AN APPROACH TO SAFETY FOR A ROBOTIC EXCAVATOR

Conrad J. Pace [1] and Derek W. Seward [2]

[1] Department of Manufacturing Engineering, University of Malta, Malta
[2] Department of Engineering, Lancaster University, U.K.

Abstract: Safety is a critical control issue for any system interacting with its environment. This paper presents an overview of such safety concerns for a robotic excavator, and how such considerations may be integrated within the system's control framework. The provision of this integration allows the handling of safety issues at different control levels and provides an essential starting point for system safety validation and verification.

Keywords: Safety, Hazard and Risk Assessment, Hybrid Control Architectures.

1 INTRODUCTION

It quickly becomes apparent that any industrially feasible autonomous robotic system needs to comply with various safety standards. Such standards do not only consider the final product safety but also take into account a safety development procedure. Indeed, one of the principle standards for computer based systems, the IEC 61508 [1], clearly identifies a safety life cycle development process, in order to provide safety compliance of the system.

In autonomous mobile robotics, such as a robotic excavator, a major safety issue is the system complexity level which gives rise to substantial difficulties in assessing system safety, and in complying with safety life cycle requirements [2]. Worse still, situations may arise where it is practically impossible to carry out the required safety analysis retrospectively [3] when minimal safety consideration has been taken during development.

In addition, when it comes to mobile construction robotics, a key safety problem is the need to interact with an unbounded and unstructured environment. Such systems need to perceive their environment, which, at best, is only partially known, and act in a manner to safely modify their environment according to the specified goal. This imposes a major safety concern which is generally not considered for industrial robotic systems, where environments are well structured and fully known and perceptual requirements, if at all necessary, are very limited.

A specific safety approach is therefore required, which takes into account the above concerns. This paper describes work that has been carried out to develop a framework for an autonomous excavator safety management and control. This framework promotes safety integration without hindrance to task achievement. Here, safety integration is not only concerned with ensuring that the 'internal system' is fully functional. More importantly, it has to deal with the requirement to handle the system's limitation in observing its environment and in assessing the system's ability to interact with its environment in a safe manner. In addition, such a framework may form a basis for validating the autonomous excavator's safety, through a safety reasoning approach within the architectural framework.

The paper is divided as follows; Section 2 will present a review of the safety concerns for a robotic excavator. Section 3 will then present the basis for an architectural framework specifically designed to incorporate the identified safety concerns. Sections 4 and 5 will finally outline the main issues regarding the handling of safety within the control architecture.

2. ANALYSING SAFETY IN AN AUTONOMOUS EXCAVATOR

Dhillon and Fashandi [4] have presented a valid foundation on which to base an analysis of the safety and risk implications within robotic systems. They present the following five issues which have to be taken into account when analysing safety;

- The necessity to Consider Safety *(Why?)*
- The Sources of Hazards *(What?)*
- Entities at risk and Responsibility for Managing Safety *(Who?)*
- Consideration of Safety Aspects during Implementation *(When?)*
- Location of Safety Considerations *(Where?)*

These issues have been considered in the context of an autonomous excavator, providing the basis for a study on the safety requirements for such a system. The following sections outline such considerations.

2.1 The Problem of Safety for an Autonomous Eexcavator

Safety has been a major concern in industrial robotics for a number of years, and research in the field of robotic safety has been quite substantial [5]. To augment this concern, various standards have been published on robot safety [6][7][8], which outline installation, operational and maintenance aspects, together with issues on reliability and safety testing. The focus of these standards, though, has always been on industrial robotic manipulators, with little or no concern to mobile robotics, let alone systems such as an autonomous excavator.

Such standards are very limited in their scope when applied to robotic excavators. One main reason for this limitation is the issue of mobility. Whereas in industrial robotic systems the environment is largely bounded and structured, in free ranging mobile robots such as autonomous excavators, this is hardly the case. This difference has implications on the robot's interaction with the environment and the ensuing analysis of such an interaction.

Typically, for an industrial robot, a safety analysis for determining potential hazards becomes a relatively deterministic exercise, the outcome being the elimination or containment of the identified hazards, giving rise to a substantial risk reduction before the system is put into operation. A safety analysis on the autonomous excavator's operational characteristics, on the other hand, becomes a vastly more complex exercise, due to the countless modes of excavator – environment interaction. Furthermore, hazard elimination may not be practical, as the system's designer will have no control on the environmental characteristics in which the excavator may operate. Hazard containment, which will be dependent completely on the mode of interaction between excavator and environment, thus becomes the principle mode of managing risk and avoiding accidents during operation.

2.2 Identification of Hazard Sources and Risks in Autonomous Excavators

Undoubtedly, any safety analysis requires the identification of hazard sources and a quantification of the risks attributable to these hazards. A hazard and risk analysis is the tool for such hazard source identification. Such a hazard analysis has been carried out as a preliminary exercise for a typical autonomous excavator [9] and an immediate outcome of this analysis defines two main groups of hazard sources. These groups can be defined as sources originating internally within the system and sources originating externally from the system.

Internal hazard sources are considered to include failures occurring within the hardware and control software of the system such as motor and sensory failure, etc.. Most standard safety design approaches are aimed at minimising or containing these types of hazards, by promoting the use of various techniques such as fault tree analysis (FTA) and failure mode and effect analysis (FMEA) for identification purposes. Handling of internal failures has also been achieved successfully for robotic systems, through various techniques such as the inclusion of system redundancy and diagnostic routines [10][11].

External hazard sources, though, are of a completely different nature. They generally do not arise due to some system fault, either hardware or software, but rather, as the outcome of the nature in which the autonomous excavator interacts with its environment. The operation of the excavator within a highly complex environment may give rise to chains of events, which themselves may be the sources of hazards. These may include events leading to collisions with various elements within the environment, excavator toppling due to both the excavator's operation and the type of terrain on which the excavator is working, and even hazards resulting from the modifications imparted to the environment by the excavator itself.

Independently from the type and form of hazard, external hazards may be considered to originate due to two main system deficiencies:

1. The excavator's inability to perceive a chain of events that give rise to a hazardous situation. This inability is mainly attributable to limitations in the system's perceptual abilities, which is mostly due to constraints on the sensory system and the related processing of the sensory data.
2. The inability to react to a perceived hazard. Even if a sequence of events may be perceived as a potentially hazardous situation, weaknesses in the decision making process, may still give rise to accidents, due to the absence of a correct reaction to the perceived hazard.

A major contributing factor to both perceptual and action deficiencies is considered to be the real-time constraint in interacting with a dynamic environment.

2.3 Safety and Risk Management

In industrial robotic systems, the ensuing risk management exercise following the risk analysis can be fully implemented before the system is put into operation. In such cases, it is possible to identify and manage the potential hazards which the system may encounter, during system development. However, in systems such as a robotic excavator, risk assessment and management cannot be thoroughly done before the system is operational. Once more, this is a result of the nature in which the excavator interacts with its operational environment. Risk management becomes even more complex when considering the perceptual and action limitations as described earlier in section 2.2. In such a situation it is not possible to limit environmental features to what can be perceived and acted upon by the excavator. Rather an integral part of the risk management exercise will rely on the ability of the system to identify its limitations, and then to achieve its task within these limitations. This signifies that most risk assessment and management decision making will have to be carried out in real time, and furthermore, by the excavator itself, rather than by the system designer.

Safety management thus becomes an issue of the excavator being able to correctly assess its perceptual and reaction capabilities and limitations in the context of the task it is carrying out and the environment in which it is operating. This shift of responsibility in the determination of what may be considered as a safe or unsafe operational state, requires the autonomous excavator to be able to ‘reason’ on the consequences of its actions in a safety context. Avoiding the inclusion of the ability to ‘reason’ on the system’s safety and to act accordingly will only result in a system which is either unsafe or so constrained by the safety requirements, that its task achieving abilities will be drastically hindered.

2.4 Development Process Aspects for Safety – When and Where

The nature of the hazards and the requirement to shift safety decision making from the designer to the autonomous system undoubtedly requires the consideration of safety during all design stages. Consideration right from the start of designing an autonomous excavator should ensure that safety is integrated within both the system’s hardware and control and decision making software. Furthermore, an early consideration of safety provides a basis on which sensory and actuation requirements may be outlined, on both safety and control requirements. Yet, as stated earlier, safety concerns must not stop with the choice of hardware and control software, but more importantly, safety management requirements must be defined and integrated within the control architecture. It is only through this integration that an embedded safety system will be capable of assessing its limitations and accordingly, to correctly interpret its perceived environment, and hence, the resulting state of its operation.

3. A FRAMEWORK FOR HANDLING SAFETY REQUIREMENTS

3.1 Satisfying Safety Requirements

The approach for dealing with the hazard sources and safety management requirements has to be based on the management of perception and action and the handling of the related limitations. Such an approach results in the following generic safety requirements for the control architecture:

- To perceive and act in real time to satisfy both the task at hand and the safety constraints.
- To reason on whether the perceptual information is correct and detailed enough to perform the required task safely, both regarding the immediate and long-term effects.
- To reason on whether the system reaction time is expected to be fast enough according to the perceived changes and rate of change of characteristics of the environment, and the system can perform the required actions to maintain safety in both the short and long term.
- To obtain real-time feedback on the system’s capabilities to assess perception and reaction abilities and to modify such assessments according to the feedback.
- Following the above requirements, to be able to modify the task at hand to suit such limitations and maintain safety.
- The ability for the system to override goal achieving actions by safety actions where necessary to maintain safety as a priority over task achievement

These requirements may be summarised as the ability to reason on the safe interaction between the excavator and the environment within the context of the task at hand, on different time and spatial scales.

To satisfy such requirements, a safety framework, which encompasses both real-time reaction and long-term consequence analysis abilities, is needed. Such a framework may be based on the mode in which control architectures have been developed to handle real-time requirements and long term plans. For this purpose, hybrid architectures which include low level reactive and high level abstract ‘reasoning’ modules, have been proposed and successfully implemented [12][13]. James Albus [14] also proposes a mode in which such reasoning requirements may be handled. In his architecture, control is divided into layers where each layer handles perception and action on different time and spatial frames and handling different levels of detail. Furthermore, each layer should contribute in the perceptual and action abilities of the layers above and below it.

This same hybrid approach can be adopted for managing safety, starting with lower reactive type, behavioural layer, to higher abstract reasoning layer which take into consideration longer chains of events and trends within the system’s operation, utilising past experiences and embedded knowledge within their reasoning processes. In addition, the provision of transferring perceptual information from one layer to another contributes to the ability of other layers, particularly higher layers to assess the adequacy of their assessment on perception and action. Such an architecture, based on control layers which handle situations on different temporal and spatial time scales, allows the system to cope with both real time safety situations and long term safety activities.

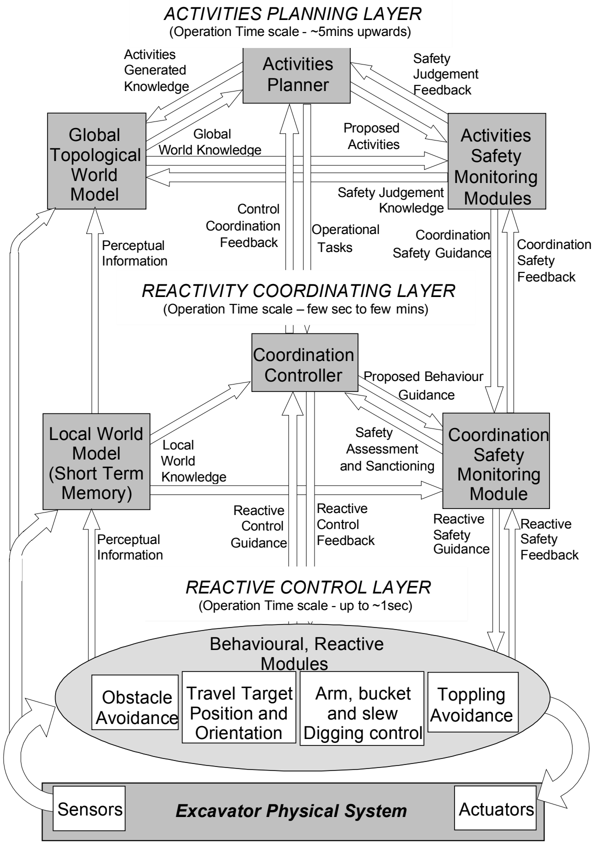
3.2 A Safety 'Conscious' Architecture

The hybrid layered approach has been developed into an architectural framework for managing safety. The framework consists of a three layer architecture;

1. A lower reactive control layer,
2. An intermediate reactivity coordinating layer
3. A top activities planning layer

In such a framework, both safety and control requirements are mapped on the individual layers. Figure 1 depicts these layers together with the communication requirements and layer structure.

In this architecture, the lower reactive control layer operates as a set of behaviours mainly dealing with real-time, mostly reactive, responses of the excavator to sensory data signals. These behaviours are influenced by the upper layers, allowing varying behaviour intensities to be generated depending on the perceptions of the higher levels. No world representation is generated at this level since behaviours are activated directly through the raw sensory data and through the influence of upper layers. At this level safety is managed directly by the behaviours, where the behaviours embed both safety and control features. Indeed, there is no distinction between the safety and control elements of the different behaviours.



ACTIVITIES PLANNING LAYER
(Operation Time scale - ~5mins upwards)
Activities Generated Knowledge
Activities Planner
Safety Judgement Feedback
Global Topological World Model
Global World Knowledge
Proposed Activities
Activities Safety Monitoring Modules
Safety Judgement Knowledge
Control Coordination Feedback
Operational Tasks
Coordination Safety Guidance
Coordination Safety Feedback
Perceptual Information
REACTIVITY COORDINATING LAYER
(Operation Time scale – few sec to few mins)
Coordination Controller
Proposed Behaviour Guidance
Safety Assessment and Sanctioning
Local World Model (Short Term Memory)
Local World Knowledge
Coordination Safety Monitoring Module
Reactive Control Guidance
Reactive Control Feedback
Reactive Safety Guidance
Reactive Safety Feedback
Perceptual Information
REACTIVE CONTROL LAYER
(Operation Time scale - up to ~1sec)
Behavioural, Reactive Modules
Obstacle Avoidance
Travel Target Position and Orientation
Arm, bucket and slew Digging control
Toppling Avoidance
Sensors
Excavator Physical System
Actuators

Figure 1. Layout of an autonomous excavator's control architecture for safety integration

Through the safety integration within behaviours, the lower level, apart from providing a fast reacting control for task achievement, also allows for the avoidance of obstacle and toppling in real-time, by having quick responses on specific sensory signals, typical of behaviour type control architectures [12]. In addition, safety actions taken within this layer provide feedback to the upper layers, indicating the level of real-time intervention for maintaining safety.

The intermediate coordinating layer mainly acts as an interface between the lower reactive and the upper activities planning layers as is typical of most hybrid architectures [12]. The objective at this level is to influence low-level behaviours according to the planned activities originating from the upper layer.

An egocentric world representation of the immediate vicinity of the excavator is developed at this level. The scope of this model is to act as a short-term memory of the excavator's perceived world. Such a metric 'bird's eye view', mapping should be detailed enough so as to allow, together with the outlined plans from the upper layer, the generation of the necessary influential signals to the behaviours at the lower layer. The provision of local world mapping further allows the intermediate layer to assess safety features within the excavator's close vicinity, thus providing for a preliminary hazard aversion approach, before real-time behaviours take over to contain any perceived hazard.

At the intermediate layer a partial distinction can be made between safety and control elements. This allows control-oriented elements within this level to generate behaviour influences, while safety oriented elements provide a censoring action on such influences. Feedback of this layer's safety actions is provided to the upper layer, allowing the upper layer to carry out a more thorough safety assessment.

The upper activities planning layer's main task is to generate operational plans. Here, decision making is based on a topologically constructed global world model [13][15]. The world model, based on a graph model representing locations and paths within the environment, provides a knowledge base which aids in the system's activities planning and safety assessment. The knowledge representation is highly symbolic in nature, allowing task and environment information to be mapped in a manner that is manageable for task planning and safety assessment operations. In addition the world model can be enhanced and modified through the gathered perceptual and feedback information, allowing for a varying world interpretation during operation.

Safety at the activities planning layer is managed separately from the control and task planning operation, and is based on a group of coordinated safety modules each concerned with the assessment of various safety related parameters. This complete segregation from control allows the safety modules to define aspects of perceptual and action abilities, which in turn will influence the task planning, process. Perceptual and action ability assessment is based on a number of factors. These include knowledge on the required actions, information gathered from experiences in performing such a task, and identified trends in certain sensory data. The ability to assess the risks involved in a specific action, allows the safety management modules to influence the proposed tasks to be executed, on the basis of all the available safety related information, and furthermore, on the knowledge that this information is limited by the system's constraints.

4. SAFETY WITHIN THE ARCHITECTURAL FRAMEWORK

The architectural layers allow the necessary assessment and management of safety on different temporal and spatial scales. In addition, each layer influences the assessments carried out by the other layers. The forward influence and feedback between layers ensures that the final safety assessment is coherent and based on all the knowledge available to the system. The action of one layer to contain or eliminate a potential hazardous source when fed back to upper layers, serves as an assessment of that layer's ability to determine the risk within an operation, and to assess the system's ability to perceive and act on its environment safely and reliably

Feedback provides a very powerful tool in determining not only system safety but also the efficiency of the safety modules in assessing safety. Through feedback, events encountered by the excavator will influence both the excavator's current and future behaviour through the integration of perceptual information in the topological world model. This ensures that safety related knowledge attributable to an event or state is not lost, but rather, is made available in future tasks requiring such an event or state. Indeed, it is the transfer of information back and forth through the layers which will provide for the comprehension of the perceptual and action limitations and the underlying effects on the excavator's ability to interact safely with its environment.

The varying time and spatial scales at the individual layers also allows the handling of safety aspects in diverse manners with changing safety objectives for each layer. This allows each layer to eliminate or contain encountered hazards in a different mode compared to the other layers, depending on when the hazard has been identified and how quick a reaction is required. In this manner, hazards can be identified at different stages within the planning and execution process, providing for an added ability for the excavator to handle operational risks.

A typical example of how hazards are handled in this framework is depicted in figure 2.

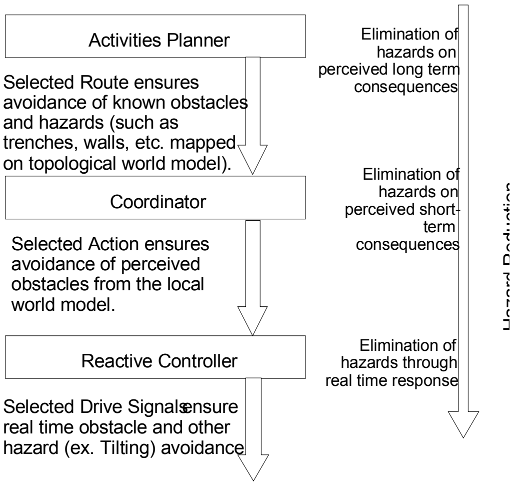

Activities Planner
Selected Route ensures avoidance of known obstacles and hazards (such as trenches, walls, etc. mapped on topological world model).
Coordinator
Selected Action ensures avoidance of perceived obstacles from the local world model.
Reactive Controller
Selected Drive Signalsensure real time obstacle and other hazard (ex. Tilting) avoidance
Elimination of hazards on perceived long term consequences
Elimination of hazards on perceived short-term consequences
Elimination of hazards through real time response
Hazard Reduction

Figure 2. A typical mode in which hazards are handled within the architectural layers

5. SAFETY INTEGRATION WITH CONTROL

One principal aspect which arises from the control architecture is the level of integration between safety and control which varies amongst the different layers. This variation is depicted in figure 3. At the lowest reactive layer, safety and control requirements are indistinguishable, with safety and control being well integrated together within all behaviours.

The level of integration, though, changes at the higher control layers. Here, the bias between control and safety of the individual modules within each layer, becomes more distinguishable. This is necessary for the following reasons;

- Separation of safety and control entities provides the required architectural subdivision for developing safety modules which, through more rigorous development procedures, are of a higher safety integrity than the control counterparts.
- Due to their higher integrity, such safety modules act as a separate 'safety conscience' at the top planning layer, being completely independent from any control decision making, and providing for the required safety assessments on which control decisions can be made.
- The grouping together of safety modules into one entity, allows the generation of a coherent, rather than fragmented, safety management procedure, which takes into consideration all safety aspects. This coherent management, allows safety judgements to influence other safety related assessments, a typical example being the influence of internal failures

influencing the interpretation of the environment and vice versa.

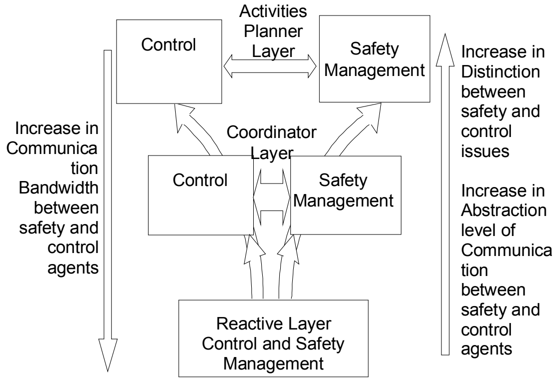

Activities Planner Layer
Control
Safety Management
Increase in Distinction between safety and control issues
Increase in Communica tion Bandwidth between safety and control agents
Coordinator Layer
Control
Safety Management
Increase in Abstraction level of Communica tion between safety and control agents
Reactive Layer Control and Safety Management

Figure 3. The progressive separation of safety from control through the architectural control layers.

- As a result of a coherent safety approach, communication requirements between individual safety modules are higher than between safety and control modules, since information sharing and influence amongst these safety modules increases heavily in the top control layer.

The net outcome is therefore an increased level of distinction and separation between entities handling control and safety aspects.

6. CONCLUSIONS AND FURTHER DEVELOPMENT

This paper has outlined the major concerns in considering safety within an autonomous excavator and has developed an approach to deal with such concerns through a hybrid architectural framework which specifically integrates safety. A layered division and the layer independence provide for better hazard management through the system's ability to visualise hazards on different spatial and temporal frames. This mode of defining safety, accommodates a representation of a safety 'awareness' or 'conscience' through its upper level, and a safety 'instinct' through its reactive level. The layered division also aids in avoiding over-reliance on a specific level or a specific module for ensuring safe operation, allowing for a more tractable and manageable design, particularly when considering issues of verification and validation.

Development of this architecture currently stands at the definition of the individual modules and their individual operational objectives. A simulation platform is also being implemented to allow an initial analysis of the specific safety concerns for each layer. It will also provide a basis on which the interaction between control and safety modules may be studied.

Further to the above, the project is also focusing on the systematic failure issues, outlining the excavator's ability to perceive its environment correctly and act on it in a safe manner. The ability to assess correct environmental perception requires the management of the underlying uncertainty within the sensory data, due to the lack of the data's completeness to generate all the required perceptual information. The handling of uncertainty as a form of managing risk will be directly embedded in the mode in which the individual modules and levels handle information.

REFERENCES

[1] International Electrotechnical Commission 'IEC 61508 – Functional Safety: Safety-Related Systems', Parts 1 to 7.

[2] Gaskill S.P. and S.R.G. Went 'Safety Issues in Modern applications of Robots' , Reliability Engineering and Systems Safety, Vol. 53 No. 3, pp. 301-307, Sep 1996.

[3] National Advanced Robotics Research Centre, 'Safety and Standards for Advanced Robots – A First Exposition', Report ARRL.92.009, July 1992.

[4] Dhillon B.S., Fashandi A.R.M., 'Safety and Reliability Assessment Techniques in Robotics', Robotica, Vol. 15, pp701-708, 1997.

[5] Graham, J.H., editor 'Safety, reliability, and human factors in robotic systems', New York: Van Norstand Reinhold, 1991.

[6] ANSI/ Robotics Industries Association, 'American National Standards for Industrial Roots and Robot Systems – Safety Requirements', R15.06-1986, Ann Arbor, MI, 1986.

[7] BSR/ Robotics Industries Association, 'Proposed American National Standard for Industrial Robots and Robot Systems – Guidelines for Reliability Acceptable Testing', AnnArbor, MI, 1993.

[8] Health and Safety Executive 'Industrial Robot Safety', Report HS/G 43, 1995.

[9] Pace C., Seward D. 'A Safety Manager for an Autonomous Mobile Robot', in Proceedings of the 29th International Symposium on Robotics ISR98, pp 277-282, May 1998.

[10] Visinsky M.L., Cavallaro J.R., Walker I.D., 'Robotic fault detection and fault tolerance: a survey' Reliability Engineeirng and System Safety, Vol.. 46, No 2, pp 139-158, 1994.

[11] Drotning W., Wapman W., Fahrenholtz J., Kimberly H., Kuhlmann J., ' System Design for Safe Robotic Handling of Nuclear Materials', in Proceedings of the 1996 2nd Speciality Conference on Robotics for challenging environments, pp. 241-247, June 1996.

[12] Arkin R., 'Behaviour-Based Robotics', MIT Press, 1998.

[13] Chung J., Ryu B.S., Yang H .S. 'Integrated Control Architecture based on behaviour and plan for mobile robot navigation', Robotica, Vol 16, pp. 387-399, 1998.

[14] Albus J., 'Outline of a Theory of Intelligence', IEEE Transactions on Systems, Man and Cybernetics, Vol 21, No. 3, pp. 473-509, 1991.

[15] Nehmsow U., ‘Self-Organisation and Self-Learning Robot Control’, IEE Colloquium (Digest), 026, 1996.