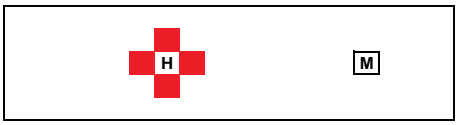

Figure 3: Risk value depicted as color code for respective robot action. Distance of '1' is defined as '*give*' '*medicine*' to '*human*'.

5.2.3 Modeling danger of burning/melting plastic

In a further scenario it is assumed, that a robot is instructed to bring the dishes to the kitchen sink. In order to perform this task the robot grips the plastic bowl and puts it on the stove under assumption that this is an optimal surface for depositing the dishes. The stove is still hot and the plastic bowl would start melting or even burning. The approaching of objects with the attribute '*plastic*' ('$A : plastic$') to objects with the attribute '*heat*' ('$A : heat$') is dangerous in principle, when the robot is not equipped with a temperature sensor in order to retrieve additional information. A simplified safety principle could be modeled as described in **Table 4.**

Object 1 (Robot/carried)	Object 2 (in environment)	Principle	Input	Risk1	@Inp1	Risk2	@Inp2
A:plastic	A:heat	LinearFunction	distance	1	1	0	3

Table 4: Relation of of objects with the attribute plastic when approached to heat sources with the help of a linear function.

The risk computation part is realized with a linear function. It is assumed that the direct contact is very risky ($dist_{A:plastic,A:heat} = 1 \rightarrow risk_{A:plastic,A:heat} = 1$). If the distance between plastic object and heat source is big enough, there is no risk anymore ($dis_{A:plastic,A:heat} = 3 \rightarrow risk_{A:plastic,A:heat} = 0$). The interpolation between these two values is realized by applying the straight-line equation (linear function). Instead of linear functions also polynomials or neural networks can be used. For respective interpolation additional parameters could be necessary. When the robot carries a '*coffee cup C*', which is specified as a plastic object (in accordance to Table 2), the approaching the '*kitchen stove S*' will result a certain risk in dependence to the distance of these two objects as shown in **Figure 4**.

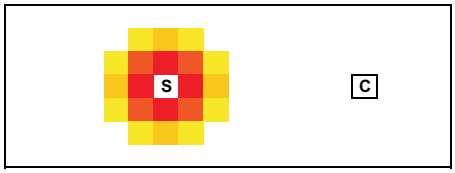

Figure 4: The approaching of '*plastic*' and '*heat*' increases risk. Putting the plastic object on top of the heat source ($dist = 1$) is very dangerous.

5.2.4 Modeling danger of scalding

When a robot carries a cup with hot tea, a collision with a human is more dangerous because there is the additional risk of scalding. It is assumed that the risk of collision depends on the relative distance and to the speed of the robot. The risk of scalding is considered insofar, that the collision risk while carrying '*hot liquids*' is higher than the simple collision risk. Therefore a safety principle is defined by combining four risk functions.

Object 1 (Robot/carried)	Object 2 (in environment)	Principle	Inp	Risk1	@Inp1	Risk2	@Inp2
A:hot liquid	H	StepFunction	distance	0	1.01	1	
*		LinearFunction	distance	1	2	0	5
*		LinearFunction	speed	1	1	0	0.3
+		StepFunction	distance	1	1.01	0	

Table 5: Combined risk function to model the increased scalding danger by fast movement while transporting '*hot liquid*'. Giving ($dist = 1$) '*hot liquid*' (cup of coffee) to a human is modeled as not dangerous.

The first risk function is a mask function in order to define the geometric distance considering risks. The second function is a linear interpolation in dependence of the relative distance. The third risk function, which is an interpolation in dependence of the speed, is multiplied with the first and second risk function. Hence, a 2D risk map is generated with respect to geometrical dependencies. The logical risk of '*giving*' the object to human is realized with fourth function, which is added as inverted mask function as shown in **Table 5**.
When the robot is carrying a coffee cup which basically can contain hot liquid (in accordance to Table 2), the risk is assumed to depend on the speed of the robot and the relative distance to humans. When the robot is driving with full speed while carrying a coffee cup containing hot liquid, the risk of scalding rises when the robot is approaching a human (see **Figure 5** upper part). When the speed is lowered to half speed, the risk is lowered in principle (see Figure 5 lower part). From that follows, the approaching to a human while transporting hot liquids can only be allowed when the speed is reduced adequately. The intention of handing over of the cup of coffee to a human is defined as not dangerous.

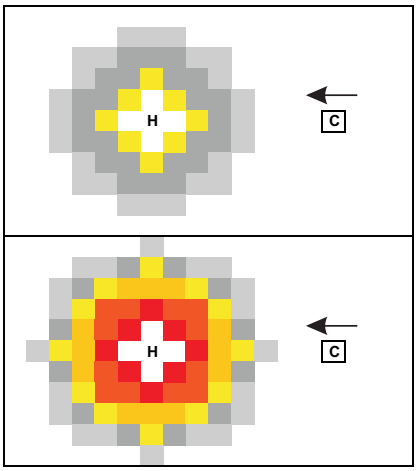

Figure 5: Risk value depicted as color code for respective robot position. Risk in relation of distance and speed (without consideration of moving direction), when approaching to a human, while robot is carrying a cup of coffee. Resulting risks are shown, while robot is driving half speed (top) and full speed (bottom). The intention to give the cup of coffee to a human is modeled as not dangerous ($dist = 1$).

6 Conclusion and Outlook

Flexible and intelligent behavior however can only be achieved by mobile service robots with a high degree of autonomy. Thus, a strict and static safety strategy would exclude the application of service robots quite fundamentally. This contribution is addressing such autonomy dependent problems. The safe state of operation and the safe behavior are identified as fundamental aspects with regard to safety of service robots. The capability to change behavior as well as internal functional structures during the operating time implies the need to observe the robot itself continuously with the help of suitable active measures. In this contribution is shown how safety components can be integrated in a hybrid robot architecture and how their active measures are generated and used within that structure. The first safety component is designed to take into account the safe state of operation of an actuator. In this connection, a low level safety component is introduced in order to exclude potential dangerous operation modes. In addition, a component is outlined which enables to survey the behavior at a planning stage. Hence, a risk assessment method is introduced in order to reflect risks that arise when acting in dynamic environments.

Initially, the applied safety principles are kept very simple, in order to show primarily the functionality of the presented method. Surely, there will be various problems when extending this method to further real world problems. This will be addressed in future research work.

References

[1] F. H. Allgöwer, Ed., *Nonlinear model predictive control*, ser. Progress in systems and control theory ; 26. Basel: Birkhäuser, 2000.

[2] P. Ertle and D. Söffker, "Towards risk analysis to enable safe service robotics," in *Interface and Interaction Design for Learning and Simulation Environments*, ser. DAAD-German Summer Academy at the University Duisburg-Essen; revised contributions, N. Baloian, W. Luther, D. Söffker, and Y. Urano, Eds. Berlin: Logos-Verlag, 2010, pp. 33–35.

[3] R. R. Murphy, *Introduction to AI Robotics*, 1st ed. Cambridge, MA, USA: MIT Press, 2000.

[4] D. Söffker, "Interaction of intelligent and autonomous systems - part i: Qualitative structuring of interactions." in *MCMDS - Mathematical and Computer Modelling of Dynamical Systems*, ser. 14, vol. 4, 2008, p. 303–339.

[5] I. Sommerville, D. Seward, R. Morrey, and S. Quayle, "Safe systems architecture for autonomous robots." in *Proc. of the 5th Safety- critical Systems Symposium*, Brighton, UK, 1997.

[6] H. Voos and N. Hochgeschwender, "Verification of autonomous UAV systems: A perspective." in *Proc. of IASTED Symposium on Robotic Applications RA 2007*, Würzburg, Germany, 2007.

[7] H. Voos and P. Ertle, "Online risk assessment for safe autonomous mobile robots - a perspective," in *7th Workshop on Advanced Control and Diagnosis*, Zielona Góra, PL, 2009.

[8] A. Wardzinski, "Dynamic risk assessment in autonomous vehicles motion planning," in *Information Technology, 2008. IT 2008. 1st International Conference on*, May 2008, pp. 1–4.

[9] A. Wardziński, "Safety assurance strategies for autonomous vehicles," in *SAFECOMP '08: Proceedings of the 27th international conference on Computer Safety, Reliability, and Security*. Berlin, Heidelberg: Springer-Verlag, 2008, p. 277–290.