Conceptual Design of a Dynamic Risk-Assessment Server for Autonomous Robots.

Philipp Ertle
University Duisburg-Essen, Chair of Dynamics and Control (SRS), 47057 Duisburg, Germany

Michel Tokic, Tobias Bystricky, Marius Ebel
University of Applied Sciences Ravensburg-Weingarten, Institute of Applied Research, 88250 Weingarten, Germany

Holger Voos
University of Luxembourg, Faculté des Sciences, de la Technologie et de la Communication, L-1359 Luxembourg

Dirk Söffker
University Duisburg-Essen, Chair of Dynamics and Control (SRS), 47057 Duisburg, Germany

Abstract

Future autonomous service robots are intended to operate in open and complex environments. This in turn implies complications ensuring safe operation. The tenor of few available investigations is the need for dynamically assessing operational risks. Furthermore, there is a new kind of hazards being implicated by the robot's capability to manipulate the environment: Hazardous environmental object interactions. Therefore, the realization of the *Dynamic Risk-Assessment* approach with special scope on object-interaction risks is addressed in this paper. A server-based architecture is proposed facilitating a feasible integration into robotic systems and realization of software and hardware redundancy as well.
Area of conference topics (keywords): Area 6 - Safety (service robots, safety, dynamic risk assessment, safety principle)

1 Introduction

The development of service robots is a steadily ongoing process with the goal that *Next-Generation Robots* (NGR) will be *capable of performing such tasks as house cleaning, security, nursing, life-support, and entertainment - all functions to be performed in co-existence with humans in businesses and homes* [1]. The more extensive the capabilities and the intelligence of such robots becomes, the more important the consideration of their ethical and moral integrity becomes as well. Especially with regard to human's health and life, the safety of such systems is one of the most important topics, which has to keep pace with the achievements enabling robots to perform more and more complex tasks, following the tenor of Jonas [2]: *Act so that the effects of your action are compatible with the permanence of genuine human life* .

It is widely accepted that robots, operating in complex human-like and typically unstructured environments, need autonomous decision-making capabilities for adequately considering current environmental conditions. Intelligence of such autonomous robots, understanding intelligence as any kind of generation, refinement or transfer of knowledge about their environment, plays a central role as well, because of *the practically impossible problem of pre-identifying the full range of kinds of situations robots [...] will get into during normal interaction with their environments* [3]. Consequently, safety assurance efforts must be extended insofar that the robot's changing basis of decision-making ex post - after the design stage - is adequately considered. Hence, in [4] the *Dynamic Risk-Assessment* (DRA) approach was introduced providing risk information to the autonomous system's decision-making process. In [5] a fault tree-based risk model is applied for deriving risk information. It was also presented in former research work how risk models can be generalized with the help of so-called *safety principles* [6, 7].

In this contribution a DRA server architecture is proposed and realized. Generalized safety principles are utilized for generating situational risk information. A small experiment finally evaluates the proper operation.

2 Methodology

In the sequel the fundamentals are described. First, the understanding of how dynamic processes can be internally represented in a robotic system is outlined. Furthermore, it is explained how such internal representation is extended with risk information.

2.1 Describing Dynamic Processes

There are various robotic architectures and description languages available. The presented *Situation-Operator Model* (SOM) approach can be used as a meta-model for un-

derstanding the robot's internal representation of dynamic processes, especially taking the interaction of a system with its environment into account. The SOM approach was utilized as well for modeling control problems [8], for realizing autonomous cognitive robots [9], cognitive (hierarchical) architectures, cognitive functions such as learning, planning [10], perception and so forth.

The elements of the SOM notation are in general situations, characteristics, operators and relations (**Figure 1**). The situations (gray ellipses) are time-fixed and an event-discrete description of dynamic processes. Changes within the considered system are denoted as operators. Changes in the process result in situation-operator chains. The situation itself consists of characteristics c_i (black dots) and relations r_i. The relation represents an inner structure of the situation, which allows the linking of characteristics to each other through arbitrary functions [10]. Basically, relations are from the same quality as operators (both white circles). The characteristics can be measured physical values or (by relations) abstracted information, such as fused sensory information, or even be recognized environmental objects.

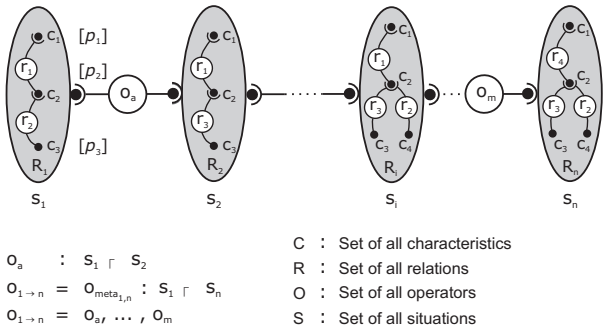

Figure 1: Sequence of operators changing the situations (arbitrary example) [8].

2.2 Modeling Context-Related Risks

In [5] a fault tree-based risk model is described, being static during the operating time. Risks might be very specific due to different operating modes and contexts of the robot. The variety of operating modes (capabilities), and as well the contexts a NGR can get into, need for a set of risk models. Robots being aware of different situations and considering their related risks are required [4].

Basically, for determining such situational risks, knowledge about possible hazards is necessary. Such *safety knowledge* is applied for deriving risk information based on situational perceptions or the anticipation of the robot.

The SOM approach considers dynamic processes as sequences of situations. Consequently, those situations can be assessed. Hereby, the description of situational risks is assumed to be sufficient because the dynamicity of processes is already described within situations. In such, operators represent the instantaneous change of parameters being related to such dynamic processes. The situational risk can therefore be transferred to transitions (actions/operators) leading to respective situations.

In order to identify situational hazards and to determine their risks, generally an approach analogous to traditional safety-assurance methods is favored: In a first step hazards are identified, in a second risks are assessed [11]. The hazard identification is based on scanning for the presence of hazard-causal factors by the system itself. Therefore, a logical hypothesis of participated causal factors (cf) confirms the mere presence of hazards (hazard-presence premise: $\mathrm{cf}_1 \wedge \mathrm{cf}_2 \wedge \ldots \wedge \mathrm{cf}_n \Rightarrow \mathrm{hazard}_x$). In case the presence of a hazard is confirmed, the present risk leading to a respective accident needs to be determined. Consequently, the risks of present hazards can be determined in a subsequent step. A set of risk determination instructions or a risk function, for instance, a fault tree-based risk model [5], generates the related risk value. Relevant hazard-related parameters, derived from measurements, requested from a database and so forth, can be considered as inputs.

2.2.1 Object-Interaction Risks

Most safety-related considerations in the robotic domain focus on hazards stemming from the robot itself. In former research work [6], hazards stemming from interaction of *real-world* or *scene objects*, being handled by the robot itself, were identified as currently not considered types of hazardous energy sources. Those are a special class of hazards that becomes important especially for robots being intended to manipulate different objects in object-rich environments (human environments are typically of this kind). Basically, it is assumed that objects being handled by the robot switch over from the set of environmental objects to robot-related objects, becoming in fact a part of the robotic system instead. The robotic system including such objects interacts further on with its environment, which is still required to take place safely. Hence, the robot system becomes 'responsible' for the objects it manipulates ('transition into robot-responsible objects' metaphor).

Hence, the occurrence of robot-related hazards depends on the kind of gripped objects and present environmental objects. Basically, dangerous interactions can be identified by an intensive investigation of potential objects being present in the future robotic environment. In order to adequately consider the typical problem of potential unknown 'open' environments, the risk models are generalized: The relevant object properties can be utilized for identifying upcoming hazards. For instance, the reason for fire hazard is not originated by the interference of the *chopping board* with the *hot cooking plate*, as rather wood catches fire if strongly heated. Hence, the presence of object properties *wood* and *heat source* relate to an identification principle considering inflaming wooden objects (wooden toys, bowls, cutlery, rolling pin etc.) by reason of strong heat sources (chimney, electric iron, grill, candle etc.).

The risk-related factors, encouraging the hazard becom-

ing an accident, are modeled in an instruction or function part. In the example mentioned above, factors such as the relative position of objects, temperature of a heat source or exposure time are surely important factors. Therefore, functions or instructions have to be modeled such that risk values can be computed from available data. Fuzzy methods, an evidence-based approach [6], fault trees [5] or even sophisticated functions approximators might be applied.

3 Software Architecture

A simplified overview highlighting the integration of a DRA server into robotic architectures is sketched in **Figure 2**. For the communication between components a middleware (e.g. such as SmartSoft[1] or ROS) is typically used. Such approach allows, besides others, for distributed redundancy and software diversity (various compilers). Without using a DRA server, the decision making takes place based on the interpretation of the current situation and on the anticipation capability (planning). The perception including recognition of objects from the current scene is summarized as a *Scene Interpretation / Object Recognition* (SIOR) module. The physical interaction of the robot with its environment is abstracted as a *Behavior Planning / Anticipation* (BPA) module. Both modules are treated very briefly, but are single fields of research.
The DRA server module lies in between the SIOR and BPA modules, and might be realized by multiple instances; e.g. compiled by various compilers, or being redundant on multiple machines. A subsequent voter receives redundant responses from each DRA server, allowing to detect deviations due to software errors or failures in general. According to the DRA approach, the DRA server provides risk information to a subsequent decision-making process (within the BPA module).

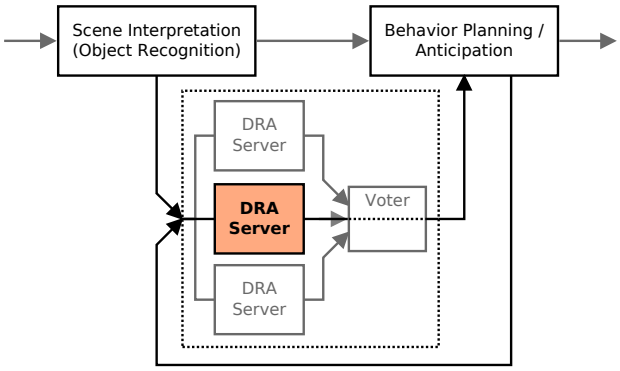

Figure 2: Sketch of inter-process communication in the DRA architecture.

Essentially, the separation of knowledge from the implementation is enforced. Furthermore, it is considered that a part of the knowledge is related also to nonsafety-relevant aspects and must therefore be accessible also to other components of the system, as it might be the case for the object database, for instance. The safety knowledge is realized as a separate knowledge base in order to provide a better overview and maintainability. A part of the UML structure is shown in **Figure 3**.

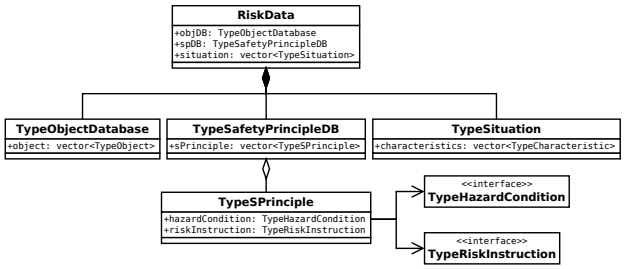

Figure 3: UML extract of the DRA server structure.

The hazard condition and risk determination parts are realized as interfaces for facilitating the integration of new features as further interfaces. With regard to this, a separation of knowledge and code is not fully taken into account.

4 Experiment

A small grid-world example is shown in **Figure 4** containing several potential hazardous object interactions. In all situations, the robot has gripped a *coffee bowl* and is following a predefined path, so far not considering any risks. A *human* enters the scene drying hairs with a *hairdryer*. The *hairdryer* is deposited afterwards and the human crosses the scene. On the robot's path it approaches first a *power plug*, *cooking plate* and later on the deposited *hairdryer*. In general, the scene consists of ten situations that are passed consecutively to the DRA server.

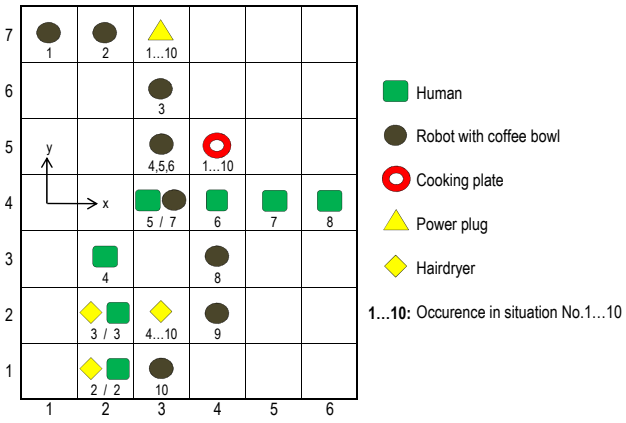

Figure 4: Simulation scene with different objects having potential for dangerous interactions.

[1] `http://smart-robotics.sourceforge.net` [Online; accessed 08-December-2011]

The DRA server is realized within a SOAP[2] middleware offering a comfortable possibility to serialize and deserialize XML structures. Hence, the situational descriptions and the knowledge are realized via XML. The DRA server is prepared with initial object knowledge. The attribute-based generalization is realized by defining object attributes according to **Figure 5**.

Objects	Attributes
Coffee bowl	hot liquid, liquid, plastic
Human	obstacle
Cooking plate	obstacle, extreme heat
Power plug	obstacle, high voltage, grippable
Hairdryer	obstacle, high voltage

Figure 5: Attributes of respective objects stored in a database.

At first, the initial safety knowledge has to be loaded into the DRA server. This safety-principles database, shown in **Figure 6**, contains four safety principles relating objects or their attributes to hazards as premises for the presence of respective hazards. The risk of each hazard is computed by the risk determination part[3]. For illustration, all risks are interpolated as linear functions in dependence of respective object distances. The interpolation takes place between two extremes: Being a very dangerous constellation on the one hand, and on the other hand, being a constellation assumed to be safe. The last two columns of the table depict the two coefficients of a linear equation. As mentioned already, any functions or instructions can be utilized in this context.

Modeled risks	Gripped obj.	Scene obj.	Dangerous		Uncritical		Function	a	b
			Dist.	*Risk*	*Dist.*	*Risk*	*Tyoe*		
Electric shock	liquid	high voltage	*1*	*0,6*	*3*	*0*	<linear>	*-0,3*	*0,9*
Scaling human	hot liquid	human	*0*	*0,3*	*2*	*0*	<linear>	*-0,15*	*0,3*
Collision risks	<any>	obstacle	*0*	*0,1*	*2*	*0*	<linear>	*-0,05*	*0,1*
Melting plastic	plastic	extreme heat	*0*	*0,8*	*2*	*0*	<linear>	*-0,4*	*0,8*

Figure 6: Example of safety knowledge (safety principles) as risk models for object-interaction accident potentials.

The results of the DRA server are illustrated in **Figure 7**. Risks are plotted situation-wise in a stack chart. The DRA server responds single risk values. The sum of these results the overall situational risk value. Hence, it becomes apparent that the assessment of anticipated situations allows for balancing of single risks, overall risks with regard to comparisons of risks, risk thresholds or even modeled task benefits.

The values chosen for risk models are surely not realistic. For real world approaches adequate approximations must be found by a respective safety-engineering process, for instance, with the help of an evidence-based approach [6].

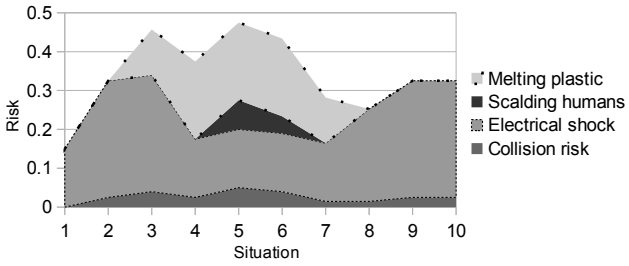

Figure 7: Illustration of DRA server's outcome being the risks comprised in the simulation example.

5 Conclusion

Dynamic Risk Assessment is assumed to be essential for ensuring safety of future *Next-Generation Robots*, which are intended to operate in and manipulate an open and unstructured environment. Thus, an architecture was suggested for realizing a *Dynamic Risk-Assessment Server* being capable also taking environmental object interaction-related hazards into account. The resulting measures describing risks of assessed situations can be considered separately in a subsequent decision-making process. Finally, a small simulation example illustrated the basic intention and operation of the proposed approach.

Acknowledgement

This work was also conducted within the collaborative center for applied research *ZAFH-Servicerobotik*. The authors gratefully acknowledge the research grants of the state of Baden-Württemberg and the European Union.

References

[1] Yueh-Hsuan Weng, Chien-Hsun Chen, and Chuen-Tsai Sun. Toward the human-robot co-existence society: On safety intelligence for next generation robots. *International Journal of Social Robotics*, 1(4):267–282, November 2009.

[2] Hans Jonas. *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. University of Chicago Press, 1985.

[3] Tim Smithers. Autonomy in robots and other agents. *Brain and Cognition*, 34:88–106, 1997.

[2] `http://www.cs.fsu.edu/~engelen/soap.html` [Online; accessed 08-December-2011]

[3] Risk is defined as the product of accident probability and severity. The severity is normalized to sev $\in (0, 1)$. Thus, the risk is $0 \leq risk \leq 1$, whereas 0 is absolutely safe and 1 represents a worst-case accident.

[4] Andrzej Wardzinski. Safety assurance strategies for autonomous vehicles. In *SAFECOMP '08: Proceedings of the 27th International Conference on Computer Safety, Reliability, and Security*, pages 277–290, Berlin, Heidelberg, 2008. Springer-Verlag.

[5] Derek Seward, Conrad Pace, and Rahee Agate. Safe and effective navigation of autonomous robots in hazardous environments. *Autonomous Robots*, 22(3):223–242, 2007.

[6] Philipp Ertle, Holger Voos, and Dirk Söffker. On risk formalization of on-line risk assessment for safe decision making in robotics. In *7th IARP Workshop on Technical Challenges for Dependable Robots in Human Environments*, pages 15–22, 2010.

[7] Philipp Ertle, Dennis Gamrad, Holger Voos, and Dirk Söffker. Action planning for autonomous systems with respect to safety aspects. In *IEEE International Conference on Systems Man and Cybernetics (SMC) 2010*, pages 2465–2472, 2010.

[8] D. Söffker. *Systemtheoretic Modeling of the knowledge-guided Human-Machine Interaction (In German)*. Habilitation Thesis, University of Wuppertal, 2001. Logos Wissenschaftsverlag, Berlin, 2003.

[9] E. Ahle. *Autonomous Systems: A Cognitive Oriented Approach Applied to Mobile Robotics*. Shaker, 2007.

[10] Dennis Gamrad and Dirk Söffker. Architecture for cognitive technical systems allowing learning from interaction with unknown environments. In *7th Workshop on Advanced Control and Diagnosis*, 2009.

[11] DIN EN ISO 12100-1. *Safety of machinery - Basic concepts, general principles for design - Part 1: Basic terminology, methodology*. Beuth Verlag, Berlin, 2004.