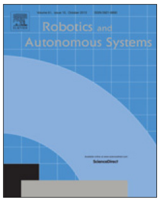
Robotics and Autonomous Systems 94 (2017) 43–52

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot



Safety-critical advanced robots: A survey

Jérémie Guiochet [a,b,*], Mathilde Machin [a,b], Hélène Waeselynck [a]

[a] *Université Toulouse, LAAS-CNRS, France*
[b] *Université Toulouse, UPS, France*

HIGHLIGHTS

- Survey on dependability techniques used for increasing safety in robots and autonomous systems.
- Identification of main issues and challenges for insuring safety.
- Review a large scope of application domains.

ARTICLE INFO

Article history:
Received 11 February 2016
Received in revised form 18 January 2017
Accepted 10 April 2017
Available online 18 April 2017

Keywords:
Dependability
Safety
Collaborative autonomous robot

ABSTRACT

Developing advanced robotics applications is now facing the safety issue for users, the environment, and the robot itself, which is a main limitation for their deployment in real life. This safety could be justified by the use of dependability techniques as it is done in other safety-critical applications. However, due to specific robotic properties (such as continuous human–robot physical interaction or non deterministic decisional layer), many techniques need to be adapted or revised. This paper reviews the main issues, research work and challenges in the field of safety-critical robots, linking up dependability and robotics concepts.

1. Introduction

Even if fictional fantasies are still far from real robots, technological improvements make them approaching reality. Technical development of the functions of such systems is a crucial issue, but if we plan that some of these fantasies come to reality in next decades, another issue can be raised, which is *what is our confidence in such systems*? A major contributor for this confidence is the justification of achieved safety. It has already been a main challenge in critical applications, like ground transportation, aeronautics or nuclear applications, the deployment of which has relied on a corpus of dependability means, as defined by [1]. Safety will obviously be a core challenge for robots deployment as well. Nevertheless, even if such systems actually belong to more general classes of systems such as embedded or cyber–physical systems, the *collaborative* and *autonomous* abilities induce important issues in the application of dependability techniques.

Dependability, and more specifically safety, has become a major challenge in robotics research projects. For instance, several recent European projects consider safety as the main challenge of human–robot cooperation like [2–5] or as a key objective together with maintainability in [6–9]. National projects such as [10] in the UK, [11] in Germany, [12] in the USA, and dedicated research teams (e.g., [13] in USA) or institutes (e.g., [14] in Japan) also focus on robot reliability and safety. Although many work in the robotics community focus on robot functions linked with safety (e.g. intrinsically safe robot[1] [15], actuators compliance [16–18], collision avoidance control or human aware motion [19]), we focus in this survey on work which addresses safety of robotic abilities by considering dependability means such as fault avoidance and treatment techniques. By safety, we will not only consider human integrity, but also the environment or the robot itself integrity.

We first introduce in Section 2 the context of our survey, i.e. abilities of considered robotic systems, such as autonomy and interaction, then some examples of induced hazards, and current European robotic safety standards. Then, we present a survey on dependability means for such robots in Section 3. Section 4 provides a selection of main challenges in the field of safe robots and Section 5 concludes this survey.

2. From industrial to advanced robots—New hazards

Among the large diversity of robotics applications and their associated social and ethical issues [20], safety is not a new concept.

---

* Corresponding author at: Université Toulouse, LAAS-CNRS, France.
E-mail address: jeremie.guiochet@laas.fr (J. Guiochet).

[1] See for instance, products such as the LWR LightWeight Robot III commercialized by KUKA, or UR5 from Universal Robots.

http://dx.doi.org/10.1016/j.robot.2017.04.004

Table 1
Core properties of industrial and advanced robotics, and examples of induced hazards.

		Industrial robotics	Advanced robotics	New hazards examples
Autonomy	Robot control	Automatic	Decisional autonomy	Hazardous decisions
	Workspace	Structured	Non-structured (uncertainties)	Adverse situations / uncertainties in perception
Collaboration	Motion	No robot motion in human presence	Simultaneous motion (human and robot)	Bad synchronization between human and robot / Non-human-legible movements
	Human–robot closeness	Human is far	Human is close/Physical interaction	Collisions, contact forces too high
	Human–robot communication	Remote device	Advanced interaction (cognitive)	Mode confusion / communication errors
Task	Mechanical architecture	Heavy / Stiff / Powerful	Light / Compliant / limited power (“intrinsically safe” [15])	Precision hazards / energy storage due to compliance
	Task complexity	Mono-function	Multi-functions	Safety rules not adapted (diverse and evolving rules)

It has been studied for years in manufacturing applications, particularly for industrial robots. But the emergence of advanced robots with new abilities, such as decisional autonomy and physical interaction with humans, forces consideration of hazards that did not exist in traditional industrial robots. Table 1 presents a comparison between industrial and advanced robotics. The distinguishing characteristics in terms of autonomy and collaboration, as well as the induced new hazards, are detailed in the subsections below. We then discuss the status of safety standards with respect to advanced robots.

2.1. Autonomy and collaboration

Autonomy is made possible by the introduction of a decisional software layer in the robot architecture. Such a layer exists in service or field applications, and in robotized systems like UAVs (Unmanned Aerial Vehicles), spacecraft, or self-driving cars. These systems are able to act deliberately regarding their mission, in diverse environments (referred as “non structured” workspace in Table 1). It exists a wide range of degrees between what we can call automatic systems (automatic control for industrial robots in Table 1) and fully autonomous systems (see the 11 levels in the European SPARC roadmap [21], 3 in [22], or 5 for vehicles in the USA roadmap [23]). Nevertheless, in this survey, we mainly use a basic automatic/autonomous dichotomy, without intermediate degrees. Cases where the degree of autonomy may impact the applicability of a dependability technique are discussed specifically in Section 3. Technically, these degrees of autonomy may be implemented in a variety of robotic architectures. Fig. 1 presents the well-known abstraction into three layers:

Decisional layer: It receives objectives from another system, or an operator and generates some plans according to an abstract representation of the system and its environment. Functions for deliberation (e.g., planning, learning or goal reasoning [22]) are usually based on knowledge specific to the application domain (such as heuristics or an environment model) and an inference mechanism used to solve problems by manipulating this knowledge. Execution time is not guaranteed and outputs/results are not deterministic. The use of heuristics is not guaranteed to be optimal or perfect, but sufficient to find solutions.

Executive layer: It converts plans sent by the decisional layer, into primitive functions for the functional level.

Control/Functional level: It is in charge of feedback control loops coupling sensors to actuators, of perception facilities and trajectory computation.

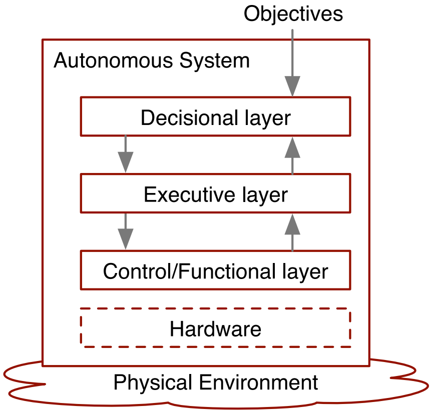


Fig. 1. A three layer architecture for decisional autonomy.

Removing the protective fences around robots, led to the development of human–robot interaction, where human and robot share task execution and may interact to synchronize their actions. As presented in Table 1, such a collaboration is based on human robot closeness (*far* for industrial robots, and *physical Human Robot Interaction* – pHRI – for advanced robots), on communication means (remote devices or cognitive signals such as voice or posture) and on simultaneous motion of the robot and the human. We proposed in the PHRIENDS project [2] to use an interaction classification [24] mixing the *closeness* and *motion* properties of Table 1 (for medical robots, defined as active medical devices, such a classification is given in the European Directive 93/42/CEE [25]):

Far: No pHRI possible, human and robot are not sharing the same workspace; a direct physical contact should be not possible.

Close: Accidental pHRI possible, human and robot are sharing the same workspace. Since the human is within the robot’s reach there is a risk of unwanted, potentially harmful physical contact.

Touching without simultaneous movement: pHRI only takes place when the robot stops, the robot shares its workspace with the human. Both are simultaneously moving through the workspace, but physical contact with the moving robot is avoided.

Touching with simultaneous movement: pHRI possible and intended, the robot shares its workspace with the human. Both are moving simultaneously and physical interaction is possible and intended.

Supporting: Continuous pHRI, physical interaction occurs continuously over extended periods of time.