FINAL REPORT OF THE TASK FORCE ON SAFETY CRITICAL SYSTEMS

1. Introduction

The Task Force started work in September 2002 by setting-up a scoping workshop which took place the EUNITE Annual Conference in Albufeira. At that meeting it was decided to constitute a steering group jointly chaired by Prof. Paulo Lisboa (Liverpool John Moores University) and Dr. Jens Strackeljan (Technische Universität Clausthal) and comprising also:

- Prof. Ali Hessami (Atkins Consulting)
- Dr. Paul Turner (AspenTech)
- Prof. Kauko Leiviska (Oulu University)
- Prof. Derek Linkens (Sheffield University)
- Karl Lieven (ELITE Foundation, Aachen)
- Dr. George Magoulas (Brunel University).

The group was tasked to deliver:

- A road-map of EU funded projects with intelligent systems in safety-related applications, including pointers to good practice in key methodologies together with a few detailed papers on generic frameworks and case studies as appropriate.
- A questionnaire to end-users with the aim of populating the road map, validating the selected baseline methodologies, and identifying gaps in the technology transfer process that are in need of central funding.

The task force was managed in four stages:

- The first was a workshop in Albufeira, Portugal, to set the terms of reference for the work;
- the second consisted of a survey of EU funded research in safety-related areas, to ascertain good practice and compile the road map;
- the third stage was a workshop in Oulu, Finland, to discuss the findings of the survey with researchers engaged in state-of-the-art methodologies;
- finally there was a questionnaire to verify the Task Group findings with the user community.

2. Road map of EU funded projects

A survey of the *CORDIS* database was carried out to identify EU funded projects in safety-critical areas. This showed a preponderance of funding in the automotive industries, process industries and environmental monitoring, which could be organised into a road map shown in fig. 1.

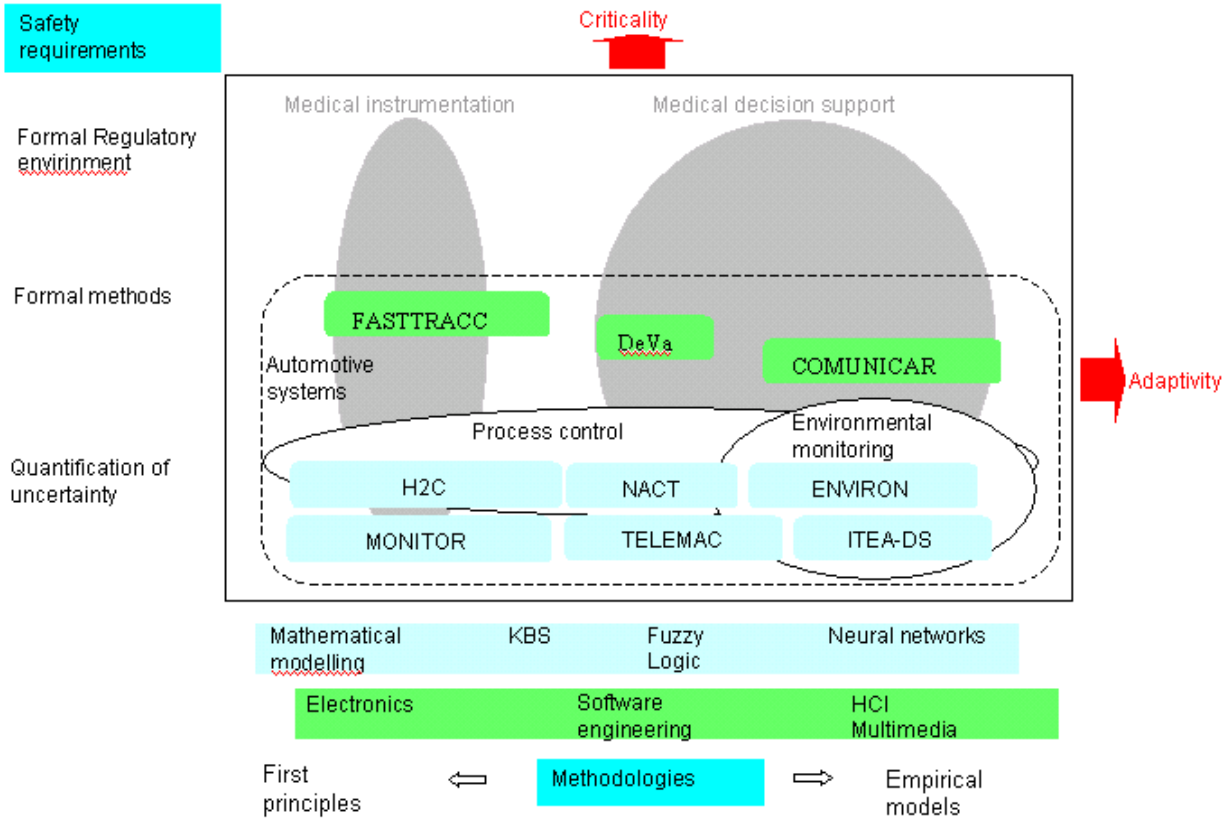

Figure 1. Organisation of EU funded projects from the CORDIS database by order of increasing adaptivity in terms of user interaction and safety requirements towards requiring a formal safety case. The administrative details of the projects are listed in Appendix 1.

Notice that control of adaptivity is progressing along two parallel tracks, one of which is the design of algorithms towards self-learning, and the other the development of increasingly versatile software to manage user interactions. At the same time, EU funded projects are developing formal methods capable of assuring safety in a range of essentially different methodologies, from monitoring through identification of dynamic systems, to closed-loop control. These projects have mostly resulted in developments of the state-of-the-art that were successfully transferred into top-of-the-range industrial design.

In contrast, funding for formal methods of software and hardware design appear to have been less taken-up their developers.

The road map should also be read in the context of other more established safety-critical developments including:

1. Specialist products already in commercial use, including the fire detector FirePrint. These products have been reviewed elsewhere (http://www.hse.gov.uk/research/crr_pdf/2001/crr01327.pdf).

2. Biometric and biomedical applications, including neuromorphic modelling for access control using iris identification, and a wide range of medical applications some of which have also been reviewed elsewhere (Lisboa, P.J.G. 'A review of evidence of health benefit from artificial neural networks in medical intervention', Neural Networks, 15, 1, 9-37,2002).

These areas are indicated broadly in the road map but they were not followed-up in detail by this task force since research in biomedicine is proceeding along well-coordinated networks both within and outside the EU funding framework. Consequently the task force focused on issues that are current and contentious, with the aim of identifying targeted areas in need of further development funding.

3. Project details

This section provides pointers to EU funded research carried out over the last 10 years that was identified in the road-map survey.

a. Automotive industry

i. FASTTRACC: Formal analysis and specification of hardware and software design for cruise control. This work seems to have been discontinued after termination of funding.

ii. **NACT, H2C:** The two projects were run contiguously and broke new ground in principled theoretical frameworks for the design of neural networks for identification and control of dynamical systems and for heterogeneous control:

→ Neural Adaptive Control Technologies, spearheaded the development of practical public domain software for the design and use of Local Model Network/Local Model Controller methods, which were integrated into other applications at Daimler-Benz.

→ Hybrid Heterogeneous Control (H2C) , that is the control of systems with combined continuous and discrete components, led to proofs of stability using convex dynamic programming, for the design of LQR controllers for piecewise linear with state and input constraints. These methods were demonstrated in an actual vehicle and benchmarked against the gold standard for the automotive industry.

iii. COMUNICAR: Integrated on-vehicle information with multimedia HMI

b. Intelligent control and scheduling

i. TELEMAC: This is a current project on telemonitoring and remote control for wastewater treatment plants. The aim is to introduce new

technologies into what is a very conservative field, that of water treatment, to enable reliable plant supervision by a remote expert aided by advanced methods for automatic control fault detection and isolation. Fuzzy logic methods are already being trialled in anaerobic reactor pilot plant.

MONITOR: The objective of this project was to develop automatic monitoring systems capable of driving preventative maintenance programmes for reciprocating engines used in shipping, power generation and petroleum refining. It focused on multisensor fusion and included signal processing methods such as wavelet analysis for feature-based fault identification.

ii. **ITEA-DS:** This is also a recent project that addresses important situation assessment and environmental monitoring in emergencies in the context of a regulatory pressure fro greater accountability of crews of vessels navigating congested waters, set against commercial pressures to reduce ship crew sizes.

c. **INTAS:** This project is focused on environmental monitoring with machine learning algorithms. It has used recent developments in Support Vector Machines to produce robust ozone and air pollution forecasting, including model selection with automatic hyperparameter tuning which is regarded as an example of good practice for these advanced inference methods. Another methodological development led by this consortium is wavelet analysis residual kriging to model multi-scale correlation and non-stationarity in spatial models for environmental monitoring and risk assessment.
These methods can be used to validate the normality assumptions made by generic non-linear regression models, such as traditional neural networks. The use of residual analysis as a diagnostic tool to monitor the goodness of fit of non-linear models is good practice but is currently not widespread.

d. **DeVa: Software design for validation involving reusability, object-oriented software design, fault tolerance.**

4. Workshop at EUNITE 2003

The workshop was well attended, with 18 participants. The programme was centred on the presentation of key deliverables for the report, starting with the road map introduced at the start of the workshop, and moving onto three key papers, outlining a generic framework for elicitation and risk modelling within a graphical network, a framework to assure neural networks for decision support, and a practical industrial case study to points out important inherent limitations in the use of standard neural network architecture for closed-loop control.

10.07.2003, 14:00-16:00	AC3 Room: Linna
Task Force Meeting: Safety Critical Systems	Chair: Paulo J.G. Lisboa, United Kingdom

Introduction to the Workshop

P.J.G. Lisboa, Liverpool John Moores University

Safety Management of ComplexTechnology

A. Hessami, Atkins

Developing Artificial Neural Networks for Safety Critical Applications

Z. Kurd, York University

Automated Product Grade Transitions: Exposing the inherent and latent dangers of Neural Networks in Manufacturing Process Control

P. Turner, AspenTech

Open Discussion

The workshop closed with a discussion of the main contentious remarks made by the speakers, namely that:

- assuring neural networks for decision support inevitably pushes us towards symbolic level verification, requiring rule extraction;
- closed-loop control of safety-related plant should use transparent models either based on first principles, see NACT and H2C later, or with explicit control of model gain, as in the Bounded Derivative Network proposed by Paul Turner, or using Takagi-Sugeno fuzzy logic. The discussions on closed-loop control concluded that the difficulties with vanishing and inverting gains were a characteristic of the feedforward neural networks, but are better handled by other technologies including fuzzy logic control.

A write-up of the papers presented at the Workshop is in Appendix 2.

5. Conclusions

The road map and pointers to good practice debated at the workshop formed the basis for an end-user questionnaire to validate the conclusions reached in Oulu, which yielded the following conclusions:

- MLP and fuzzy logic models applied to industrial process control and medical applications including anaesthesia.

- In decision support, neural network design should include generalisation control preferably through a Bayesian framework evaluated either with Markov Chain Monte Carlo methods or using the evidence approximation. Error bars for regression are regarded as essential in safety-related applications, while rule extraction is required to assure assignment to binary categories.

- In process control, classical models for identification and control remain essential. These are often explicitly built into fuzzy logic models and feature in the Local Model Network/Local Model Controller, as noted above. However, some form of gain control is seen as essential to prevent interpolation errors in model identification from causing unwanted control actions. The difficulties with vanishing and inverting gains are a characteristic of the feedforward neural network, and are better handled by other technologies such as fuzzy logic or models with analytically constrained gains.

- Substantial testing across the operating envelope is a requirement for verification and validation of black-boxes utilised for closed-loop control in safety-related applications. Stability proofs should also be a requirement but in non-linear identification and control this is largely an open question.

- Risk analysis should ideally be implemented with probabilistic graphical models evaluated by Monte Carlo sampling.