Metadata of the article that will be visualized in OnlineFirst

Please note: Images will appear in color online but will be printed in black and white.

1	Article Title	**Environmental Hazard Analysis - a Variant of Preliminary Hazard Analysis for Autonomous Mobile Robots**	
2	Article Sub-Title		
3	Article Copyright - Year	**Springer Science+Business Media Dordrecht 2014 (This will be the copyright line in the final PDF)**	
4	Journal Name	Journal of Intelligent & Robotic Systems	
5	Corresponding Author	Family Name	**Dogramadzi**
6		Particle	
7		Given Name	**Sanja**
8		Suffix	
9		Organization	University of the West of England
10		Division	Bristol Robotics Laboratory
11		Address	Bristol, UK
12		e-mail	sanja.dogramadzi@uwe.ac.uk
13	Author	Family Name	**Giannaccini**
14		Particle	
15		Given Name	**Maria Elena**
16		Suffix	
17		Organization	University of the West of England
18		Division	Bristol Robotics Laboratory
19		Address	Bristol, UK
20		e-mail	maria.elena.giannaccini@brl.ac.uk
21	Author	Family Name	**Harper**
22		Particle	
23		Given Name	**Christopher**
24		Suffix	
25		Organization	University of the West of England
26		Division	Bristol Robotics Laboratory
27		Address	Bristol, UK
28		e-mail	cjharper@avian-technologies.co.uk
29	Author	Family Name	**Sobhani**
30		Particle	
31		Given Name	**Mohamed**
32		Suffix	

33		Organization	University of the West of England
34		Division	Bristol Robotics Laboratory
35		Address	Bristol, UK
36		e-mail	None
37	Author	Family Name	**Woodman**
38		Particle	
39		Given Name	**Roger**
40		Suffix	
41		Organization	University of the West of England
42		Division	Bristol Robotics Laboratory
43		Address	Bristol, UK
44		e-mail	roger.woodman@brl.ac.uk
45	Author	Family Name	**Choung**
46		Particle	
47		Given Name	**Jiyeon**
48		Suffix	
49		Organization	University of the West of England
50		Division	Bristol Robotics Laboratory
51		Address	Bristol, UK
52		e-mail	jiyeon.choung.2011@my.bristol.ac.uk

53	Schedule	Received	25 May 2013
54		Revised	
55		Accepted	27 December 2013

56 Abstract

Robot manufacturers will be required to demonstrate objectively that all reasonably foreseeable hazards have been identified in any robotic product design that is to be marketed commercially. This is problematic for autonomous mobile robots because conventional methods, which have been developed for automatic systems do not assist safety analysts in identifying non-mission interactions with environmental features that are not directly associated with the robot's design mission, and which may comprise the majority of the required tasks of autonomous robots. In this paper we develop a new variant of preliminary hazard analysis that is explicitly aimed at identifying non-mission interactions by means of new sets of guidewords not normally found in existing variants. We develop the required features of the method and describe its application to several small trials conducted at Bristol Robotics Laboratory in the 2011–2012 period.

57	Keywords separated by ' - '	Hazard analysis - Environmental survey - Autonomous - Mobile robot - Safety
58	Foot note information	The online version of this article (doi: 10.1007/s10846-013-0020-7) contains supplementary material, which is available to authorized users.

Electronic supplementary material

(PDF 124 KB)
(PDF 76.5 KB)

J Intell Robot Syst
DOI 10.1007/s10846-013-0020-7

Environmental Hazard Analysis - a Variant of Preliminary Hazard Analysis for Autonomous Mobile Robots

Q1 **Sanja Dogramadzi · Maria Elena Giannaccini · Christopher Harper · Mohamed Sobhani · Roger Woodman · Jiyeon Choung**

Received: 25 May 2013 / Accepted: 27 December 2013

Abstract Robot manufacturers will be required to demonstrate objectively that all reasonably foreseeable hazards have been identified in any robotic product design that is to be marketed commercially. This is problematic for autonomous mobile robots because conventional methods, which have been developed for automatic systems do not assist safety analysts in identifying non-mission interactions with environmental features that are not directly associated with the robot's design mission, and which may comprise the majority of the required tasks of autonomous robots. In this paper we develop a new variant of preliminary hazard analysis that is explicitly aimed at identifying non-mission interactions by means of new sets of guidewords not normally found in existing variants. We develop the required features of the method and describe its application to several small trials conducted at Bristol Robotics Laboratory in the 2011–2012 period.

Keywords Hazard analysis · Environmental survey · Autonomous · Mobile robot · Safety

1 Introduction

As autonomous mobile robots become a commercial reality, attention must be paid to the problem of assuring their safety. In almost every application of mobile robots other than toys, the size, power or speed of robots will be such that potential hazards will be associated with their operation or malfunction. Legal regulations in most countries require that any such safety critical system be designed so as to reduce the risk of accidents caused by these hazards to less than some required threshold, or at least as low as is reasonably practicable.

The achievement of safety in engineering systems requires a combination of different approaches of safety requirements specification, analysis, design and manufacturing inspections, and product testing. The objective of these is to determine what hazards are associated with the system, to specify and implement

Electronic supplementary material The online version of this article (doi:10.1007/s10846-013-0020-7) contains supplementary material, which is available to authorized users.

Q2 S. Dogramadzi (✉) · M. E. Giannaccini · C. Harper · M. Sobhani · R. Woodman · J. Choung
Bristol Robotics Laboratory,
University of the West of England, Bristol, UK
e-mail: sanja.dogramadzi@uwe.ac.uk
URL: http://www.brl.ac.uk

M. E. Giannaccini
e-mail: maria.elena.giannaccini@brl.ac.uk

C. Harper
e-mail: cjharper@avian-technologies.co.uk

R. Woodman
e-mail: roger.woodman@brl.ac.uk

J. Choung
e-mail: jiyeon.choung.2011@my.bristol.ac.uk

features of the design that act to reduce the probability of an accident, and then to confirm whether each product that is actually manufactured does indeed possess the intended properties when operating in its intended environment(s).

This paper presents the results of recent research performed by the authors at Bristol Robotics Laboratory (BRL) into methods of analysis of robotic systems for the identification of potential hazards associated with autonomous operation in diverse environments. Much of the work was carried out as a background activity to the European INTRO project (www.introbotics.eu), and some work as internal research and postgraduate projects solely within BRL. The results of the application of Hazard Analysis in INTRO research conducted in BRL is summarized in the work of [10]. Several studies have been performed on different robotic applications, and lessons learned in early efforts have resulted in proposals for a new method, Environmental Surveys, which have then been applied in later trials. In this paper, we present the work that was performed, and draw conclusions about the effectiveness of the new method and ideas for future work that emerge from these studies.

1.1 The INTRO Project

INTRO (www.introbotics.eu) seeks to better understand issues in Human-Robot interaction and, ultimately, endow the robot with cognitive and physical intelligence sufficient to deal with complex situations and safety of typical interactions. The 4 year long, Initial Training Network project, sponsored by the European Commission*, has trained 8 young researchers to prepare them for careers in the fast developing area of service robotics. They explored various aspects of interactions - from learning by demonstration, intention and emotion recognition, to gesture analysis, intelligent interfaces and safety factors. The individual topics will be integrated into two different scenarios designed and developed by two post-doctoral researchers on the project employed by two European robotic companies – Space Applications (Belgium) and Robosoft (France). The two scenarios – Search and Rescue and Robot-waiter have been selected to be best to demonstrate what robots need to do in situations that require communication between humans and the robot and that are placed in noisy and dynamic environments. In both cases, hazards and faults are inevitable.

1.2 Industry Safety Standards for Autonomous Robots

In addition to existing research into safety issues for mobile autonomous robots, BRL has also supported UK participation in the ISO TC184 SC2 (Robots and robotic devices) committee in its development of a new industry standard ISO 13482 [22], which specifies safety requirements for (non-medical) personal care applications of service robots. These include domestic service robots, physical assistant robots (e.g. exoskeleton-type assistive robots or human load-sharing mobile robots) and person carrier robots (autonomous mobile passenger carts). The standard includes lists of hazards that are predicted to be commonly encountered, so standard levels of safety performance can be specified that can offer a baseline performance level which can be assessed and certified. ISO 13482 is due for public release in late-2013, and at time of writing is in its final draft stage. The work in this paper is intended to supplement the publication of the standard by offering guidance on how to perform the hazard identification task for the kinds of robots covered by ISO 13482.

1.3 Structure of this Paper

In Section 2 of this paper we review existing work on the topic of hazard identification of autonomous mobile robots. In Section 3 of this paper, we present a review of current methods for functional hazard analysis, as developed in numerous existing (non-robotic) industry sectors. In Section 4 we present the initial hazard analysis study, and we discuss the problems facing the task of hazard identification for systems that operate autonomously in open environments, which led us to develop the new method of Environmental Surveys. In the Section 5 we present the new method and in Section 6 we present its initial trials. In Sections 7 and 8 we discuss the results and present our conclusions about the effectiveness of the work and how it should progress in the future.

2 Background

In this section we discuss the main safety issues associated with designing an autonomous service robot.

2.1 Safety of Autonomous Robotic Systems

Autonomous robots are a class of robot system which may have one or more of the following properties: adaptation to changes in the environment; planning for future events; learning new tasks; and making informed decisions without human intervention. Although commercially available autonomous robots are still few, [12] report that there is increasing demand for both personal robots for the home and service robots for industry.

At present, much of the research into robotic safety is looking at improving design of safety mechanisms, for example collision avoidance [19, 24] or fault detection and tolerance Petterson 2005, object manipulation [13], or human contact safety [17]. This has led researchers to suggest that safety of human-robot interaction requires both high-precision sensory information and fast reaction times, in order to work with and around humans [11, 25]. Work by [2] suggests that for autonomous systems to support humans as peers, while maintaining safety, robot actions may need to be restricted, preventing optimum flexibility and performance. Other work in robotic safety focuses on risk quantification, for example [16] and [21].

In contrast, our work is concerned with initial identification of hazards and their associated safety requirements. It is not concerned with risk assessment, or the design and implementation of safety mechanisms and fault detection such as the work described by Petterson 2005. The only work we are aware of, which is similar to this paper, is that of Guiochet and Baron [14], Guiochet et al. [15], Martin-Guillerez-et al. [28] (see Section 2.2 for a detailed discussion).

One of the principle requirements for dependability in autonomous robots is robustness. This means being able to handle errors and to continue operation during abnormal conditions Lussier et al. 2004. To achieve this it is important that the system should be able to support changes to its task specification [4]. These changes are necessary as, in a dynamic environment, the robot will frequently find itself in a wide range of previously unseen situations. While this is not a subject covered in this paper, our work does also lead us to similar conclusions – see Section 8.2.

It is clear from the literature that little research has been done on the day-to-day operation of personal robots, and all the safety risks associated with this. One reason why this may be the case, is that currently personal robots are only tested in 'mock' home conditions that have been heavily structured and the majority of real world hazards removed. Therefore there has been no need to conduct a survey of many of the real environments, in which personal robots may be required to operate.

2.2 Results of Robot Studies Using Hazard Analysis

One of the few research works for hazard analysis of service robots has been published by [15]. Their research considers the MIRAS RobuWalker, which is a robotic assistant for helping people stand up from a seated position and support them while walking. The RobuWalker can be used in two modes, a user controlled mode and an automation mode. The user controlled mode is used when the human is supported by the robot in a standing position. The automated mode is required when the human is in a seated position. This mode allows the user to request the robot to move from its stored position, which could be anywhere in the room, to the location where the human making the request is located. This involves the robot navigating the environment with no assistance from the user. Based on the hazard analysis results that have been published, it is clear that only hazards associated with the normal operation of the robot have been considered. For example there are no hazards recorded associated with other non-task related entities that may be present in the robot's operating area. This issue of not analysing hazards that are not directly associated with the robot's task has also been identified in other projects. A study by [6] examined a therapeutic robot for disabled children. To analyse the safety of this device, the researchers used the hazard analysis technique HAZOP. This method examined how the child and robot would interact and considered the potential safety risks. However, as with the previous example, no consideration is given to the types of hazard that the robot may encounter outside the predefined tasks.

The PHRIENDS project [1, 28] performed hazard analysis on a wheel-based mobile robot with a manipulator arm that was designed to pick up and move objects around the environment. This robot, which was required to work collaboratively with a human user, was designed to safely navigate a dynamic environment that could contain multiple humans. This represents the largest scale hazard analysis of a personal robot found in the literature. Their analysis considered the safety risks of the robot from a number of positions, including the potential hazards of each major component of the robot failing, the risks associated with human users, and the types and severity of collisions that may occur.

As has been discussed in this paper, traditional hazard analysis methods for service robots can result in safety risks outside the normal operating scenarios being missed. To address this issue, research by [38] has proposed the use of a hazard analysis check list. This check list highlights a number of environmental and user risks that need to be considered when assessing the risk of a personal robot. Although this research concludes that the check list cannot be shown to identify all the potential safety risks.

The following section presents the findings of the experiments conducted at the BRL, and discusses their implications for the safety analysis of service robots.

3 Hazard Identification Analysis

Hazard identification analysis (often referred to simply as 'hazard identification' or 'hazard analysis') is required as a safety assurance activity during the requirements specification and early design stages of any safety critical system (it is often required as a mandatory activity by industry safety standards). This section provides an overview of the subject, and discusses the issues that affect the analysis of autonomous mobile robots.

3.1 Conventional Theory and Methodology

In most countries, national laws require that all reasonable steps be taken to ensure that products or processes sold to consumers or used in workplaces are safe as far as is reasonably practicable. Depending on the legal codes and practices of a given nation, the mandate for "reasonableness" is either written explicitly into legislation as in the UK Health & Safety at Work and Consumer Protection Acts [36, 37] or it is implicit within the legal code as in many other European countries [8]. In either case, the result is the same – it is incumbent on manufacturers and employers to ensure that risks are reduced *"so far as reasonably practicable (SFAIRP)"* or *"as low as reasonably practicable (ALARP)"* (these terms are synonymous, but the latter is more popular). It is generally considered, at least in the UK [8], that the risk of harm cannot be reduced as low as reasonably practicable unless the following can be shown *objectively* (i.e. without allowance for any personal qualities of a manufacturer, employer, or vendor):

- the harm was not foreseeable,
- the safety measures taken were not reasonably practicable, or
- the harm was outside the scope of the undertaking (manufacturers/employers are not liable for that which is outside the scope of their responsibility).

Of these three criteria, the first and third present particular challenges to developers of mobile autonomous robots, and are the ultimate objectives to which the methods proposed in this paper are dedicated.

In order to satisfy these criteria, engineers perform a variety of *safety assurance* tasks during the design of a safety critical system. Methods and processes for safety-directed design and testing are outside the scope of this paper, but safety assurance also includes a number of procedures to identify potential sources of harm, and for delineating the scope of consideration to the boundaries of the manufacturer's responsibility. These methods and procedures are generally referred to as *hazard analysis* or *hazard identification*.

3.1.1 Background on Hazard Identification

The hazard identification process is the start of the safety assurance process of any safety critical system. The general objective of hazard identification is to define all the possible hazards that might occur in a system throughout its operational life. However, the unbounded definition of the operational time and of the environment of a system means that it cannot be guaranteed formally whether all possible hazards have been identified. So typical hazard analysis methods seek to try and provide a *systematic classification* of hazards, which can identify all the logical *types* of hazards but not all the specific *instances* of hazards (the events themselves), which safety assurance engineers must determine based on their knowledge and intuition.

Hazard identification is first started at an early stage in the system development process, typically once the initial version of the system requirements specification is available. Hazard identification analysis done at this stage is often referred to as *Preliminary Hazard Analysis* or *Identification* (*PHA* or *PHI*), because it is often the case that the only design information available for analysis are the most abstract (high level) and basic functional requirements defining what the system is to do – details about the general nature of the actuation mechanisms or the interfaces between the system and its environment have not yet been specified. Later, as the general physical structure is defined and the details of the boundary interfaces are specified, the hazard analysis is often referred to as *Functional* or *System Hazard Analysis* (*FHA* or *SHA*).

3.1.2 Contemporary Hazard Identification Methodologies – a Review

A number of variants of preliminary and functional hazard identification methods have been developed over the years, often for different industrial sectors reflecting the particular technological domains, design practices, conventions and terminology. This section describes the general principles, and reviews some of the more widely used methods from different industry sectors.

Hazard Identification Analysis – General Principles The aim of hazard analysis is to identify all plausible and reasonably foreseeable hazards associated with a system's operation in its environment. For identification of functional hazards this is typically achieved by two general approaches, which are canonical so their use is equivalent in functional term.

The two approaches are based on two variations in the modelling of failures and their effects within system functional models, which are illustrated in Fig. 1. In general, system functions are modelled as input/output processes encapsulated within the system's boundary and interacting with the outside world via the system interface. Hazards arising from defects within the system can then be modelled by defining *failure conditions* of the elements of the system model, in the two respective viewpoints.

The first approach – the function-oriented view- is to model failures as defects of the functional processes. The requirements of each system function are inspected, and fault or error conditions associated with each requirement are identified and assessed for their consequences on the external environment via the system interfaces. The hazard analysis builds up a classification table or diagram of system failure conditions on a function-by-function basis, with interface behaviour being a secondary description within each function-based classification category.

In contrast, the second approach – the interface-oriented view – models failure conditions at the boundary interface of the system. Fault or error conditions are identified for all the parameters that define the interface, and the consequences of each parameter failure on the performance of the system functions is assessed for its consequences, and the hazard analysis table or diagram is built up in terms of system interfaces and the failure of their parameters.

With respect to system functional safety, the two approaches are canonical: a system failure cannot have any effect on safety unless it affects the way in which the system interacts with the outside environment. An internal fault or error that causes no change in the behaviour of the system at its interface to the outside world has no effect on safety, so the only defects that are of interest are those where failure conditions at the boundary are paired with failure conditions of functional processes, so if one can provide a complete classification of either then all relevant failure conditions will be identified.

Example of Function-oriented Hazard Identification – Aircraft Industry FHA Functional Hazard Assessment (FHA) was originally developed in the aerospace sector, although the name and methods have been carried across to other industries. The standard procedures and practices for performing this method in the civil aerospace sector have been codified in the ARP 4761 standard [3]. The general approach is to examine the functional requirements specification of a system, and then to identify three generic *failure conditions* associated with each functional requirement:

- Failure to operate as/when intended
- Unintended or inadvertent operation
- Malfunction (a.k.a. misleading function)

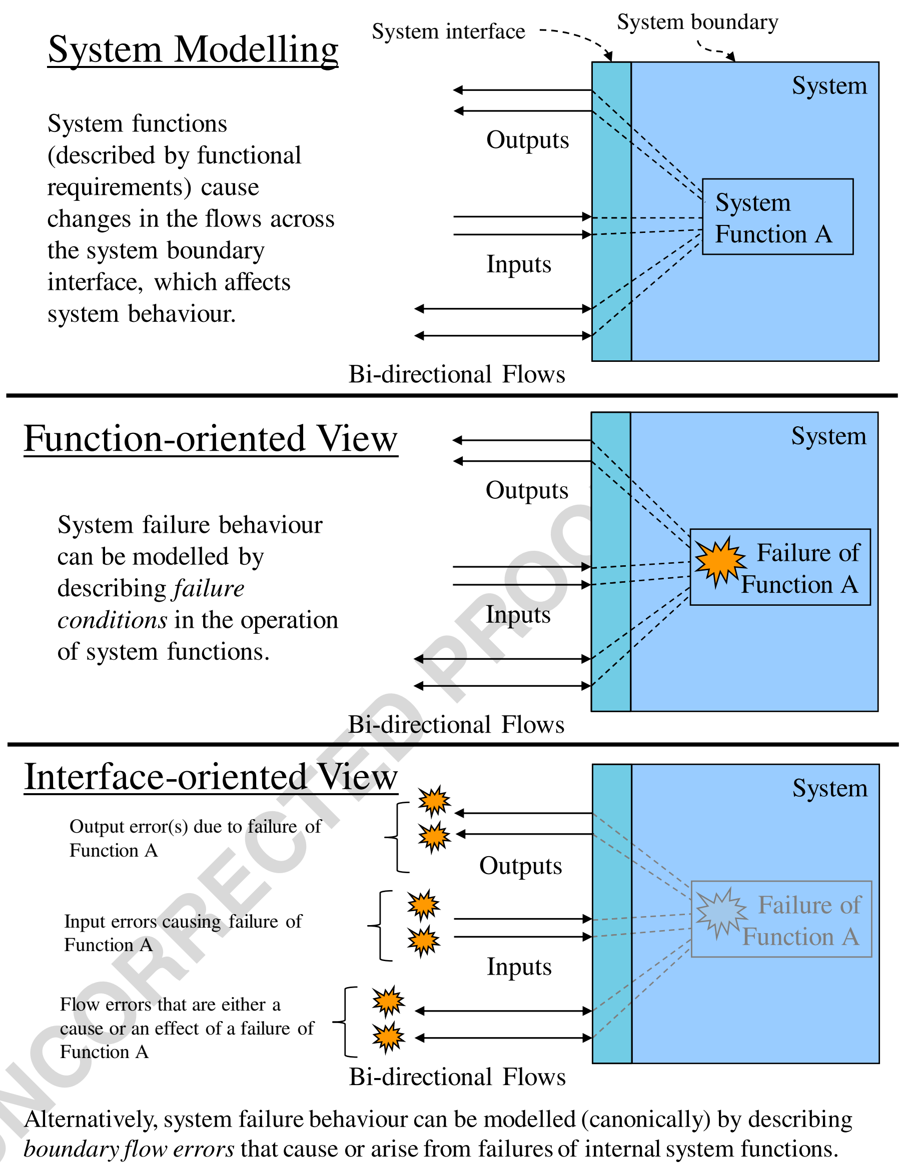


Fig. 1 Canonical representations of failures typically used in hazard identification analysis

The method proceeds by generating three hypothetical failure conditions (one of each type) for each functional requirements of the system. Hypothetical conditions that are implausible can be ignored, but for all others a precise description of the failure condition is defined. Then, for each failure condition the consequences of the condition are identified. Since the nature of the system's environment often varies throughout the operational use of a system, the consequences are assessed over different partitions of the system mission (in an aircraft these are its flight phases such as take-off, landing, cruise, etc.) in order to identify different consequences of the same failure condition if it was to occur in different environmental circumstances. The severity of harm of each distinct consequence is determined, usually in terms of the number and degree of injuries caused to persons (crew, passengers or third parties). These hazard identification results are then used as the basis of a risk assessment, where the probability of occurrence of each failure condition is assessed and if found to present an unacceptable risk then the system function can be redesigned so as to eliminate the problem, or safeguards built into the design to reduce the expected probability of occurrence to such a level that the risk is acceptable. The results of the FHA are usually presented in tabular format similar to the example shown in Table 1.

Example of Interface-oriented Hazard Identification – HAZOP One of the most widely known interface-oriented analysis methods is HAZOP (HAZard and Operability studies). This method was originally developed in the chemical process control industry, and has since been codified in the IEC 61882 standard [20]. As discussed earlier, HAZOP proceeds by a systematic analysis of failure conditions in the *flow parameters* across the boundary interface of the system. In general, flows are any information (data, signals), energy (electrical or mechanical power), fluid flow (chemical reagents, fuel), or mechanical force (structural loads and stresses, mechanical actions) that pass across the system boundary.

HAZOP identifies a number of *guidewords* which have the same role as the generic failure conditions of aerospace industry FHA. Guidewords are generally tailored to the technological domain of the system being analysed, i.e. different keyword sets for electrical/hydraulic/pneumatic/mechanical machines, fluid dynamical interfaces or mechanisms, analogue or digital electronics, software processes. However, most keywords relate to the flow of energy, force, information, or physical material across the system boundary interface, and generally identify deviations in the value, timing, or provision of service across a boundary interface. The guidewords that were originally identified for the original HAZOP version (as specified in IEC 61882 [20]) are listed in Table 2.

The method proceeds by developing an *interpretation table* for the flow parameters of the system, where the keywords are applied to the parameter types and specific definitions of the failure conditions are defined, if the combination is plausible. Some examples of guideword interpretations are provided in Table 3. Then the relevant interpretations are applied to the parameters of the boundary interface and the effects on system functions and consequences on its interaction with the environment are assessed. The results are tabulated in a similar manner to the format shown in Table 1.

Since HAZOP was originally developed for industrial process control systems, variants of HAZOP have been proposed for computer systems and software, which follow the same general methodology but propose guidewords that are more appropriate for flows of data and electronic signals than fluid and mechanical forces. Two variants of note are defined in the UK Defence Standard 00-58 [35] and the SHARD Method, developed at the University of York [32]. The former uses the same guideword set as basic HAZOP but offers guidance that is more tailored to the study of computer-based systems. The latter is notable in that it proposes a different set of guidewords developed from a survey of computer/software failure cases. The new guidewords are related to the functional service that is provided through a given flow parameter, and are described in Table 4.

Although the guideword set is different to HAZOP, the procedural methodology of SHARD is otherwise unchanged, with interpretation tables being developed for the range of software/electronic interface flow parameter types, and then the specific failure conditions being applied to the actual parameters of each such interface to determine the functional failures and their consequences.

The SHARD guideword set is interesting; its definition of failure types in service provision terms and flow behaviour terms is (respectively) both function-oriented and interface-oriented. This was one of the reasons why the SHARD guideword set was used in the initial hazard analysis studies of a robot waiter at BRL, which are described in Section 4.

3.1.3 Other Keyword Based Safety Analyses: FMEA

Hazard analysis is not the only safety analysis technique to use a keyword-driven approach – another widely used technique is Failure Modes and Effects Analysis (FMEA). FMEA differs from FHA in two principal ways – the keyword set and the level of design detail used as the information on which the analysis is based. FMEA is typically applied at a much later stage of system development, when a detailed design is available for the system and its components. The keywords used are often related to very specific fault types of physical components (e.g. short-circuit faults, varying parameter values). FMEA was employed as a safety analysis technique on one of the BRL projects discussed in this paper. In one of the SAR robot design studies, FMEA was used to analyse a particular robot task (tele-operated navigation).

Table 1 Example hazard identification analysis table format

#	Model Element	Keyword	Mission Phase/Mode	Failure Description	Consequence Description	Consequence Severity	Possible Corrective	Residual Probability	Cause(s)	Design Recommendations
1	Function A	Omission	Normal operation	Function does not operate when intended	Robot fails to perform service	Marginal	1. User action 2. Redundant subsystem 3. Diverse function	1. 10^{-6} hr^{-1} 2. 10^{-4} hr^{-1} 3. 10^{-8} hr^{-1}	Faults or design errors in subsystems performing Function A	System shall incorporate a diverse function for Function A. • Function A: SIL 1 • Diverse function: SIL 1
2		Commission	Protective stop	Function operates when not intended while a protective stop is in progress	Inadvertent operation prevents safe stop – major injuries to robot user(s) and/or third parties	Critical	Etc.	Etc.	Etc.	Etc.
3		Early/Late	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.
4		Coarse error	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.
5		Subtle error	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.
6	Input A	Omission	All phases	Loss of input signal	Function A fails to operate	Marginal	1. Input validation mechanism 2. Redundant input 3. Diverse input	1. 10^{-5} hr^{-1} 2. 10^{-6} hr^{-1} 3. 10^{-8} hr^{-1}	Fault in Input A interface element	• Validation mechanisms shall be provided for Input A • Input A shall be dual redundant • Function a shall receive Input B as a diverse check against Input A
7		Commission	All phases	N/A – input is required to be permanently active	N/A	N/A	N/A	N/A	External system/ process transmits information erroneously via Input A.	N/A
8		Early/Late	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.	Etc.

In many practical industrial hazard analyses, the process includes both hazard identification and *risk assessment*, where the severity and predicted probability of occurrence of a failure condition are assessed. Residual probability estimates frequency of occurrence of specific failures after all safety measures have been taken into account. Typically, probability estimates are obtained by reliability analysis of the system design or from established references e.g. reliability databases. This paper is not concerned with the problem of risk assessment, only the problem of how to identify a set of hazards that is as complete as is reasonably foreseeable. Therefore, probability analysis will not be discussed any further

For example, in the SAR Robot design problem, an initial assumption was that when the rescuer offers a piece of rubble he or she knows the robot gripping size capacity. However, it is possible that a fatigued rescue worker picks a wrong-size piece of rubble and passes it to the robot. Thus, the robot needs a software module to assess the offered piece. As an initial design step, Hierarchical Task Analysis (HTA – see Appendix A) was used to identify interaction-related tasks, to define a basis on which possible failure modes can be identified using FMEA. A well-known task analysis approach, HTA provides a description of the system operations toward achieving system end goal by clarifying relationships between tasks and sub-task and their order of execution [23]. The task hierarchy is developed by assigning ultimate goal of the system at top and then defining each tasks involved in goal attainment. In each level, a plan describes the order of execution of tasks. FMEA was originally established for system components reliability analysis and later its application extended to human error analysis. This technique provides compact information about the system failures in a tabular format. Hence, it was expected to be a strong tool to address failures of both sides of interaction; the robot and a human rescuer. One row of the obtained FMEA table [34] for one of the tasks failure is presented in Table 5. Failure of tele-operated navigation is when operator tries to send the robot to a position, while the robot obstacle avoidance module prevents it to move to get there. This failure can be due to either lack of the operator's situation awareness or a fault in the robot reasoning or sensory information.

This analysis provides a concise frame work for investigating different aspects of the system, qualitatively. FMEA outcome is fed to a Fault Tree Analysis (FTA) to investigate the role of each involved element for each revealed failures modes. Originally developed in the aerospace and defence industries, FTA is a powerful method utilized to assess reliability of multifaceted systems. A tree-like diagram structure is used to demonstrate the contribution of the basic events and their relative importance in a specific system failure mode. A fault tree is developed for each failure mode revealed in the FMEA. For each tree, the relationship between contributed elements toward the system failure is described by Boolean algebra and finding minimal cutest expression. This analysis can potentially provide both qualitative and quantitative frameworks for prioritizing role and importance of each faulty component. Although qualitative FTA has been insightful, performing a quantitative analysis is faced a serious challenge of finding failure and success rates and probabilities. For hardware components it is possible to have such data based on their reliability tests, nonetheless, finding failure rate of software modules and human error probability is far more difficult and challenging. Even the performance of hardware components can differ from their published reliability values when the robot is in an unpredictable and dynamic disaster environment. It is also noteworthy that qualitative FTA has been performed for a semi-autonomous robot and based on a certain restricted scenario [26] in which all the basic events have been predicted in advance, while for a fully autonomous robot predicting all the basic events is difficult to achieve.

Table 2 HAZOP generic guidewords

Guide word	Meaning
No or not	Complete negation of the design intent
More	Quantitative increase
Less	Quantitative decrease
As well as	Qualitative modification/increase
Part of	Qualitative modification/decrease
Reverse	Logical opposite of the design intent
Other than	Complete substitution
Early	Relative to the clock time
Late	Relative to the clock time
Before	Relating to order or sequence
After	Relating to order or sequence

Table 3 Sample HAZOP guideword interpretation table

Parameter/ guide word	More	Less	None	Reverse	As well as	Part of	Other than
Flow	high flow	low flow	no flow	reverse flow	deviating concentration	contamination	deviating material
Pressure	high pressure	low pressure	vacuum		delta-p		explosion
Temperature	high temperature	low temperature					
Level	high level	low level	no level		different level		
Time	too long/ too late	too short/ too soon	sequence step skipped	backwards	missing actions	extra actions	wrong time

4 Initial Experiments in Hazard Analysis of Robots – Robot Waiter Application

The research at BRL began as an exercise to support the authors' contributions to the development of the ISO 13482 industrial safety standard for mobile service robots. The standard includes a list of hazards that are expected to be common to many robot designs, and the original aim of the exercise was to conduct a hazard analysis of a proposed design to determine other possible hazards that could be submitted to the list. A partial mobile robot application design was developed to a point where a preliminary hazard analysis could be conducted, although it was not envisaged that the design would be taken through to full implementation.

The original intent of the analysis study was to apply existing hazard analysis techniques that have been developed for conventional industrial systems, with the secondary aim of evaluating the suitability of existing design and analysis methods to autonomous system applications. However, the attempt revealed a number of problems, the result of which was the proposal of a new method.

In this section we describe the specification of the robotic application that we studied, the hazard analysis technique that was applied, and we discuss the results that were obtained from the analysis sessions.

4.1 Robot Waiter Task Specification

Preliminary hazard analysis requires at least a high-level/abstract system model on which to operate, so it was necessary to produce a basic specification and architecture model of the Robot Waiter as input to the PHA process. A basic task specification of the robot was developed using Hierarchical Task Analysis (HTA, see Appendix A) and a preliminary system architecture model was developed using the NASA Goddard Agent Architecture reference model (see Appendix B). This allowed a basic identification of the functional processes that might serve as architectural components of such a system. The task-process model was then taken as the basis for the PHA. The Robot Waiter task involves an autonomous mobile robot acting as a human waiter, delivering drinks to a human customer. Specifically this requires the robot to be

Table 4 SHARD generic guidewords

Service failure	Guideword	Meaning
Service provision	Omission	Functional service not provided when intended
	Commission	Functional service provided when not intended
Service timing	Early	Functional service provided earlier than intended
	Late	Functional service provided later than intended
Service value	Coarse	Value of functional service parameters is coarsely incorrect (illegal value)
	Subtle	Value of functional service parameters is subtly incorrect (value is legal but incorrect)

Table 5 The first row of the FMEA table

Task	Failure mode	Causes	Fault/error type	Failure effect	Potential recovery type	Severity
1.1-Tele-operated Navigation	Paradox	Lack of situation awareness	Human-made	Unreachable Destination/ Damage to Robot	Rollback-Roll forward, Compensation	Marginal
	Incomplete Input	Rescuer out of the field of view	Human-made		Rollback-Roll forward	Marginal
	Delayed Input	Delayed/ Disrupted Communication	Hardware		Rollback- Roll forward	Marginal
	No Input	Camera doesn't Work	Hardware		No Recovery: repair action required	Critical
	Paradox	Ranger/Proximity Sensor Fault	Hardware		Rollback- Roll forward, Isolation	Marginal

capable of taking a drink order from a customer, fetching the correct drink and finally delivering the drink to the customer. In defining the Robot Waiter task specification a number of assumptions were made about the robots design and operating environment. These assumptions are as follows:

In order to maintain consistency between different design studies, these assumptions should be carried over to future work. The following section discusses the functional design of the Robot Waiter task (Table 6). The HTA results for the Robot Waiter task are included in Extension 1 to the online version of this paper. The hierarchical decomposition of the robot's tasks in textual form is provided in a tabular form in Extension 2. This table starts from the top level Task 0 "Deliver Ordered Drink to Customer". This top level task is achieved by performing the sub-tasks of waiting in the waiting location and scanning the room for a customer, attending the customer to take a drink order, getting the requested drink from the bar, delivering the drink to the customer, and then asking the customer if everything is satisfactory. The analysis also considers some of the principal error situations that may occur in performing this service, such as where the requested drink is unavailable at the bar, or if the customer is missing when the drink is delivered. Each task is assigned a Behaviour Type, which classifies the task according to the NASA Goddard Agent Architecture Model [33] – see Appendix B and Table 14. This model has been used to identify the nature of the cognitive processes that are required in order to perform the task. This model allows other design analyses such as preliminary functional failure / hazard analyses to be performed without requiring explicit details about the implementation, which are not available at this stage of development.

4.2 Robot Waiter Functional Architecture Model

The functional architecture of the Robot Waiter was developed by a three-step procedure:

a) Identify the Behaviour Type of each task, as defined in the NASA Goddard Agent Model (see Table 14)
b) For each task, identify the cognitive processes employed within the task, as implied by the task behaviour type and the relevant processes for that type as shown in Figs. 9–16 of Appendix B.
c) For each cognitive process, identify any essential parameters or global variables used by the process, any special hardware required, and the data flow across the boundary of the process (the interface).

Table 6 BRL Robot waiter study - design assumptions

Category	Assumptions
Mechanical assumptions	• The robot will have only one manipulator for carrying drinks.
	• The robot will transport drinks in an internal compartment.
Environmental assumptions	• All drinks to be served will be placed in specific areas on a table surface (the bar), which are pre-determined and known by (programmed into) the robot.
	• The environment is a single-storey flat surface with no stairs to be climbed.
	• An area of the environment is reserved as a waiting location while the robot is not serving customers.
	• A number of specific environments were envisaged for the robot: ◦ A laboratory lounge area ◦ A restaurant ◦ A bar ◦ A demonstration area of a robotics conference ◦ At home
	• It is assumed that drinks will be provided in the following types of container: a) A stiff polystyrene cup, of cylindrical or inverted (upside-down) conic section profile, with a lid attached to the top and without any handles b) A near-cylindrical plastic bottle (e.g. mineral water bottle) with no handles
	• It is assumed that bar tables will have their own drainage to capture spilled drinks, or that any such spillages will be promptly cleaned up by bar staff. It is assumed that spillages at the bar table will not leak onto the café / restaurant main floor.
Operational assumptions	• The robot will only have a drinks serving (waiter) role; drinks preparation (bartending) role is outside the scope of this design. It is assumed that requested drinks will be prepared and placed into the correct areas on the bar by another agent – the bartender – who may be human or artificial.
	• The robot will take an order, transport and serve a drink one at a time.
	• The robot will wait to be called (reactive), not to offer drinks proactively.
	• The robot may optionally hand over drink to customer, place drink on a table, or leave drink on tray.
	No special behaviour is required for particular drinks, for example if they were to be served in different mugs, cups and saucers, or other types of drink container. It is assumed that all types of drinks to be served can be handled in the same manner, and that no special behaviour is required because a drink is hot, cold, or unusually delicate in some manner.

The result of this design step was a large task-process model, which is provided in Extension 3 to the online version of this paper.

4.3 Hazard Analysis Methodology of the Experiment

The hazard analysis of the robot waiter design model proceeded as a set of six sessions over the April – June 2011 period. The authors were the participating team for all of the sessions. The procedure adopted for the analysis was to use the SHARD guideword set listed in Section 3.1.2 and work through the Task-Process Model of the Robot Waiter applying the SHARD guidewords to the task description. Causes of any plausible hazards were identified as functional failures of the Goddard reference architecture elements that were relevant to the task as defined in the Task-Process Model.

The SHARD method was selected because it has both function-oriented and interface-oriented aspects, and since the functional architecture model described in Section 4.2 contains elements of both types of model, it was considered to be the most appropriate. The SHARD guidewords shown in Table 4 were used in the analysis.

The analysis proceeded in a typical manner for this type of analysis, with the team discussing each element of the model in turn and assessing the potential consequences of its failure. The consequences were logged in a hazard analysis table, a fragment of which

Table 7 Sample fragment of preliminary hazard analysis table from BRL robot waiter design study

Model element	Failure type	Failure description	Operating phase	Consequence description	Cause description	Corrective action (design only)	Design recommendations/ safety requirements
Task 3.2 Pick Up Drink Pick up one example of the requested type of drink (and put it in the storage compartment)	Omission	Arm fails to move	At Bar	Loss of service; no safety effect	–	–	Assumptions: • Drinks provided in stiff plastic/ polystyrene cup, with lid attached to cover the top, or will be (near-) cylindrical plastic bottles (e.g. mineral water bottles) • Drinks cups will not have handles
		Arm fails to move wrist to correct table location	At Bar	Robot knocks over other drinks, causing spillage on bar table			Assumption: • Bar table has drainage or will prevent spillages from leaking onto floor
		Arm drives into table surface	At Bar	Damage to arm table; other drinks knocked over (on bar table) Subsequent hazards may occur due to damaged arm	• Execution: controller fault/error • Perceptors: sensor faults • Modelling & State: errors in world mapping or object (drink cup) mapping • Effectors: arm motor faults	Provision of proximity sensors on arm	Robot shall use proximity sensor positioned about its arm, to detect potential collisions with table surfaces If robot drops a cup at bar table, then it will repeat task afterwards
		Arm drives into drink cup/bottle	At Bar	Spillage of one or more drinks on bar table	–	–	Arm trajectory design (sliding motion not chopping motion)
		Gripper fails to move	At Bar	Drink knocked over (on bar table) if arm subsequently moves sideways	–	Continuous contact detection and real-time monitoring; if no contact then arm must reverse its trajectory (exit strategy)	

Table 7 (continued)

Model element	Failure type	Failure description	Operating phase	Consequence description	Cause description	Corrective action (design only)	Design recommendations/ safety requirements
		Gripper fails to grip cup with sufficient strength	At Bar	Drink slides out of the gripper, causing spillage over bar table	• Execution: controller fault/error • Perceptors: sensor faults • Modelling & State: errors in world mapping or object (drink cup) mapping	Redundant or independent pressure sensing separate from task controller Use of conically-shaped cups	
				Delayed drink cup slippage will cause spillage on floor	• Effectors: gripper motor faults	Use of deformable soft-touch sensors Gripper force feedback detection	
		Gripper fails to grip cup in appropriate position	At Bar	Cup spins within grip, causing spillage on bar table		Gripper design	Robot shall employ a gripper design that prevents the cup spinning within the robot's grip (e.g. four-fingered gripper)
		Robot fails to open storage compartment door	At Bar	Robot arm smashes into door, causing spillage onto floor and possible damage to robot (sharp pieces on floor)			If robot is away from bar table, it shall stop and indicate spillage after cup is dropped or knocked over
		Robot fails to put drink inside storage compartment	At Bar	Spillage of drink over robot and floor	Drink already stored		

is shown in Table 7. Since functional hazard analysis is very time consuming, a complete analysis (all keywords applied to all model elements) was not performed, only a subset sufficient to demonstrate the method.

4.4 Discussion of the Results

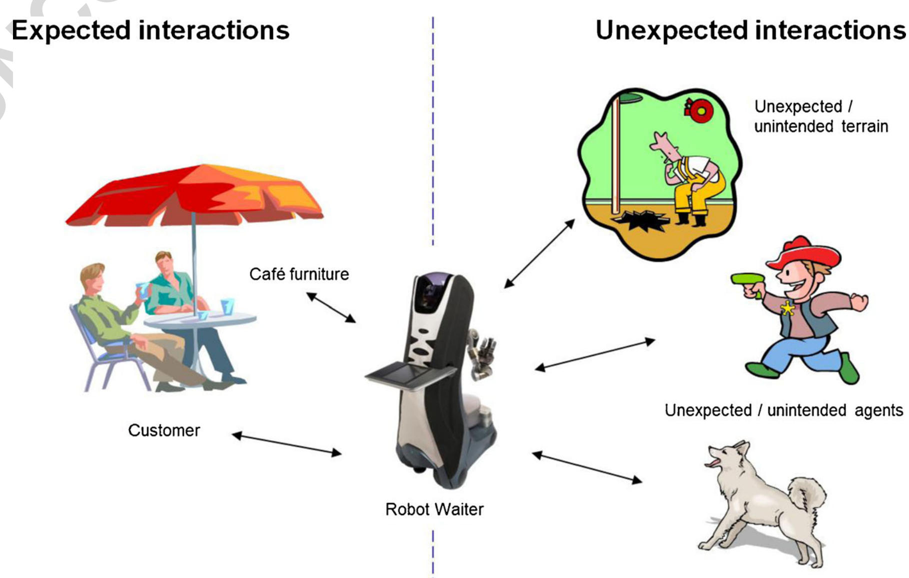
Table 7 provides a representative sample of the results that were generated in the analysis sessions. In many respects, this is similar to the kind of results that are achieved in similar analyses of non-robotic systems and as it stands the results show that this kind of analysis can yield useful safety requirements. However, the results themselves do not reveal the issues that drove the research described in this paper, which emerged from the flow of the discussions that formed the process itself.

As the analysis sessions proceeded, it became apparent that the analysis guide words were not directing the team discussion in the manner intended; the failure conditions of individual elements of the model became less significant in the discussion than the identification of the circumstances of the robot's situation in its environment and the features of the environment with which the robot must interact. It was very difficult to determine the exact consequences of a robot's action and their severity until it is known with what the robot might be interacting.

For example if a robot moves across a room at high speed, either due to its control system or due to a motor failure, there may be the potential for a collision with some object in the environment. However, the precise consequences and the severity of those consequences will depend on what collides with the robot. If the object is a chair or a table, then the consequence (a damaged table or chair knocked over) is not particularly severe. If the object is a person, especially a child, then the consequences are significantly higher in severity and it may be necessary to design safety features into the robot to reduce the risk of this occurrence.

During the analysis, it became clear to us that the guide words being used for the analysis were not encouraging the team to consider different types of environmental interaction. The guide words were applied to elements of the internal design of the robot, albeit at an abstract level, and were effective in identifying a comprehensive range of internal errors, but did not assist with the identification of external features with which the robot might interact in its intended environment. The only external features that were mentioned were those that were inherent to the robot's intended mission, which had been identified in the tasks developed in the hierarchical task analysis design process. Other features that can plausibly be considered to be present at least occasionally are not mentioned, and there is a very real risk that the analysis process may overlook potential hazards that are

Fig. 2 Types of interactions for autonomous systems



reasonably foreseeable, which may lead to accident risks not being reduced to acceptable levels. Furthermore, the apparent completeness of guide word sets such as SHARD and HAZOP may mislead manufacturers into believing that their hazard assessment is as complete when it is not, which could have serious implications for their liability and for the risk to the public of their products.

The conclusions reached by the team during this initial trial study suggested the concept that while the team had specified those tasks that were required of the robot to perform its intended duty, there were potentially a lot of tasks that may be required of a robot simply to exist in its environment and survive long enough to be available to perform its intended tasks without causing any undesirable situations or unacceptable accidents.

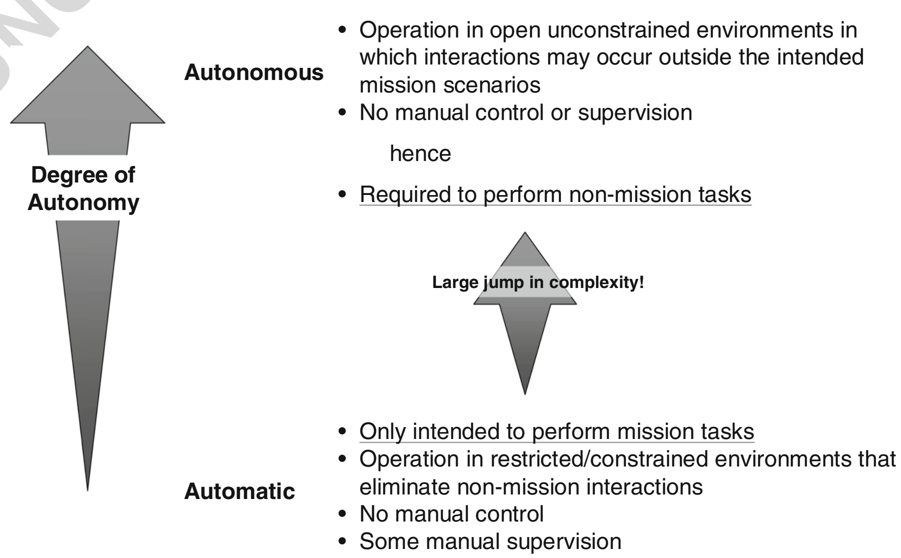
This revelation led us to define the concept of *mission tasks* and *non-mission tasks*, as illustrated in Fig. 2.

Mission tasks are defined as those task required for the robot to fulfil its intended function or mission, which are typically identified by design processes such as hierarchical task analysis or similar methods. Mission tasks handle the *expected interactions* of the robot with its environment – those that are likely to occur in most instances of its mission.

Non-mission tasks are those tasks other than mission tasks that are necessary to allow a robot to 'survive', i.e. to maintain its state of operational readiness whenever a mission is not in progress or to perform a task at any time that prevents the occurrence of hazards (or reduces their risk). Non-mission tasks handle the *unexpected interactions* – those that are reasonably foreseeable but not expected to occur often.

The proliferation of non-mission interactions in comparison to the mission interactions, which were identified by the team in BRL Robot Waiter hazard analysis sessions, led us to understand that the non-mission tasks may well comprise the great majority of the robot's functionality or behavioural repertoire. It also led to the idea that the ability to cope with non-mission interactions may be a defining aspect of the difference between an automatic and an autonomous system. Automatic systems are designed to perform mission tasks without human intervention, but do not include any provision within their design for handling non-mission interactions. These are handled either by designing the environment of the system to exclude the possibility of any interactions other than those related to its mission, or else humans remain in the system in a supervisory mode, handling or preventing any non-mission interactions while the automatic system performs the mission task(s). Industrial machines and automatic (driverless) railways are good examples of this concept. In contrast, autonomous systems have no human control or supervisory input whatsoever, and are generally expected to operate in environments that have not been pre-prepared for its operation. Robot waiters in cafes and wheeled rovers on other planets are good examples of this concept. Thus, the mission vs. non-mission task classification concept

Fig. 3 Comparison of automatic and autonomous systems



offers an intriguing insight into what the differences are between these classes of system.

This relationship between the categories of automatic and autonomous systems can also be seen as defining a *degree of autonomy* measure, at least in a qualitative sense, as represented in Fig. 3. The more non-mission interactions a system is required to handle by itself without any human intervention or without prior preparation of its environment, the greater its degree of autonomy.

Non-mission interactions are what makes the hazard analysis of autonomous agents (such as mobile robots) more difficult than conventional systems - it requires an additional analysis step to identify the non-mission interactions of an autonomous system as a necessary first step before proceeding to identify hazards derived from internal failures in the traditional manner. Since there may well be many more non-mission tasks required of a robot than mission tasks, this additional step becomes the dominant design/analysis activity in the development of a robot. The increased effort required for the design of non-mission tasks will make the development process of the robot more expensive than an equivalent automatic system with manual supervision, and the determination of the most appropriate level of automation will be a crucial design decision having a significant effect on a system's development costs and timescales and its operating costs.

Hazard analysis methods intended for identifying potentially hazardous non-mission interactions and defining safety requirements must therefore provide a systematic method for identifying potential hazards associated with non-mission tasks, when those tasks may not be defined in the robot's functional requirement specification. Therefore, new methods, or variations on existing methods, are needed to fill this gap and provide a more effective method for performing preliminary hazard analysis of autonomous systems such as mobile robots. The method we propose is called Environmental Survey Hazard Analysis, which is described in Section 5.

5 Environmental Survey Hazard Analysis

In this section we propose a new variant of hazard analysis, called Environmental Survey Hazard Analysis (ESHA), which is intended on identifying non-mission interactions and the potential hazards that may be associated with them, as a preliminary hazard analysis exercise that should be performed prior to the more traditional internally focused hazard analysis exercises that are typically performed for conventional non-robotic systems [18].

5.1 Objectives of New Method

As discussed in Section 3.1, the objective of any hazard analysis method is to provide an objectively demonstrable basis for demonstrating that all reasonably foreseeable hazards have been identified. This must also be the objective of any method that seeks to identify hazards associated with non-mission interactions. The method must provide a classification framework that can be argued as providing complete coverage of the range of foreseeable non-mission interactions at some level of abstraction, and since it is not practicable to identify every instance of any foreseeable interaction in any possible robotic application in or operating environment, a classification scheme is necessary at a higher level of abstraction, which provides full coverage of the abstract model but leaves it to the human analysts to supply all reasonably foreseeable examples of each category for the target application and environment. However, this criterion in and of itself does not offer any guidance as to what the hazard classification scheme should be, and therefore any such choice will be arbitrary with respect to the above objective. Therefore it is necessary to draw on other ideas to provide the framework.

Our current proposal is based on an abstract model of the situated-ness of a robot in its environment. An autonomous mobile robot is an agent embedded in its environment, perceiving the world through its sensors and taking action using its effectors (motors, manipulators etc.) to change its state or the state of features in the external environment. One way to classify features of the environment, in a manner that may be convenient to the design of safety mechanisms, could be to classify them abstractly in terms of size or shape *as perceived by the robot through its sensors*. Therefore, instead of classifying hazards based on the precise identity of particular features, which would lead to an open-ended list, we propose to classify them in terms of abstract properties that we can be certain cover all possible features.

Given this frame of reference, we argue that the entire environment perceived by the robot through its sensors can be divided into the following categories:

- **Environmental Features**: these are features associated with the background environment itself, rather than any object situated within it, and their state is fixed to the frame of reference of the environment.
- **Objects**: these are features that are embedded or situated within the environment, but are assigned their own distinct identity and state, and are often assigned their own frames of reference.

We argue that everything in the environment can be considered either a background feature or an object, and thus this level of classification is complete.

Background environmental features can be further sub-divided into invariant and varying features, the former including *terrain features* and the latter including *ambient conditions*. Terrain features describe features of the structre or configuration of the environment itself (i.e. not with any object situated in the environment) that generally remain fixed or constant during the operation of the robot. These include geographic areas, for example “urban”, “indoors” or “marine”, particular types of surface such as “paved road” or “grass” or terrain features such as ‘lakes’ or ‘pathways’. Variable environmental features do change over time, the most common of which are *ambient conditions*, such as temperature or pressure.

We have classified Objects by means of several abstract properties. One obvious abstract property of an object is its *shape*. To provide a classification that covers all possible shapes, we have proposed a set of categories based on the dimensionality of their shape – point-like (0D), linear (1D), surface (2D), and volumetric (3D). Everything in the environment that has a shape will fall into these categories. A second property we have used is *motion*. Objects may either be stationary or moving; the former may either be immovable (fixed in place) or may be movable, either by the robot itself or by the action of others. The third property we have used is *agency*, which is considered for moving objects, in which we consider whether an object is moving purposefully or not.

In all these categorizations, we have applied wherever possible logically exclusive definitions, so that the hazard analysis guidewords derived from them cannot admit any other possibilities. This means that by following the guidewords human safety analysts are assisted in achieving the aim of identifying all reasonably foreseeable hazards, because the logical structure of the classification is complete.

While it must be admitted that the choice of classification is arbitrary, it is guided heuristically by an understanding of the domain problem. One of the aims of this research is to assess whether the classification scheme is useful in guiding human analysts towards an effective identification of environmental interactions and their potential hazards. If the proposed classification was unhelpful in this respect, we should expect to receive feedback from analysts claiming that it was difficult to apply the guidewords constructively, and that the guidewords hindered them from thinking clearly about the problem. The discussion in Section 6 describes the feedback we have received so far from our experiments to date.

Following the above argument, the ESHA classification scheme is shown in Fig. 4, in which all of the categories mentioned above are integrated together.

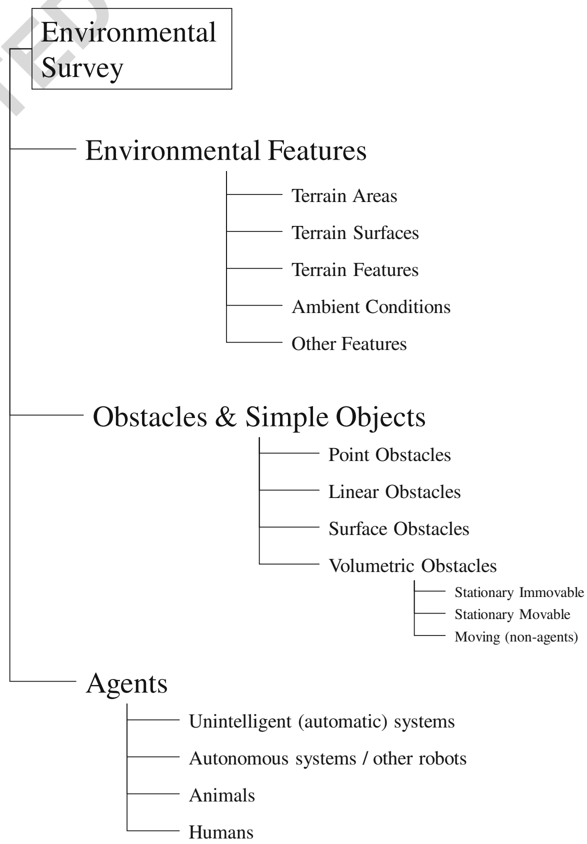


Fig. 4 Classification scheme used in environmental survey hazard analysis

Q5 **Table 8** Environmental survey hazard analysis – standard worksheet template

Ref. No.	Object: (Environment feature/obstacle/agent)	Interaction details	Interaction failure type/keyword	Interaction failure details	Consequence	Safety measures

The initial classification of environmental features combines the basic feature types with the complexity of their behaviour, dividing the complete environment into three possible classes:

- Environmental features – these are invariant, large-scale and semi-permanent features of the environment that provide the reference frame within which other objects exist.
- Obstacles and Simple Objects – these are objects that are situated within the framework of the static environmental features described above, which may be fixed, movable, or even actively moving, but whose behaviour is not goal-directed in any way, i.e. their behaviour cannot be defined as purposeful in any way.
- Agents – these are objects that are moving in the environment in a purposeful way, i.e. their behaviour is goal-directed.

This classification of features maintains its logical completeness as discussed in previous paragraphs, and requires no default alternate category to do so (as is done for Environmental Features, as discussed below).

For the Environmental Features category, we have defined the following principal sub-categories: terrain surface types, terrain areas, terrain features, and ambient conditions. The argument is that the robot will perceive the world as one or more different areas, each of which has a given type of surface and contains a set of terrain features and ambient conditions. Since this classification scheme is not logically closed, we must admit to the possibility of other types of environment feature that do not fall into the secondary scheme; therefore we have added a default secondary category that covers all features not covered by the first four. This closes the logical completeness of this level of the classification, and although it does not provide positive guidance to analysts it will at least remind them that they must consider other possibilities and encourages analysts to search for any exceptional features that are not covered by the initial classification.

For the Obstacles and Simple Objects category, we have defined four shape/structure categories that reflect how these features may be perceived by a robot: Point Obstacles (0-D), Linear Obstacles (1-D), Surface Obstacles (2-D) and Volumetric Obstacles (3-D). We argue that all objects in the environment will be perceived by the robot as having a shape or structure that is point-like, line-like, surface-like or will have a perceived volume. Therefore, by encouraging analysts to search for features that have these shape characteristics, we argue that they will search through all reasonably foreseeable features within the target environment. Since this is a logically closed classification it does not require any default category called "other types" or similar. We have also further sub-divided the volumetric obstacles into a further sub-category based on whether its movement can be influenced by the actions of the robot: Stationary Immovable (i.e. obstacles that cannot be pushed out of the way), Stationary Movable (obstacles that can be pushed out of the way by the robot or due to other actions) and Moving (obstacles that do move, but not in any purposeful way i.e. they are not agents).

For the Agents category, we have defined four categories that capture the full range of behaviour patterns that any agent may exhibit, which is perceived by the robot. The secondary categories are: Automatic Systems (performing mission tasks only), Autonomous Systems and Other Robots (which perform both mission and non-mission tasks), Animals (autonomous biological creatures exhibiting purposeful but non-sentient behaviour) and Humans (autonomous biological creatures exhibiting purposeful and sentient behaviour).[1]

These classification categories are being tested in on-going design studies and trials at Bristol Robotics Laboratory, the first tranche of which are reported in Section 6 of this paper. It is anticipated that the classification scheme and the associated guide words (see Section 5.2) will evolve over time depending on how useful they are in guiding analysts in the systematic identification of non-mission interactions and tasks. As discussed in Section 7, it is anticipated that

[1]Until the existence of other sentient species is proved, we consider humans to be the only category of autonomous biological creatures exhibiting purposeful and sentient behaviour, and hence no other species need be named in this category. The subcategories of agents are only developed for the purposes of our classification and have no authority for any other purpose.

the classification scheme may evolve significantly as different classes of robotic applications are studied or developed.

5.2 Procedure of New Method

For the trials described in Section 6, we developed a set of aids for performing an ESHA analysis:

1. An ESHA Procedure Checklist, which contains the classification categories mentioned in Section 5.1 above, and provide non-exhaustive lists of examples as an aid to the analyst(s). The checklist contains a number of questions designed to guide the analyst(s) in thinking through the application of the ESHA classification guide words as shown in Fig. 4. The checklist is provided in the text boxes on the following three pages.
2. A generic ESHA worksheet (shown in Tables 8 and 9) which provides a tabular format for recording the results of the analysis. It is similar in layout to Table 1, but the column titles are aligned to the output of the ESHA procedure information.

The full worksheet template and checklist have also been provided as Extensions 4 and 5 to the online version of this paper.

The Procedure Checklist consists of three parts, for Environmental Features, Obstacles and Simple Objects, and Agents. Each part comprises a series of steps, characterised by questions, in which the classification scheme mentioned previously in this section is applied to identify potential environmental interactions (mission and non-mission related), and then to determine whether the interactions have potential hazards and to identify possible safety measures that may reduce or eliminate the risk of those hazards. These safety measures would then become system safety requirements for the robot, to be incorporated into its design.

The standard Worksheet Template is matched to the Procedure Checklist, and is intended to provide a tabular format for recording the results of the assessments and decisions of the hazard analysis process, so that they can be reviewed afterwards for the purposes of safety assurance, or to repeat/revise the results if necessary.

The checklist and worksheet template have been applied in some (but not all) of the experiments conducted to date, and the assessment of that work is discussed in Sections 6 and 7.

Table 9 Fragment from environmental survey hazard analysis worksheet – INTRO project 3rd workshop – robot waiter demonstrator

Ref. No.	Object: environment	Interaction details	Interaction failure type/keyword	Interaction failure details	Consequence	Safety measures
	Water, liquid or broken glasses on the Floor	Moving on the floor	Slipping		The robot could fall over: hazard	Travel slowly, sensor that can detect irregularities on the floor coupled with a system that can avoid them
			Losing your point in space	For odometry in navigation	Inaccurate localization: loss of function	When the robot stops then it must always recalibrate, sensor that can detect irregularities on the floor coupled with a system that can avoid them
	Doorstep	Go past doorstep problems	Robot falling		Hitting people: hazard Damage property: damage	Set up an environment without small steps
					Robot sensors could get damaged and that could later become an hazard	Include in the robot design a sensor that at the floor

ENVIRONMENTAL SURVEY HAZARD ANALYSIS PROCEDURE

1: Analysis of Environmental Features

Are there any specific examples of the following features of the environment in which the robot is intended to operate?

- What specific areas exist in the environment?
  - e.g. Interior: rooms, corridors, stairs, elevators, escalators, slide-ways
  - e.g. Exterior: lawns, sidewalks, roads, fields, woods/trees, scrubland (low vegetation), marshland
- What types of terrain surface? (e.g. Interior: floor surface types)
  - e.g. Exterior: terrain types: paved, grass, mud, sand, gravel, rocky, water, paved/unpaved paths
- What types of terrain feature?
  - e.g. Interior: walls, doors, windows, barriers, prohibited areas)
  - e.g. Exterior: barriers, fences boundaries, prohibited areas, flower beds, trees, ponds
- What ranges of ambient conditions?
  - e.g. Lighting levels, air temperature
  - e.g. Special conditions such as steam/water vapour, snow/ice, smoke/fire, corrosive atmosphere, salt atmosphere/spray
- Are there any other features not yet identified?

For each environmental feature, identify how the robot should interact with it.

- What should the robot do? (e.g. approach / avoid / track / manipulate)
- What are the characteristics of the interaction? (e.g. short or long range, immediate or delayed response, reflexive, deliberative, reactive, social/communicative)

For each interaction, what could go wrong?

- Failure to interact when intended?
- Inadvertent interaction?
- Partial interaction?
- Reverse interaction?
- Actions taken are too much, too little, more than required, less than required?

For each interaction failure, what are the consequences?

- Injury
- Damage to property
- Damage to the environment

For each consequence, what measures can be taken to reduce the likelihood of the consequences?

- Inherent safety measures (re-design the robot to eliminate the problem)
- Safeguards or protective devices (protection systems)
- Instructions to robot users (less likely if the robot is fully autonomous)

2: Analysis of Obstacles and Simple Objects

Are there any specific examples of the following obstacles or simple objects in the environment in which the robot is intended to operate?

- What types of Point Obstacles are there in the environment?
  - e.g. light/heat/sound/odour sources
  - e.g. linear or volumetric obstacles viewed from a long distance
  - Interior: clutter objects (at far range), light sources (e.g. lamps)
  - Exterior:
- What types of Linear Obstacles are there in the environment?
  (boundary lines/edges, vertical posts/pillars, volumetric obstacles viewed edge-on from a distance)
  - e.g. Interior: power cables, carpet edges, doorsteps, staircase edges
  - e.g. Exterior: kerbs, barriers/fences, paving-stone ruts ('crazy-paving')
- What types of Surface Obstacles are there in the environment?
  - e.g. Interior: surface spills {water, detergent, foodstuffs, domestic chemicals}, open trapdoors
  - e.g. Exterior: surface water/flooding, ice patches, surface spils {oil, detergent, fuel, chemicals}, manholes, trenches, ramps, drains, safety mirrors/reflectors
- What types of Volumetric Obstacles are there in the environment?
  - Stationary-immovable
    - e.g. Interior: permanent furniture, food/drinks machines, power cables, tape / stretch / rope barriers, cones, furniture (tables, chairs, desks, office furniture), staircases, large tables, desks, beds, domestic furniture, bathroom furniture, cookers, washing machines, fires/fireplaces, computer equipment cabinets, food/drink machines, photocopiers
    - e.g. Exterior: shelters, road works, garden benches
  - Stationary-movable:
    - e.g. Interior: clutter objects (at close range), chairs, small tables
    - e.g. Exterior: tape/stretch barriers, bollards/cones
  - Moving (non-agents)
    - e.g. Interior: toys, trolleys, moving decorations (e.g. wind-chimes, hanging sculptures, childrens' mobiles), ventilation fans,
    - e.g. Exterior: sliding doors, giant folding doors, turnstiles,
- Are there any other features not yet identified?

For each obstacle, identify how the robot should interact with it.

- What should the robot do? (e.g. approach / avoid / track / manipulate / other?)
- What are the characteristics of the interaction? (e.g. short or long range, immediate or delayed response, reflexive, deliberative, reactive, social/communicative)

For each interaction, what could go wrong?

- Failure to interact when intended?
- Inadvertent interaction?
- Partial interaction?
- Reverse interaction?
- Actions taken are too much, too little, more than required, less than required?

For each interaction failure, what are the consequences?

- Injury
- Damage to property
- Damage to the environment

For each consequence, what measures can be taken to reduce the likelihood of the consequences?

- Inherent safety measures (re-design the robot to eliminate the problem)
- Safeguards or protective devices (protection systems)
- Instructions to robot users (less likely if the robot is fully autonomous)

1120

3: Analysis of Agents

Are there any specific examples of the following agents in the environment in which the robot is intended to operate?

- Will there be any unintelligent systems in the environment?
  (e.g. Vehicles, automatic systems)
- Will there be any other autonomous systems or robots in the environment?
- Will there be any other animals (living, non-sentient) in the environment?
- Will there be any humans in the environment?
  - Maturity: child / adolescent / adult / elderly
  - Strength: stronger / weaker / handicapped
  - Height: tall / short
  - Weight: light / heavy / very heavy (i.e. obese)
  - Gender: male / female *(although this is not foreseen to be a likely issue)*
  - Impairment?
    e.g. vision, hearing, touch, taste, smell (olfaction), thermal sense, balance, manipulation ability speech, *others?*
  - Intelligence
    - Literacy: literate / illiterate / non-native language or alphabet / dyslexic
    - Numeracy: numerate / innumerate / dyscalculic
    - *others?*
  - State:
    - Conscious/Unconscious
    - Movement: stationary/crawling/walking/running/jumping
    - Attention Level:
    - attentive (to the robot/system and its situation),
    - distracted (not attentive to any external object or situation),
    - focussed elsewhere (attentive toward other object or situation)
- Are there any other agents not yet identified?

For each environmental feature, identify how the robot should interact with it.

- What should the robot do? (e.g. approach / avoid / track / manipulate)
- What are the characteristics of the interaction? (e.g. short or long range, immediate or delayed response, reflexive, deliberative, reactive, social/communicative)

For each interaction, what could go wrong?

- Failure to interact when intended?
- Inadvertent interaction?
- Partial interaction?
- Reverse interaction?
- Actions taken are too much, too little, more than required, less than required?

For each interaction failure, what are the consequences?

- Injury
- Damage to property
- Damage to the environment

For each consequence, what measures can be taken to reduce the likelihood of the consequences?

- Inherent safety measures (re-design the robot to eliminate the problem)
- Safeguards or protective devices (protection systems)
- Instructions to robot users (less likely if the robot is fully autonomous)

6 Trials of Environmental Survey Hazard Analysis

Having developed the initial ESHA method proposal, which we believe offers an improved assessment of mobile autonomous robot applications, we set out to evaluate the new method on further robotic application studies. This section provides an overview of the results collected.

By fortunate coincidence, at the time the proposed ESHA method was being developed, the INTRO project was in the process of developing the initial requirements and specifications for its demonstrator projects. This offered an opportunity to test the new method on the demonstrator, and at a workshop at BRL in 2011 we held two sessions in which we used Environmental Surveys to identify conceptual hazards that might be associated with the application requirements that the INTRO project was developing as design studies for the two demonstrator projects.

In addition to the INTRO demonstrator projects, two Postgraduate (MSc) Dissertation studies were performed in 2012 into safety analysis and design of robotic applications. One project (the USAR Robot study) was a precursor to further work to be done within the INTRO project, while the other (the Guide Assistant Robot) was developed as an entirely independent study.

Section 6.1 provides the description of the application of ESHA to the Robot Waiter scenario. Section 6.2 reviews the work done on the Urban Search and Rescue (USAR) application study, and finally Section 6.3 reviews the study into a Guide Assistant Robot application. Each section discusses the task requirements of the application, the (partial) ESHA exercises that were performed and presents the results that were obtained.

6.1 Application Study #1 – The Robot Waiter

The Robot Waiter scenario described in chapter 4 aims to demonstrate the behaviour of an intelligent robotic system that functions in close interaction with humans in a cafe, which is a partially unstructured and dynamically changing environment.

In this scenario, characteristics such as autonomy, an intelligent interface, high-level sensing abilities, a safe manipulator arm, visual pattern recognition and knowledge extraction in order to learn about the robot's environment, are key to achieve an efficient human-robot interaction and cooperation.

During the September 2011 INTRO Workshop, held at Bristol Robotics Laboratory (BRL), a trial of Environmental Survey Hazard Analysis (ESHA) was conducted for the first time with participants other than the authors. The general aim of the overall process is to merge the results of ESHA with the aforementioned Hazard Analysis results. The traditional Hazard Analysis would take care of the potential hazards in mission tasks caused during a system's operation in its environment, while the Environmental Survey would identify the non-mission aspects of extended operation.

In the practice session, a four-person group applied an especially drafted form for ESHA. After the tutorial a discussion session was conducted in order to collect the participants' opinions on the usefulness of the approach. The practice session lasted less than 2 hours, so the quantity of work achieved was small, but enough to offer an initial impression of the approach. A sample from the ESHA worksheet produced by this study group is shown in Tables 8 and 9. Q7

The Robot Waiter scenario was the same as the one described in chapter 4, however, the way the same scenario was approached this time is different since in chapter 4, only the mission tasks were considered, as it happens for a traditional Hazard Analysis, while during these trials the new ESHA was applied to the Robot Waiter scenario, thus all non-mission aspects and the environment where the robot operates were taken into account.

The analysis was effective since participants were able to go over multiple possible hazard scenarios involving the robot and environmental elements. The safety requirements identified for both the robot and the environment were numerous, and it was clear that many more could have been made during a longer trial.

However, the participants commented that better guidance is needed in the order to ensure that each row of the hazard analysis table must be filled. The possible resulting confusion increases the chance that parts of the analysis may be overlooked. During the trial, in order to complete the survey, guidance from the authors was necessary. In addition, the "Interaction Failure Details" column in the ESHA form was not taken in consideration by the participants, who would find that field hard to fill. Furthermore, it was necessary to explain that the "Interaction Details" column refers to normal operational times. These comments will be considered as the guidelines for a future revision of the ESHA methodology (see Section 7.2).

6.2 Application Study #2 – Urban Search and Rescue Application

In the USAR scenario, the aim is to detect and uncover surface and lightly trapped victims. "Surface" victims are visible and mostly free to move and "Lightly" trapped ones are partially covered by light and small pieces of rubble. The first phase of rescue response, after setting coordinating command centre up, is reconnaissance of affected region to identify cold, warm and hot zone. The INTRO USAR scenario considers human robot collaboration in this phase. Using rescue robots in this phase helps to speed up the search for victim and reduces risks that the human rescuers are exposed to. Additionally, robots can assist in uncovering lightly trapped victims. The search for victims is shared between a human rescuer and an assistant mobile robot. The robot will cooperate with the human in assisting both with the visual detection and the extraction of victims by clearing away the rubble which is trapping them.

The robotic system will include a mobile platform fit for unstructured environments and a standard 6 degree of freedom manipulator. In the USAR scenario, a mobile robot assistant has three main requirements: mobility, manipulation and sensing. Mobility is ensured by the mobile outdoor platform base which is also capable of powering the auxiliary hardware installed on it. Simple manipulation tasks such as pick and place of small and light objects are provided by the manipulator. The sensors positioned on the base include rangers for navigation so that the human-robot team can navigate the ruins in search of victims to extract. A stereo vision camera is also employed for HRI and victim detection.

6.2.1 Application Specification

The scenario comprises multiple tasks. The robot searches the disaster environment controlled by teleoperation. During exploration, visual saliency detection is continuously employed to look for victims' faces and/or movement. In case of a successful detection, the robotic manipulator is pointed in the direction of the victim to inform the rescue worker of the victim's approximate position. At this point, the following robot action depends on the intention recognition cues. Depending on the rescuer's cue, the robot has two possible behaviours. In the case where the rescue worker picks up a piece of rubble and offers it to the robot, the rescuer is indicating to the robot that it must pick up the rubble and deposit it to a suitable place. Then, the robot will get ready to pick up another piece. The robot acts autonomously during this collaboration.

On the contrary, if the human directs the robot with a pointing gesture then the robot independently begins clearing out an area of the rubble. At this point, the robot continues moving the rubble until the victim is free. The robot continues finding and extracting victims until the end of the mission. The state-chart of this scenario is depicted in the Fig. 5.

6.2.2 Results of SAR Robot Hazard Analysis

At the September 2011 INTRO workshop at BRL a tutorial session on ESHA was held, to introduce the INTRO project researchers to the proposed method and to conduct an initial trial that would provide feedback on the usability of the technique. It must be noted that this workshop took place early in the demonstrator project, and the analysis was not performed on the design model illustrated in Fig. 5, which represents a later stage of development. The ESHA worksheet that was developed for the USAR Robot demonstrator in the workshop tutorial is presented in Table 10 and its accompanying notes.

Since the session was a tutorial and the first time that the participants had received any training in hazard analysis, the study group that produced the worksheet did not develop the worksheet precisely as intended in the checklist procedure. Improvement of the checklist guidelines has been identified as an area for further development (see Section 7.1). However, the general feedback from the participants was that the method encouraged them to consider issues that they might not have done before, and the worksheet and its notes show that in the limited time available the study group was beginning to identify aspects of the robot's interaction with its environment and the consequent non-mission interactions.

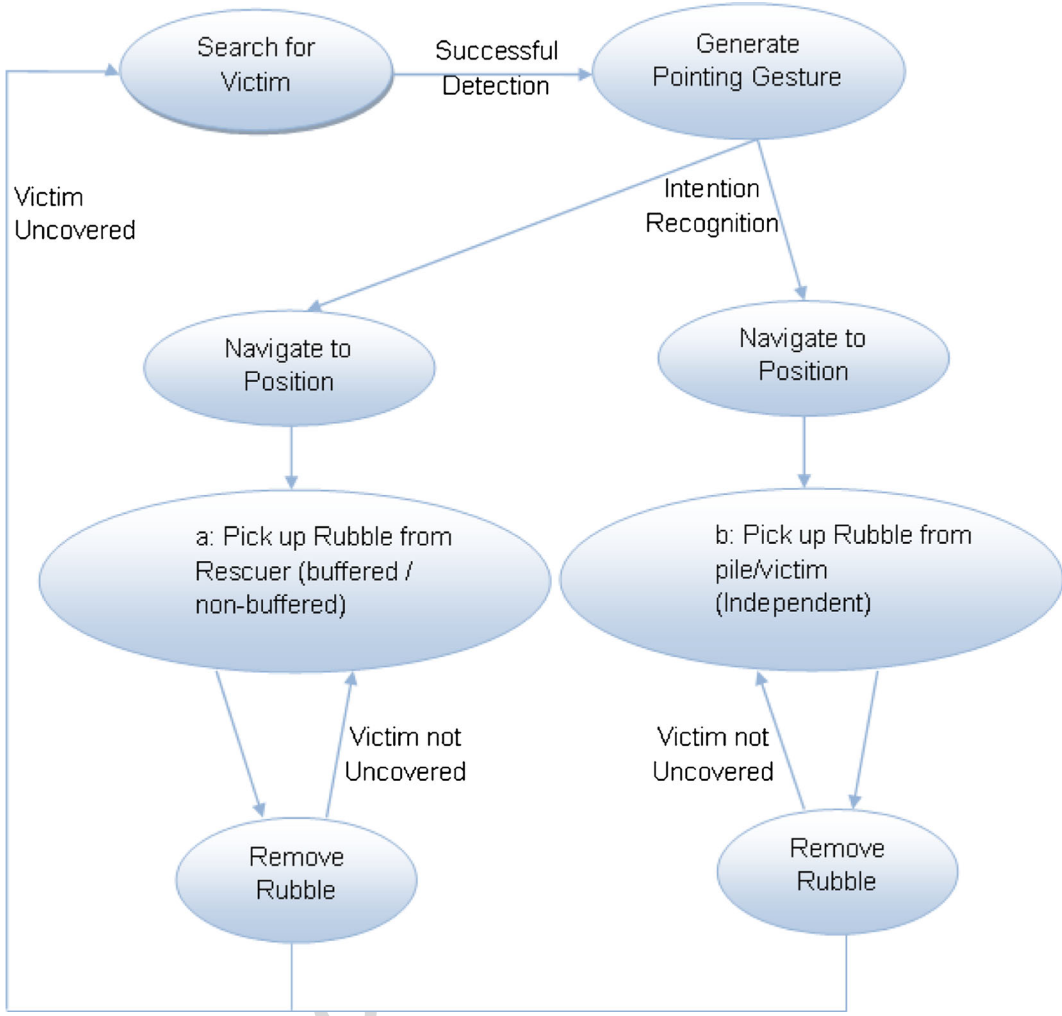


Fig. 5 USAR robot task model

6.3 Application Study #3 – Guide Assistant Robot Application

The third application study of ESHA was an MSc dissertation project carried out by one of the authors at BRL in 2012 [7]. The dissertation was a study on the requirements of a guide robot for elderly persons, in which a task analysis was performed to identify the mission tasks required of the robot, and the ESHA technique was used to identify robot hazards and the safety requirements and non-mission tasks necessary to mitigate their risks.

6.3.1 Application Specification

The basic functional requirement of the Guide Robot was developed as a task model using Hierarchical Task Analysis as the requirements capture method. This produced the task diagram shown in Fig. 6, which is presented in tabular form in Table 11.

The Guide Robot's complete functionality is described by its top level Task 0 "Guide the elderly to the destination". The robot performs this task by means of four sub-tasks: "Waiting for user's call", "Getting user's requirement", "Escorting the user to the destination" and "Finishing the journey". Further

Table 10 Environmental survey hazard analysis worksheet – INTRO project 3rd workshop tutorial – USAR robot example Q8

Object: (Environment feature/obstacle /agent)	Interaction details	Interaction failure type/keyword	Interaction failure details	Consequence	Safety measures
Burning rooms	Approach	Failure to interact	Don't find the fire	Injury Damage to robot	Inherent – temperature measurement
		Too little interaction	Don't move close enough	Injury	Inherent –make robot fire proof
				Injury	User training
		Too much interaction	Moves into fire	Damage to robot	
	Detect fire	Failure to interact	Fails to detect a fire	Injury Damage to robot Fails to warn fire -fighters	
	Detect people				
	Notify/warn				
Edge to vertical drop	Avoid	Failure to interact	Drives over drop	Injury to people below the drop Damage to robot	Terrain scanning Sensors mounted high up on the robot Diverse scanning with sonar, vision, laser, sound, etc. Inherent: hooks on the back of the robot that can grab the surface and avoid a fall Inherent: Explosive bolt at the back that secures the robot and avoids a fall Inherent: Long robot with large mass in the centre to avoid it from falling even if it passes over an edge

Table 10 (continued)

Object: (Environment feature/obstacle /agent)	Interaction details	Interaction failure type/keyword	Interaction failure details	Consequence	Safety measures
Circumstances					
Collapsed building meaning that path planning from old drawings isn't possible					
Wheeled robot with single manipulator					
Fire in the building					
There is a human present to cooperate with the robot					
The robot can lift approximately 7 kg					
The robot can push things					
The robot can do reconnaissance					
Analysis of environmental features					
Specific areas					
Interior: rooms (possibly broken), corridor (possibly broken), stairs (possibly broken), rubble					
Exterior: rubble, streets, garden,					
Types of terrain surface					
Floor, stairs, rubble					
Types of terrain features					
Rough, damaged, uneven, cracks, water, mud, gravel					
Ambient conditions					
Daylight outside and dark inside, sharp contrasts, any kind of light, outside temperature, smoke and fire					
Analysis of obstacles and simple objects					
Point-like obstacles					
Fire, exposed electrical cable					
Linear obstacles					
Stairs, edge to a vertical drop, cables, cracks in the floor					
Surface obstacles					
Collapsed flat objects					

subdivisions of these tasks are described in Table 11. The task analysis only considered essential sub-tasks to achieve top level task and assumed some of the potential error situations that may occur in performing this scenario.

The nominal mission of the Guide Robot is as follows: the robot is intended to remain stationary at a pre-determined standby location, and continuously scan for calls from prospective users of the robot, and when a call is detected or received to go to that user. Once called by a given user, the robot will not be able to accept any other call until the conditions arise where the mission is complete. By returning to a standby location, the robot ensures that it does not block the environment by waiting at the location where its last mission ended. User interactions such as asking a question or getting a user's request are intended to be done by means of a touch screen, or by gesture or speech recognition.

It is assumed that the robot has a built in map of the operating environment (a care home for the elderly) which provides pre-planned paths for given destinations, allowing the robot to plan a journey automatically after confirming the destination from the user.

Escorting and guiding a user to a destination requires the robot to move carefully so as to maintain pace with the user, who may well not be able to move fast, and particular stages of the journey (especially at the start and end) may require the robot to announce

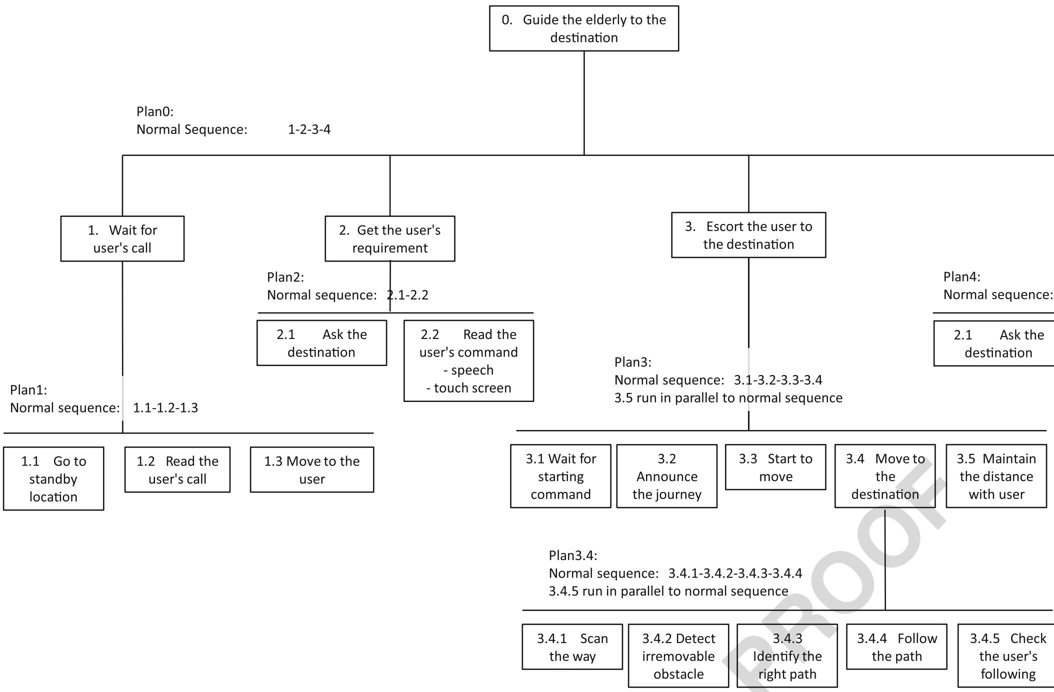


Fig. 6 Guide robot hierarchical task diagram

its intentions so that the user is not confused about the robot's intended behaviour. It is intended that the user places a hand on top of the robot while moving so that the robot can use touch/pressure sensors to detect that it is in pace with the user or when the user leaves the robot (intentionally or unintentionally). As the robot moves it guides the user around obstacles as well as following the planned path.

6.3.2 Results of PC Robot Hazard Analysis

Having completed a basic task specification using HTA, the design was subjected to a preliminary hazard identification analysis using the ESHA technique. However it should be noted that for reasons of practicality this list was developed by the research student as a 'brainstorming' exercise, not by conducting a physical on-site survey of a care home. Therefore, while it was sufficient to develop design and simulation models for the purposes of a student dissertation, it should not be seen as sufficiently or reasonably foreseeably complete for the purposes of a commercial product without being supported by such a direct survey of a target environment. However, the exercise was sufficient to allow an initial overview of the practicability of the ESHA method.

Following the guidelines described in Section 5, a list of Environmental Features, Obstacles/Simple Objects, and Agents to be found in a care home was drawn up by the research student. This list is shown in Table 12. Some of the items in the list were used to develop a set of ESHA worksheets, in which the potentially harmful interactions with those items were identified and a set of safety measures were identified that could reduce their risk (i.e. reduce their severity or probability). A sample of these worksheets is provided in Table 13, and the full set that was developed in the MSc Dissertation is included in an Extension 6 to this paper.

The safety measures in Table 13 and the ESHA worksheets were classified into Inherent safety measures, Safeguards and protective mechanisms, and Instructions to users. This is consistent with the practice of the risk reduction methodologies underlying international standards for industrial and service robots (ISO 10218 [22]). Inherent safety measures are passive constraints or built-in properties of the robot t11.1 **Table 11** Guide robot task descriptions

	Task name	Task description	Task plan (S)
t11.3	0 Guide the elderly to the destination		PLAN 0:
t11.4			• Normal sequence: 1-2-3-4
t11.5	∟ 1 Wait for user's call	Remain stationary and look for a user's call	PLAN 1:
t11.6	∟ 1.1 Go to standby location	∟ Go to standby location and wait there	• Normal sequence: 1.1-1.2-1.3
t11.7	∟ 1.2 Read the user's call	∟ Receive and match signal or sign from user (speech or button to call robot)	
t11.8	∟ 1.3 Move to the user	∟ Approach user close enough to get user's request within safety distance	
t11.9	∟ 2 Get the user's requirement	Obtain an order for the destination of user	PLAN 2:
t11.10	∟ 2.1 Ask the destination	∟ Interact with user to obtain user's requirement	• Normal sequence: 2.1-2.2
t11.11	∟ 2.2 Read the user's command	∟ Receive and confirm the destination from user interface or speech	
t11.12	∟ 3 Escort the user to the destination	Guide user until reaching the destination	PLAN 3:
t11.13	∟ 3.1 Wait for starting command	∟ look for user's starting command (speech and touch on top of robot)	• Normal sequence:
t11.14	∟ 3.2 Announce the journey	∟ Notice the journey to user using voice	◯ 3.1, 3.2, 3.3, 3.4
t11.15	∟ 3.3 Start to move	∟ Start to move with user	◯ 3.5 executes in parallel to normal sequence
t11.16	∟ 3.4 Move to the destination	∟ Go to the destination with user	PLAN 3.4:
t11.17	∟ 3.4.1 Scan the planned path	∟ Scan the planned path	• Normal sequence:
t11.18	∟ 3.4.2 Detect irremovable obstacle	∟ Interact with environment to find irremovable obstacle within range of sensor	◯ 3.4.1, 3.4.2, 3.4.3, 3.4.4
t11.19	∟ 3.4.3 Identify the right path	∟ Confirm the right path	◯ 3.4.5 executes in parallel to normal sequence
t11.20	∟ 3.4.4 Follow the path	∟ Move along the right path	
t11.21	∟ 3.4.5 Check the user's following	∟ Monitor user's following during the journey using touch sensor	
t11.22	∟ 3.5 Maintain the distance with user	∟ Move with estimated walking speed of user	
t11.23	∟ 4 Finish the journey	Finish the journey if robot arrive the destination	PLAN 4:
t11.24	∟ 4.1 Announce the end of journey	∟ Announce the end of journey to user to recognize the destination	• Normal sequence: 4.1-4.2
t11.25	∟ 4.2 Stop moving	∟ Stop moving slowly in order to allow user that is able to stop their following	

Table 12 Examples of environment features

Environment feature	
Specific areas	Bedroom, Bathroom, Living room, Care home common room, Kitchen, Storage room, Corridors, Lifts/Elevators, Staircase
Terrain surfaces	Carpeted surface, Smooth/polished tile floor, Wooden flooring (smooth, varnished)
Terrain features	Walls, Doors (sliding door, normal door, automatic doors, rolling shutter, saloon doors), Windows (full height windows only), Mirrors (full-height mirror, smaller mirrors)
Ambient conditions	Natural light conditions, Artificial light conditions (approximate sunlight (broad spectrum of colours), monochromatic light), Directed / diffuse light source, Air temperature (Room temperature($\approx$ 20C), Hot conditions ($\geq$ 40C), Cold conditions ($\approx$ 5C)), Water/moisture conditions (Fire sprinklers, Fluids spilt on robot (e.g. drinks), Water on floor, Humidity), Wind / air currents (e.g. through open window), Leaking gas, Salt atmosphere (near coasts)
Environment obstacles and simple objects	
Point obstacles	Media Centre / Speakers, Lights & Lamps, Cookers (chemical/odour source), Vacuum cleaners (noise source), Washing machines (noise source)
Linear obstacles	Floor surface area edges (carpet edges, tile floor edges), Vertical furniture items (lamps, potted plants, loudspeakers, coat stands, ceramic vases), Cables for portable appliances, Doorsteps or small steps, Edges of staircases, Edges of holes
Surface obstacles	Pictures & ornaments on walls, Television screens, Water spilt on the floor, Spilt beads/marbles/balls on floor, Detergent (or other slippery surface) on floor, Thick/soft carpets (which are hard to drive over), Recently cleaned surfaces marked by signs, Manholes & trapdoors, Food spilt on floor, Clutter on floor (papers, plastic bags, other objects left on the floor)
Volumetric obstacles	Large furniture (large tables, heavy chairs, bookcases, shelves, other large furniture items, appliances, beds, sofas), Portable items (walking sticks, clutter on the floor), Smaller chairs/tables, Wheeled objects (wheelchairs, trolleys, suitcases, appliances, items mounted on wheeled stands), Movable signs/barriers, Balls/toys, Trolleys/stretchers, Moving decorations, Moving ventilation fans, Waste bins, Things falling off tables
Agents	
Customer	User (attention level, native language, vision, hearing impairment, balance, speech impairment, gesture/manipulation impairment (i.e. can't keep steady hand on top of the robot), walking speed)
Animals	Pets (cats, dogs, birds, rabbits, guide dogs, exotic animals)
Humans	Other people:care home residents (with varying attention level, native language, vision/hearing impairment, walking speed, position: seated/lying down/standing-), cleaners, visitors, care workers, security, supervisors, medical personnel (walking/running speed, attention level), people in wheelchairs, people on stretchers, children ((in-)attention level, walking/running speed, size, position: seated/lying down/standing, non-malicious but deliberate misuse (i.e. playing with the robot)
Autonomous systems or unintelligent systems	Other robots: cleaning robots, other guide robots, robot pets (entertainment robots), mobile domestic servant robots, medical robots, semi-autonomous wheelchairs

Table 13 Analysis of one specific feature - staircase

Ref. No.	Object: (environment feature/ obstacle/agent)	Interaction details failure type/keyword	Interaction failure details	Interaction	Consequence	Safety measures
1	Staircase	Wheeled robot - cannot climb stairs.	Robot fails to notice stairs	Robot try to go forward; Robot recognize stairs as wall	Robot drops on the way downstairs; Robot damaged by dropping Property damaged; Robot damaged by edge of stair; Robot Hits user; Robot falls down; People damaged from running wheel (burning) because of robot's running on same position; Robot avoids stairs but moves around stairs;	• Inherent safety measure ◦ Use of inherently safe materials in the robot's wheel; ◦ Design robot's height higher than stair; ◦ Set up a care home without stairs; ◦ Put caution sign about stair on wall near stair; ◦ Set up a baby gate on beginning of stairs; ◦ Set up a soft cover on edge of stairs; • Safeguards or protective devices ◦ Protective stop function triggered by robot; ◦ Use of touch sensor/bumper on bottom of robot to recognize hit from stair; ◦ Use of compass sensor to recognize robot falling; ◦ Include in the robot design a sensor that points at the floor; ◦ Terrain scanning sensors mounted on ;robot to recognize; • Instructions to robot users ◦ Training user to notice that robot cannot climb stairs;

that ensure that an environmental interaction does not cause harm, such as limitation of motor power or use of soft materials. Safeguards and protection mechanisms are active functions of the robot that take positive action to prevent hazards occurring, for example speed controllers for robot wheelbases or force controller for manipulators. Instructions in the user manuals and guidance notes for users are sometimes required as safety measures when no inherent or safeguard measure can be provided, warning the user to take certain actions in order to avoid possible hazards, for example warnings about when to apply the emergency stop button. Table 13 shows how ESHA can be used to develop safety requirements in a manner consistent with those already found in industry standards. We consider this to be useful in assisting the production of coherent safety requirements specifications for robots.

Although only a partial set of ESHA worksheets were developed in this MSc study, they provide a clear illustration of how the method is to be applied, and these results are currently the most extensive application of the method to date. The results do show the derivation of safety requirements from a systematic review of environmental interactions regardless of their status as mission or non-mission tasks. Therefore, while details such as the ESHA keyword sets may continue to evolve in the future to improve their applicability and coverage, it is clear that an analysis process of this format is able to fulfil the objective of providing a non-mission based perspective on the behaviour of a robot.

The main limitation of this study was the fact that it was the work of a single student and not a design team including domain experts, which is the recommended practice in industry for conducting for system hazard analyses and remains equally valid for ESHA (although several analysis sessions were conducted with a group of student colleagues and supervisors). This limitation can be seen in a close inspection of the ESHA worksheets, where some of the entries appear to be based on assumptions that a domain expert might challenge. However, this limitation was inherent in the structure of the project. The issue of provision of domain expertise is discussed further in Section 7.1.

7 Discussion

In this section we discuss the themes emerging from all the application studies taken as a complete set, i.e. comments on the effectiveness of the ESHA methodology.

7.1 Findings from the INTRO & BRL Experiments

The tutorial session on hazard analysis, which was held at the 3rd INTRO project Workshop at BRL in 2011, was the first trial of the ESHA method. Details of the results of the tutorial are provided in Sections 6.1 and 6.2. There were two specific comments arising from this first trial of the ESHA method, which will be taken into consideration when refining the methodology in the future:

1. Although the intent of ESHA is that the hazard analysis process should not be biased by the mission specification, in practice it is still necessary to provide some contextual information on what general tasks the autonomous system is expected to be doing, if only to allow the relevant environmental situations to be identified in which non-mission interactions might occur. Therefore, it is still necessary to consider the mission in terms of its generalized scenarios as background information to the analysis.
2. Better guidance is needed on the order in which the tables should be completed. The guidelines were insufficiently clear about the need to ensure that each row of the hazard analysis table is complete before moving on to the next one. As a result, one of the sessions became a little chaotic in the way in which the table was completed, and it was noted that this increased the possibility that parts of the analysis may be overlooked. The comment was raised that the wording of the guidelines should be revised to make the procedure more prescriptive in the way in which the analysis steps were to be followed. This will be considered as the guidelines are revised in the light of further practice and experience.

The Guide Robot and the design study was the second phase of trials of the ESHA method, by which time more experience in applying the methods had been gained. This study showed that the general method appears to be feasible, although the major lesson learned at this stage was that like other more established variants of hazard analysis, ESHA requires a team with good domain knowledge in order to produce an analysis with good confidence that all reasonably foreseeable hazards have been identified. While the analysis of the Guide Robot could proceed because this type of robot is operated in domestic environments, for which most people have good domain experience by default, this issue was a particular problem with some of the work on the USAR Robot problem, where there was difficulty in applying the ESHA method because none of the researchers or supervisors had sufficient experience with search and rescue operations to form a confident opinion about the identification of hazards.

7.2 Improvements to Environmental Survey Hazard Analysis

Given the experience of the trials described in Section 6 and the conclusions presented in Section 7.1, we consider the following improvements of the ESHA to be needed for

- Refinements to the ESHA guidewords, to offer more usable guidance.
- Refinements to the ESHA checklist/procedure, to clarify how the ESHA worksheet tables should be completed and the order in which the work should be done.
- Development of further guidance on the composition of the analysis team and the need for persons with suitable domain knowledge or experience to participate in the process.

8 Conclusions

In this section, we discuss some of the wider issues raised by this research.

8.1 Implications for Industry Safety Standards in the Robotics Sector

Once this work gains maturity and is more widely practised and accepted, it may form a valuable tool complementing the use of robotics industry safety standards. We hope that the general principle can be written into future versions of standards such as ISO 13482 that the preliminary hazard analysis stage of any robot development project should include an environmental assessment intended to identify non-mission interactions.

8.2 Requirements for Online Hazard Analysis in Advanced Robots

Although we believe ESHA to provide a useful basis for preliminary hazard analysis by human designers of robots, there are limits to what can be achieved during the design stage. We believe the method will be able to support the claim that human designers have taken all reasonably foreseeable steps to identify hazards for relatively simple robots, which perform only a few tasks in environments that are predictable in advance of the robot's entry into service (such as the initial generation of robots anticipated in the development of the industry safety standard ISO 13482). However, as the number of required mission tasks and the required number of operating environments grows, the number of potential non-mission interactions will grow rapidly, making the task of identifying all such interactions by hand prohibitively expensive, and for more sophisticated robots designers will not credibly be able to make the above claims.

Although an ESHA-style preliminary hazard analysis will still be a useful tool in specifying safety functions for an initial set of non-mission interactions, a truly dependable robot will need to be capable of identifying new environmental features online and developing the relevant safety functions to maintain safety in the new non-mission interactions. This may well entail the use of adaptive and learning mechanisms configured to the identification of novel environmental features, and for the provision of behavioural capabilities for investigating such features and for assessing the safety of the resultant interactions.

Novelty detection and task acquisition is an ongoing field of research in robotics, for example, [4, 27, 29, 30]. Many such methods may be useable for the purpose of online hazard analysis. It may be useful to provide these mechanisms with information structures (knowledge bases, semantic networks, or similar) that encode the ESHA guidewords classification scheme, to ensure that the robot develops an analysis that is an extension of the initial human analysis done at design time. We aim to investigate this idea in future work.

8.3 Future Work

Future work in this area of research is likely to proceed in the following directions:

- The current experiments and trials have tended to focus on wheeled robots used in urban or domestic environments. We are interested in applying ESHA to different domains and applications of robotics, such as UAVs and AUVs, remote manipulation / tele-robotics in medicine, space and other environments. This will be useful in developing and adapting the guide words for ESHA, which may at the present time contain biases towards the applications we have considered so far.
- To date we have taken a breadth-first approach to our application trials, by studying as many different applications as practicable in the time and opportunities available, but to a relatively shallow (incomplete) extent. We did this to get as early an understanding as possible of the relevance and validity of the proposed ESHA guideword set and classification scheme. In future work, we propose to develop an in-depth, full and complete ESHA on an application; this will evaluate explicitly our claim that the method is comprehensive enough to claim that all reasonably foreseeable hazards can be identified for a given environment.
- Other safety analysis methods may be useful for the analysis of robotic systems. In particular, a relatively new hazard analysis methodology called STAMP [31] shows promise as it may also be usable as an externally focused analysis that may also offer a method of identifying non-mission interactions. We are interested in investigating this method in future case studies.

Acknowledgments This work has been funded by the European Commission FP7 framework. It is part of the INTRO (INTeractive RObotics Research Network) project, in the Marie Curie Initial Training Networks (ITN) framework, grant agreement no.: 238486

Appendix A: Hierarchical Task Analysis

The highest level of abstraction in the functional specification of a system is to model the system as a single element (often called a ‘black box’ specification) and to define its interaction with the environment. Typically, this requires a specification of the tasks to be performed by the system, from the viewpoint of external observers, agents or stakeholders. Many methods exist for specifying the externally-observed functionality of a system, including Use Case Design, User Stories, and Viewpoints-based Requirements Engineering. However, for the BRL Robot Waiter design study, a method called Hierarchical Task Analysis was used.

Hierarchical Task Analysis (HTA) [23] is a system analysis method that has been developed by the Human Factors Analysis community as a method for eliciting the procedures and action sequences by which a system is used by human operators. System and procedural models identified by HTA are then used as the basis for operator error analyses to determine whether the system functional or user interface design has an increased potential for of hazards due to human error.

In addition to its use as a methodology for Human Factors analysis, HTA may also be useful as a design technique for mobile robots and other intelligent autonomous systems. The tasks identified within HTA are descriptions of the externally-viewed behaviour required of a robot, which strongly resemble the task modules or behaviour modules developed in many system architectures used widely within the mobile robotics domain (behaviour based architectures). Furthermore, the hierarchical organisation of tasks produced by HTA also resembles the layered hierarchies of tasks that typical of many behaviour-based architectural schemes, such as Subsumption Architecture [5].

Therefore, it is hypothesized that HTA might be a useful candidate for a high level system requirements elicitation technique, generating behavioural (task-based) models of the functionality required of an autonomous robot and identifying their relative hierarchical ordering, without making assumptions about the manner of their implementation. This enhances the utility of HTA as a requirements technique, as it provides maximum freedom of choice to designers in the selection of implementation schemes.

HTA proceeds by the identification of the tasks required of the system, and identification of plans, which describe the order in which tasks are to be performed. Tasks are described by the general activity to

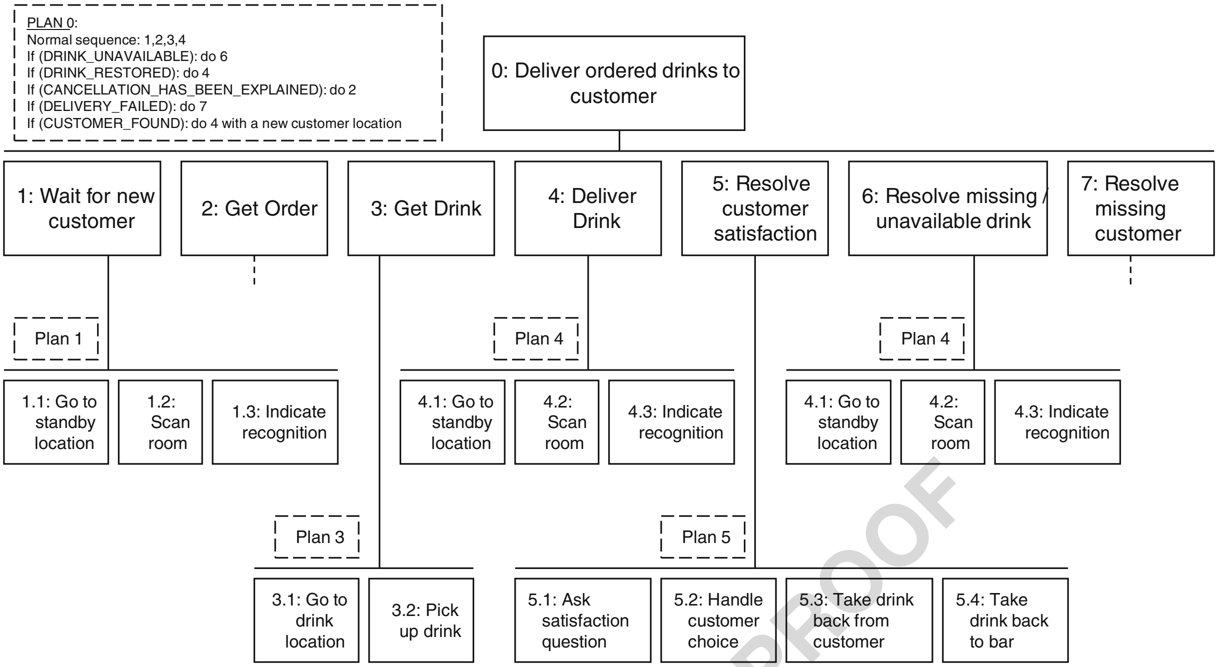


Fig. 7 Partial hierarchical task diagram example for BRL robot waiter design study

be performed and/or the desired end state of the system and its environment at the end of the activity. Each task is then successively decomposed into sub-tasks by the same procedure, as far as is reasonable for the purpose of the analysis. Each task is accompanied by its own plan specifying the ordering of the sub-tasks. The results can also be used in the construction of a hierarchical task diagram that presents the organisational structure of the tasks in a graphical format. An example HTA task diagram is shown in Fig. 7.

The tasks are numbered hierarchically (1, 2.1, 3.2.1, etc.) according to its layer of decomposition, and their associated task plans take the same number.

Each task plan is described in a standard format:

- The *normal sequence*, which describes the intended sequence of execution of the principal sub-tasks necessary to achieve the objective of the task under nominal environmental circumstances.
- *Alternate sequences* may be defined for the sub-tasks, which cater for specific circumstances which may occur but are not considered to be handled by the normal sequence. Typically alternate sequences will be triggered by changes in the environmental conditions that initiated the normal sequence, which obviate that sequence and require further activity to restore the robot and its environment to a nominal state. To take an example from the BRL Robot Waiter study, if a customer leaves the café while the robot is fetching the drink they ordered, then the robot must return the ordered drink to the bar before returning to its waiting location. The sequence “return drink” and “return to waiting location” form an alternate sequence to the normal sequence for delivering the ordered drink. Other candidate alternate sequences might include emergency actions, fail-safe actions, or user-choice actions.

In addition to hierarchical task diagrams, an alternative tabular format for presenting the task structure is shown in Table 14. This table shows an extension to Q10 the tabular format that was added in the BRL Robot Waiter design study, where for each task the behaviour type was identified as defined in the NASA Goddard Agent reference model. This was done to facilitate the development of a functional architecture model on top of the basic task specification. This is described in Appendix B.

Table 14 BRL robot waiter hierarchical task analysis results

Task name	Task description	Behaviour type	Task plan(S)	
0 Deliver Ordered Drink to Customer		[mixed]	*PLAN 0:* • Normal sequence: 1,2,3,4,5 • If (DRINK_UNAVAILABLE): do 6 • If (CANCELLATION_HAS_BEEN_EXPLAINED): do 1 • If (DELIVERY_FAILED): do 7 • If (CUSTOMER_FOUND): do 4 with new customer location	
⌞ 1 Wait for new customer	Remain stationary and look for a new customer	[mixed]	PLAN 1:	
⌞ 1.1 Go to standby location	⌞ Go to standby location and wait there	[reactive]	Normal sequence: 1.1, 1.2, 1.3	
⌞ 1.2 Scan room	⌞ Scan room to look for a customer attentional gesture	[reactive][a]		
⌞ 1.3 Indicate recognition	⌞ Indicate recognition of the attentional gesture to customer	[social]		
⌞ 2 Get Order	Obtain an order for a drink	[mixed]	PLAN 2:	
⌞ 2.2 Attend Customer	⌞ Approach customer close enough to allow use of user interface	[reactive]	Normal sequence: 2.1, 2.2, 2.3	
⌞ 2.3 Take Order	⌞ Interact with customer to obtain the drink order	[social]		PLAN 2.3:
⌞ 2.3.1 Receive Order	⌞ Receive order via user interface	[social]		Normal sequence: 2.3.1, 2.3.2
⌞ 2.3.2 Confirm Order	⌞ Ask customer to confirm that the order is correct	[social]		
⌞ 3 Get Drink	Go to the bar area and obtain the drink	[mixed]	PLAN 3:	
⌞ 3.1 Go to Drink Location	⌞ Move to the bar location where the requested type of drink is supplied	[reactive]	• Normal sequence: 3.1, 3.2	
⌞ 3.2 Pick Up Drink	⌞ Pick up one example of the requested type of drink	[reactive]	• If no drink at location: (DRINK_UNAVAILABLE)	

Table 14 (continued)

Task name	Task description	Behaviour type	Task plan(S)	
∟ 4 Deliver Drink	Deliver drink to customer	[mixed]	PLAN 4:	
∟ 4.1 Approach Customer	∟ Carry drink to customer location	[reactive]	Normal sequence: 4.1, 4.2, 4.3	
∟ 4.2 Engage Customer	∟ Interact with customer to obtain permission to serve drink and mode of service	[mixed]	• If no customer at original location: (DELIVERY_FAILED)	
∟ 4.2.1 Get customer attention	∟ Attract customer attention with a sign	[social]		PLAN 4.2:
∟ 4.2.2 Detect customer recognition	∟ Scan customer for sign of recognition	[social]		Normal sequence: 4.2.1, 4.3.2; 4.2.3
∟ 4.2.3 Request mode of service	∟ Ask customer for service mode (on table or hand-to-hand)	[social]		
∟ 4.3 Serve Drink to Customer	∟ Serve drink to customer by requested mode	[reactive]		
∟ 5 Resolve Customer Satisfaction	Ask customer if order is satisfactory and resolve any complaints	[mixed]	PLAN 5:	
∟ 5.1 Ask satisfaction question	∟ Ask customer for Yes/No answer on their satisfaction	[social]	Normal sequence: 5.1, 5.2, 5.3, 5.4	
∟ 5.2 Handle customer choice	∟ Offer customer choice of action	[social]	• If customer requests replacement drink do 3	
∟ 5.3 Take drink back from customer	∟ Pick up drink from table or from customer's hand	[reactive]	• If customer requests new drinks order do 2	
∟ 5.4 Take drink back to bar	∟ Return unwanted drink to bar (to returns area)	[reactive]		
∟ 6 Resolve Missing/Unavailable Drink	Find out why drink is unavailable and report back to customer	[mixed]	PLAN 6:	
∟ 6.1 Notify bartender	∟ Notify bartender that there is no drink	[social]	• Normal sequence: 6.1, then CHOICE:	
∟ 6.2 Wait for new drink	∟ Wait fixed time for a new drink to be supplied	[reactive]	☐ If drink is delayed then do 6.2 then do 2;	
∟ 6.3 Return to customer	∟ Return to customer, explain reason, and take new order if requested	[mixed]	☐ If no drinks left then do 6.3	PLAN 6.3:
∟ 6.3.1 Return to customer location	∟ Go back to original location of customer	[reactive]		Normal sequence: 6.3.1, 6.3.2 At end of 6.3.2: (CANCELLATION_HAS_BEEN_EXPLAINED) If no customer at end of 6.3.1 then do 1
∟ 6.3.2 Explain reason	∟ Explain reason for unavailable drink	[social]		

Table 14 (continued)

Task name	Task description	Behaviour type	Task plan(S)
⌞ 7 Resolve Missing Customer	Search for missing customer and/ or take undelivered drink back to bar	[mixed]	PLAN 7:
⌞ 7.1 Do local search	⌞ Search for customer within table area for fixed time period	[reactive]	• Normal sequence: 7.1, 7.2 then do 1 • If customer is recognised during 7.1 time period: (CUSTOMER_FOUND)
⌞ 7.2 Take drink back to bar	⌞ Return unwanted drink to bar (to returns area) [identical to 5.4]	[reactive]	

[a] This task could be considered proactive, in that the robot could be considered to proactively scan the environment for new customers

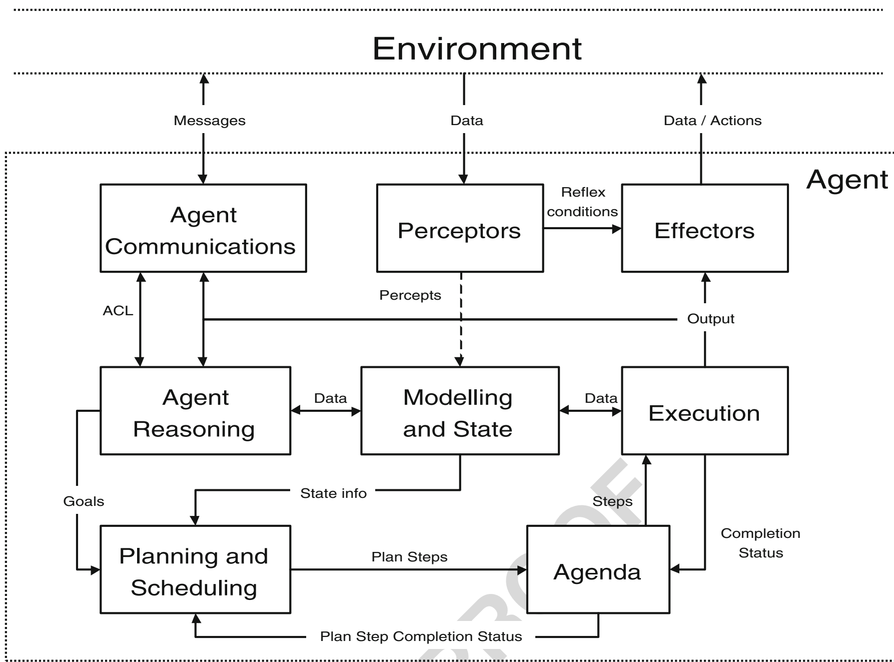
Appendix B: Use of the NASA Goddard Reference Architecture as a System Model

In the BRL Robot Waiter experiment, we decided to use the NASA Goddard Agent Architecture [33] as a reference model for the robot functional architecture design. This model identifies the general nature of the cognitive processing required in order to perform behavioural tasks of a given type. The components of the architecture model are shown in Fig. 8.

The architecture model identifies a number of cognitive processes that must be present within an autonomous agent if it is to perform various different types of task:

- *Perceptors* observe the environment and provide signals or indications (percepts) that reflect the state or condition of the environment. Perceptors may be more than just a sensor; they may include some level of signal processing in order to provide a particular item of information to the other cognitive processes of the agent. Perceptors also provide more primitive signals to the effectors, for the purposes of performing reflexive behaviour patterns (see later).
- *Effectors* are the actuators, motors, muscles, or other transducers that act physically upon the environment. Effectors may either perform physical activity, or they may provide other forms of emission of information, materiel or energy into the environment.
- The *Agent Communications* process performs explicit message-based communications directed specifically to other agents. This is the primary cognitive process associated with social behaviour patterns, which involve dialogue rather than just physical actions.
- The *Execution* process is responsible for deciding upon the specific actions to be taken in order to achieve the steps of a given plan (provided by other processes). It can be thought of as the lowest level of action planning within the agent. Actions are specified based on the action plan and the state of the world as supplied by the Agenda and the Modelling & State processes.
- The *Modelling and State* process provides the storage of all data, information or knowledge required by the agent, typically in the form of world models or knowledge bases. In general it is a passive component, merely providing a storage

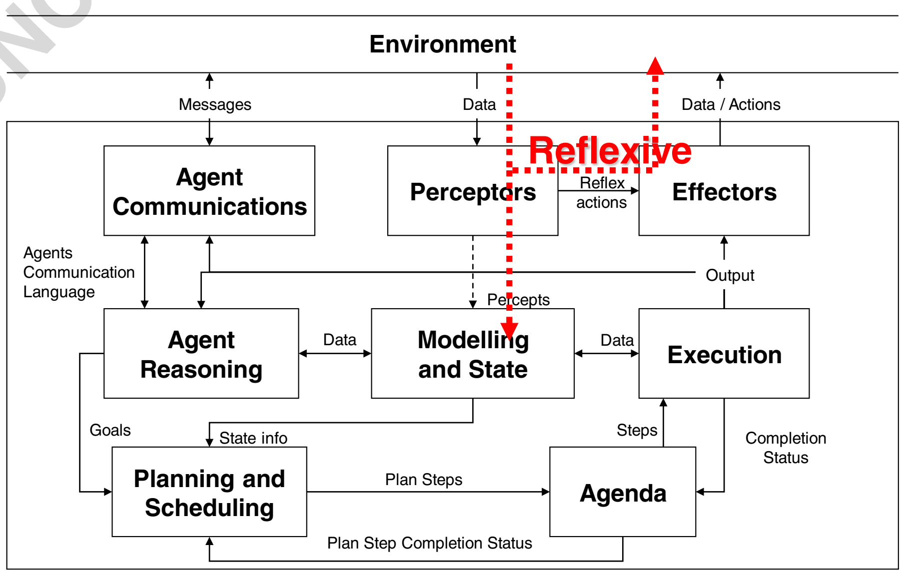
Fig. 8 NASA Goddard agent architecture reference model



and retrieval service to other processes. However, occasionally it may be the source of internally triggered or motivated behaviour patterns, if any specific data/information patterns occur within the world model.

- The *Agent Reasoning* process is the source of all logical inference and reasoning within the agent. It encodes the primary goals of the agent, and invokes the necessary deliberative, social or reflexive behaviours needed to achieve them. This process is the principal source of internally motivated (proactive) behaviour, although other processes may also do so (as above).
- The *Planning and Scheduling* process is responsible for the generation and monitoring of action plans that achieve the goals generated by

Fig. 9 Reflexive behaviour



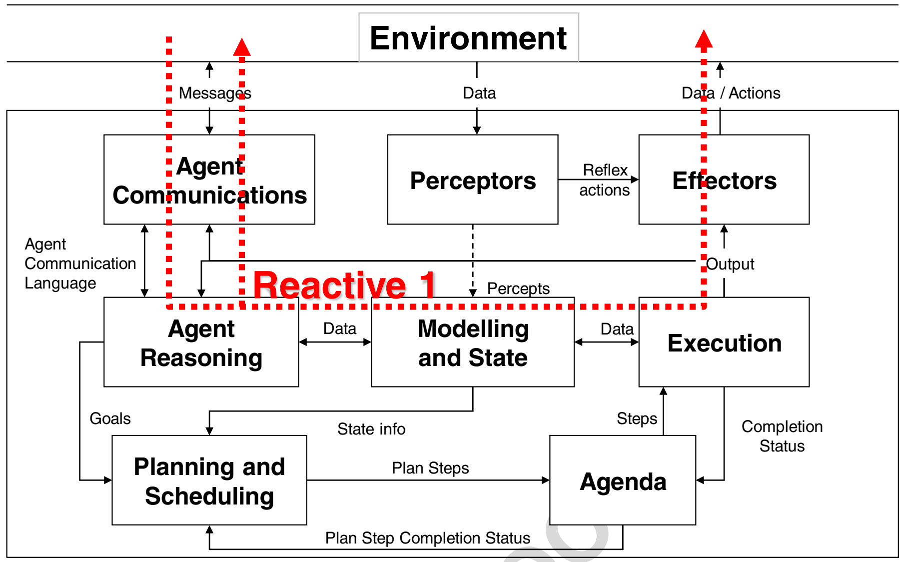


Fig. 10 Reactive 1 behaviour

the Agent Reasoning process. This process is intended to perform only a high level planning process (management or supervisory), selecting from a range of more specific plans, monitoring their completion, and reacting to failures with the selection of new plans.

- The *Agenda* process is responsible for the lower level of planning, identifying the action steps required to achieve the high level plans supplied by the Planning & Scheduling Process. It passes the individual action steps to the Execution process, monitors their successful completion, and then advises the Planning & Scheduling process as to whether a given plan has been performed successfully (or otherwise).

The processes shown in Fig. 8 define the internal cognitive mechanisms required of an agent. The Goddard Agent Architecture Model also identifies a number of different types of behaviour pattern that an agent may exhibit:

- Reactive: reasoned action initiated by events in the environment

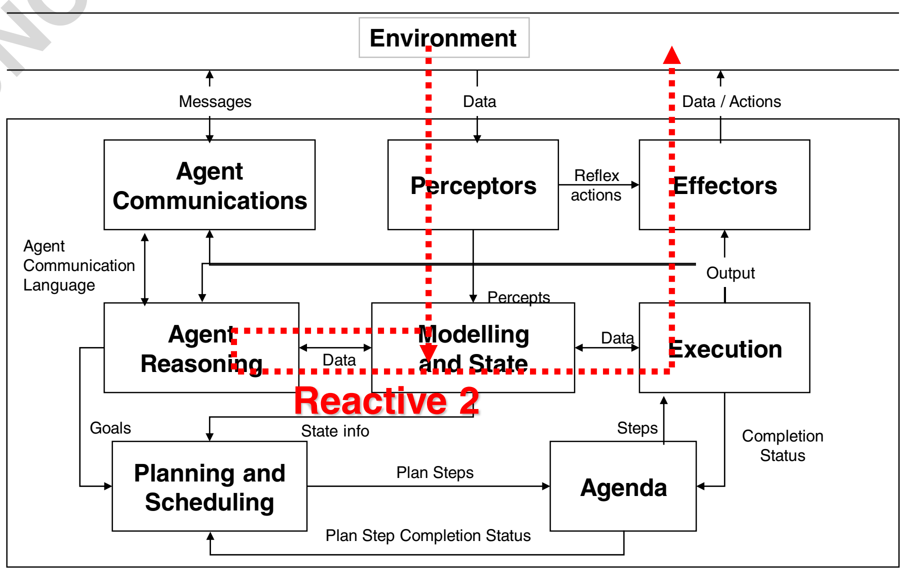


Fig. 11 Reactive 2 behaviour

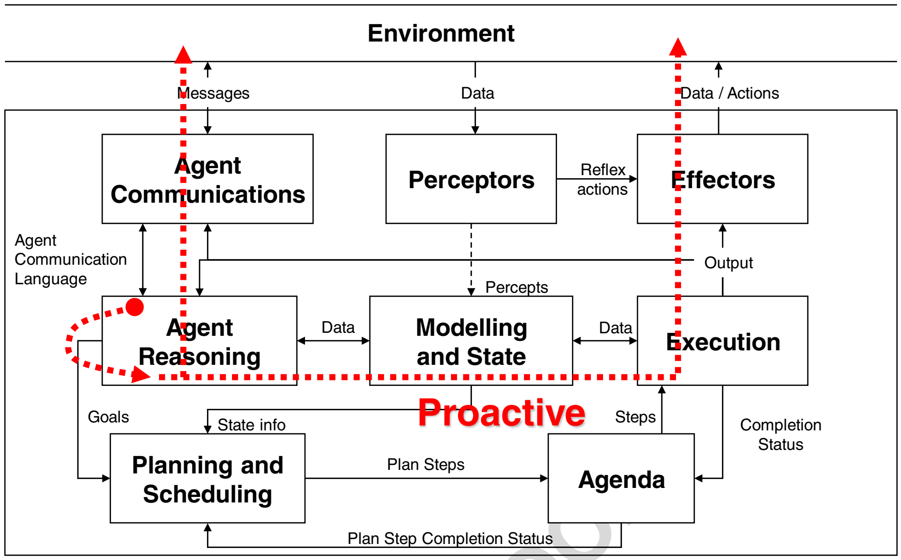


Fig. 12 Proactive behaviour

- Reflexive: fixed/stereotyped action pattern initiated directly by percepts
- Deliberative: reasoned and planned action initiated by external events
- Proactive: action initiated by the agent itself due to internal motivations
- Social: dialogue with other agent(s) which may also trigger action

These basic behaviour types are then extended by consideration of how the behaviour may be triggered or initiated, thereby producing a list of eight specific *behaviour modes*:

1. Reactive 1: triggered by another agent
2. Reactive 2: triggered by a percept
3. Reflexive
4. Deliberative 1: triggered by another agent
5. Deliberative 2: triggered by a percept
6. Proactive
7. Social 1: triggered by another agent
8. Social 2: triggered by the agent itself

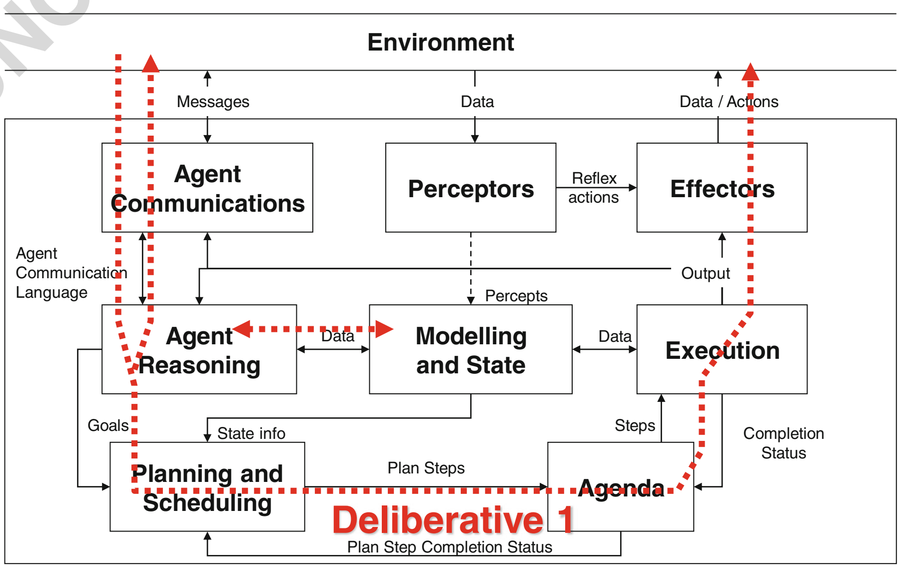


Fig. 13 Deliberative 1 behaviour

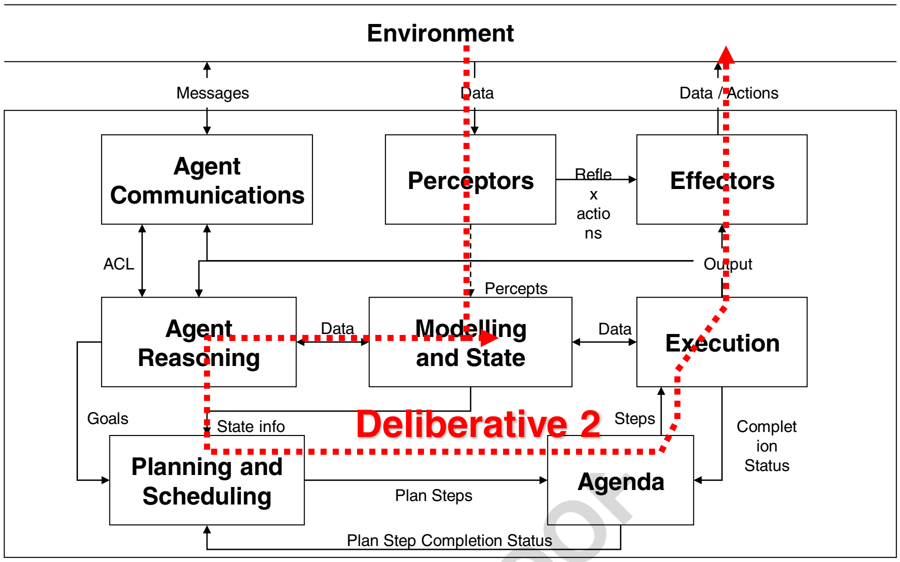


Fig. 14 Deliberative 2 behaviour

The Goddard Agent Architecture Model identifies how the cognitive processes combine to perform each behaviour mode by modelling the information flow through the process model. The various different information flow archetypes are presented in Figs. 9–16.

Although the Goddard Agent Architecture reference model is presented as a block diagram suggesting that the constituent processes must be thought of as an implementation, it need not be interpreted in this way. The model is intended to define the *cognitive* processes of an agent, not necessarily the *software* processes. There does necessarily need to be a one-to-one correspondence between the cognitive processes required of an agent and the software algorithms that are programmed into its computational equipment. Instead, the model may be interpreted as a statement of the functional requirements for performing behaviours of a given type, which could be implemented by other architectures as appropriate, as long as the cognitive processes necessary are allocated to the elements of the implementation architecture.

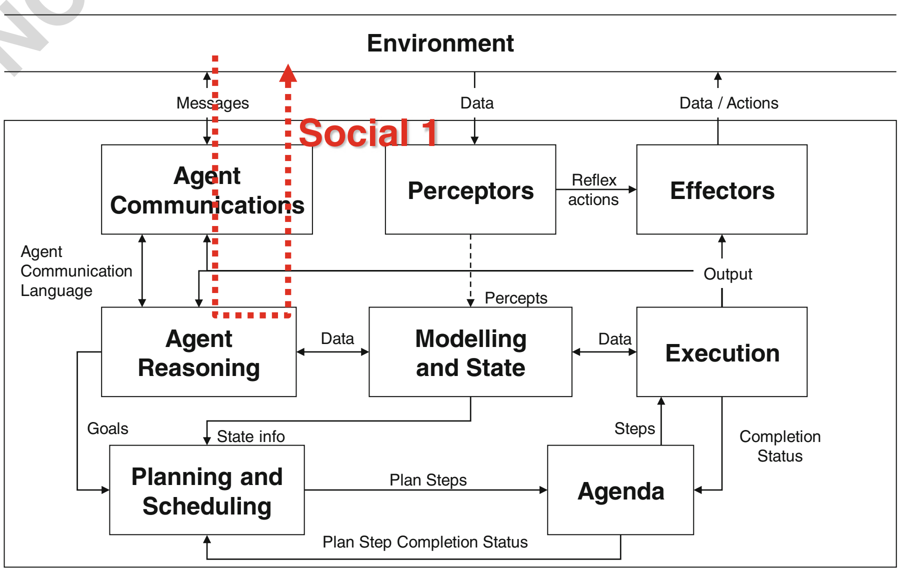


Fig. 15 Social 1 behaviour

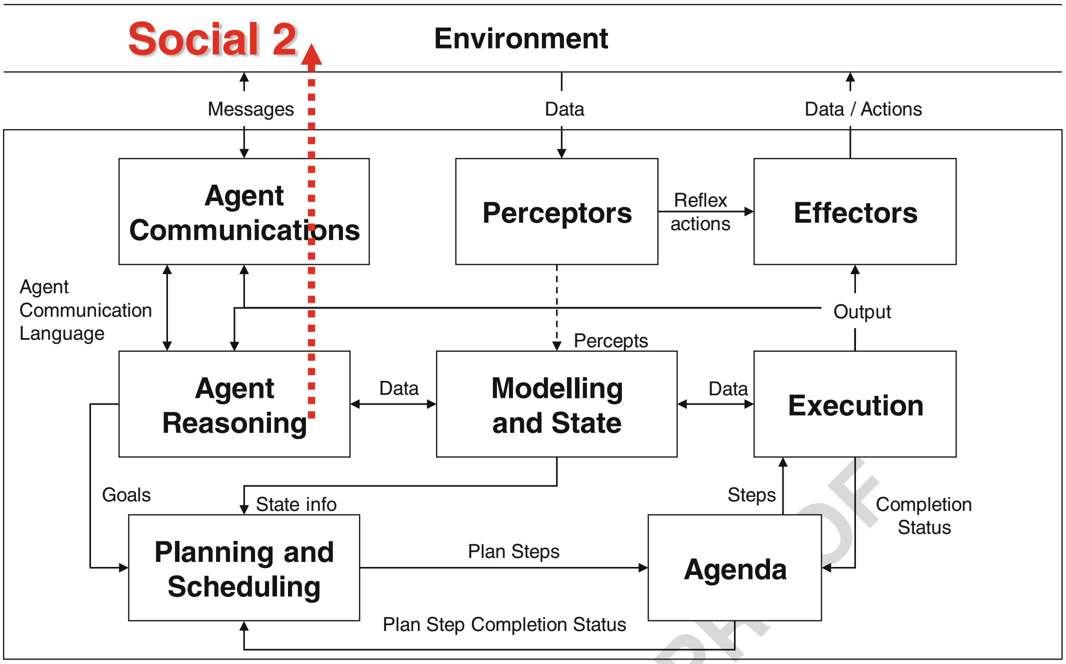


Fig. 16 Social 2 behaviour

Thus, it is possible to use the Goddard Agent Architecture Model as a reference model for functional requirements for the primitive processes of the task model, to identify the internal functionality they require. This can then be used in further design studies such as functional hazard/failure analysis, by providing some information about the internal functional processes of the system, but still retaining considerable freedom about how the design may be implemented.

References

1. Alami, R., Albu-Schaeffer, A., Bicchi, A., Bischoff, R., Chatila, R., De Luca, A., De Santis, A., Giralt, G., Guiochet, J., Hirzinger, G., Ingrand, F., Lippiello, V., Mattone, R., Powell, D., Sen, S., Siciliano, B., Tonietti, G., Villani, L.: Safe and dependable physical human-robot interaction in anthropic domains: State of the art and challenges. Proc. IROS'06 Workshop on pHRI - Physical Human-Robot Interaction in Anthropic Domains (2006)
2. Alexander, R., Herbert, N., Kelly, T.: The role of the human in an autonomous system. Proceedings of the 4th IET System Safety Conference (2009)
3. ARP 4761: Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment. Society of Automotive Engineers (1996)
4. Bonasso, P., Kortenkamp, D.: Using a layered control architecture to alleviate planning with incomplete information. Proceedings of the AAA Spring Symposium on Planning with Incomplete Information for Robot Problems, pp. 1–4 (1996)
5. Brooks, R.: Cambrian Intelligence: The Early History of the New AI. MIT Press, Cambridge (1999)
6. Böhm, P., Gruber, T.: A novel hazop study approach in the rams analysis of a therapeutic robot for disabled children. Proceedings of the 29th International Conference on Computer Safety, Reliability, and Security, vol. 6351, pp. 15–27 (2010)
7. Choung, J.: Safety analysis & simulation of a guide robot for the elderly in care home, MSc Dissertation, University of Bristol (2012)
8. Eliot, C.E.: What is a reasonable argument in law? Proc. 8th GSN User Club Meeting, York UK, 2007 December (2007) Q11
9. Fuller, C., Vassie, L.: Health and Safety Management: Principles and Best Practice. Pearson Education, Essex (2004)
10. Giannaccini, M.E., Sobhani, M., Dogramadzi, S., Harper, C.: Investigating real world issues in Human Robot Interaction: Physical and Cognitive solutions for a safe robotic system. Proc. ICRA 2013, IEEE (2013)
11. Giuliani, M., Lenz, C., Mller, T., Rickert, M., Knoll, A.: Design principles for safety in human-robot interaction. Int. J. Social Robot. **2**(3), 253–274 (2010)
12. Goodrich, M., Schultz, A.: Human-robot interaction: a survey. Found. Trends Hum. Comput. Interact. **1**(3), 203–275 (2007)
13. Grigore, E.C., Eder, K., Pipe, A.G., Melhuish, C., Leonards, U.: Joint Action Understanding improves Q12

Robot-to-Human Object Handover, to appear in Proc IROS 2013 (2013)

14. Guiochet, J., Baron, C.: UML based risk analysis - Application to a medical robot. Proc. of the Quality Reliability and Maintenance 5th International Conference, Oxford, UK, pp. 213–216, Professional Engineering Publishing, I Mech E. April, 2004 (2004)
15. Guiochet, J., Martin-Guillerez, D., Powell, D.: Experience with model-based user-centered risk assessment for service robots. Proceedings of the 2010 IEEE 12th International Symposium on High-Assurance Systems Engineering, pp. 104–113 (2010)
16. Haddadin, S., Albu-Schäffer, A., Hirzinger, G.: Requirements for safe robots: measurements, analysis and new insights. Int. J. Robotics Res. **28**(11–12), 1507–1527 (2009)
17. Haddadin, S., Albu-Schaffer, A., Hirzinger, G.: Soft-tissue injury in robotics. In: Robotics and Automation (ICRA), IEEE International Conference on 2010, pp. 3426–3433. IEEE (2010)
18. Harper, C., Giannaccini, M.E., Woodman, R., Dogramadzi, S., Pipe, T., Winfield, A.: Challenges for the hazard identification process of autonomous mobile robots. 4th Workshop on Human-Friendly Robotics Enschede, Netherlands (2011)
19. Heinzmann, J., Zelinsky, A.: Quantitative safety guarantees for physical human-robot interaction. Int. J. Robot. Res. **22**(7), 479–504 (2003)
20. IEC 61882: Hazard and operability studies (HAZOP studies)-Application Guide, IEC (2001)
21. Ikuta, K., Ishii, H., Makoto, N.: Safety evaluation method of design and control for human-care robots. Int. J. Robot. Res. **22**(5), 281–298 (2003)
22. ISO/FDIS 13482: Robots and robotic devices - Safety requirements - Non-medical personal care robot. International Organization for Standardization (2013)
23. Kirwan, B., Ainsworth, L.K.: A Guide to Task Analysis: The Task Analysis Working Group. Taylor & Francis, London (1992)
24. Kulic, D., Croft, E.: Strategies for safety in human robot interaction. Proceedings of IEEE International Conference on Advanced Robotics, pp. 644–649 (2003)
25. Kulic, D., Croft, E.: Pre-collision safety strategies for human-robot interaction. Auton. Robot. **22**(2), 149–164 (2007)
26. Lankenau, A., Meyer, O.: Formal methods in robotics: Fault tree based verification. Proceedings of Quality Week (1999)
27. Larsen, T., Hansen, S.: Evolving composite robot behaviour – a modular architecture. Proceedings of RoMoCo'05, pp. 271–276 (2005)
28. Martin-Guillerez, D., Guiochet, J., Powell, D., Zanon, C.: A UML-based method for risk analysis of human-robot interactions. 2nd International Workshop on Software Engineering for Resilient Systems, pp. 32–41 (2010)
29. Nehmzow, U.: Flexible control of mobile robots through autonomous competence acquisition. Meas. Control **28**, 48–54 (1995)
30. Nehmzow, U., Kyriacou, T., Iglesias, R., Billings, S.: Robotmodic: modelling, identification and characterisation of mobile robots. Proc. TAROS 2004 (2004)
31. Owens, B.D., Stringfellow Herring, M., Dulac, N., Leveson, N.G.: Application of a Safety-Driven Design Methodology to an Outer Planet Exploration Mission, IEEEAC paper #1279, Version 8, Updated December 14 (2007)
32. Pumfrey, D.: The principled design of computer system safety analyses. PhD Thesis, University of York (1999)
33. Rouff, C.A., Hinchey, M., Rash, J., Truszkowski, W., Gordon-Spears, D. (eds.): Agent Technology from a Formal Perspective. Springer (2006)
34. Sobhani, M.M.: Fault Detection ad Recovery in HRI in Rescue Robotics. MSc Dissertation, Bristol Robotics Laboratory (2012)
35. UK MoD: HAZOP Studies on Systems Containing Programmable Electronics. Defence Standard 00-58 Issue 2, UK Ministry of Defence (2000)
36. UK National Archives 1974, UK Health and Safety at Work Act 1974, available freely over the internet at http://www.legislation.gov.uk/. Accessed 30 Sept 2013 (1974)
37. UK National Archives 1987, UK Consumer Protection Act 1987, available freely over the internet at http://www.legislation.gov.uk/. Accessed 30 Sept 2013 (1987)
38. Woodman, R., Winfield, A.F.T., Harper, C., Fraser, M.: Building safer robots: Safety driven control. Int. J. Robot. Res. **31**(13), 1603–1626 (2012)
39. Wozniak, J., Baggiolini, V., Garcia, D.Q., Wenninger, J.: Software interlocks system. Proceedings of ICALEPCS07, pp. 403–405 (2007) Q9

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES:

Q1. Please check authors' names (conflict with manuscript draft) and contacts if correct.
Q2. Please check captured corresponding author if correct.
Q3. Petterson 2005 and Lussier et al. 2004, was cited in text, but not found in reference list, Please provide. Otherwise, delete the citation from the text.
Q4. Dummy citation for Table 6 was inserted in the 2nd paragraph of section "Robot Waiter Task Specification", please provide a sequenced table citations.
Q5. Please check presentation of table 8 if correct.
Q6. Please check table 8 missing text in body cells.
Q7. Tables were renumber (Table 8 was given twice), please confirm if correct.
Q8. Please check table 10 footnote if captured correctly.
Q9. Dummy citation for Table 12 was inserted here. Please check if appropriate.
Q10. Please check presentation of table 14, if appropriate.
Q11. References 9 and 39 were not cited anywhere in the text. Please provide. Otherwise, delete the citation from the list.
Q12. Please provide updated details for reference 13.