Design for Safety in Safecharts with Risk Ordering of States

Nimal Nissanke and Hamdan Dammag

School of Computing, Information Systems and Mathematics, South Bank University, 103 Borough Road, London SE1 0AA

Abstract

Safecharts is a variant of Statecharts intended exclusively for safety critical systems design. With two separate representations for functional and safety requirements, Safecharts brings the distinctions and dependencies between them into sharper focus, helping both designers and auditors alike in modelling and reviewing safety features. Safecharts incorporates ways to represent equipment failures and failure handling mechanisms and uses a safety oriented classification of transitions and a safety oriented scheme for resolving any unpredictable non-deterministic pattern of behaviour. It achieves these through an explicit representation of risks posed by hazardous states by means of an ordering of states and a concept called risk band. Recognising the possibility of gaps and inaccuracies in safety analysis, Safecharts do not permit transitions between states with unknown relative risk levels. However, in order to limit the number of transitions excluded in this manner, Safecharts provides a default interpretation for relative risk levels between states not covered by the risk ordering relation, requiring the designer to clarify the risk levels in the event of a disagreement and thus improving the risk assessment process.

Key words: Safety, Statecharts, Risk ordering, Risk assessment, Failures

1 Introduction

Provision of safety in critical systems is based on a combination of well–known strategies such as avoidance, elimination, prevention, evasion and tolerance of faults that could potentially contribute to hazards. These strategies address faults that could arise in different stages of the life cycle, some in

Email addresses: nissanke@sbu.ac.uk (Nimal Nissanke), dammagh@sbu.ac.uk (Hamdan Dammag).

pre-operational development stages while others during the system operation. Since they can have different and, sometimes, conflicting design goals, the design process is fraught with complexities. A systematic approach to design is, therefore, essential. In addition to being based on sound engineering principles, there are several other desirable attributes of such an approach, including the means to facilitate good practice and to address the needs specific to safety critical systems design. These have been the main aims of Safecharts (Dammag *et al.*, 1999), (Nissanke *et al.*, 2000) – a visual formalism based on Statecharts (Harel *et al.*, 1985) (Harel, 1987) and developed exclusively for safety critical systems design.

Turning to the attributes mentioned above with respect to systematic design, Safecharts addresses an important need in safety critical systems design, namely, means to represent run–time failures and safety mechanisms. The result is an integrated framework where functional design and safety design can progress hand in hand, helping to combat pre-operational faults in both kinds of design. Safecharts seeks to promote sound engineering principles through the adoption of Statecharts - a well established formalism in the area of reactive systems. In this respect, Safecharts shares in common the traditional virtues of Statecharts, namely, visual appeal, modular and hierarchical representation of systems and mathematical rigour. Its strategy is to use Statecharts as done conventionally to represent functional requirements and to provide an extended notation to capture safety requirements. Safecharts distinguishes the use of these two notations in terms of a *functional layer* and a *safety layer* and concentrates its effort exclusively on the features of the latter. Demarcation drawn in this manner between function and safety helps focusing on safety matters without being distracted by functional issues, the evaluation of implications of function on safety and the assurance of safety provisions for each and every action involving any risk. Conservative assumptions made in Safecharts with respect to missing information and non-deterministic outcomes alerts the designer to their implications and, if appropriate, to take corrective measures. These are means through which Safecharts strives to facilitate good practice.

A key feature of the safety layer is the ordering of states according to risks posed by them, reflecting both the likelihood of hazards and the magnitude of their potential consequences. One of the uses of the resulting risk ordering relation, or the *risk graph*, is the characterisation of transitions according to the nature of relative risk posed by their target and source states. This results in three categories of transitions: *safe*, *unsafe* and *neutral* transitions. Recognising the role of human competence in safety analysis and, as a result, the possibility of gaps and inaccuracies in such analysis, Safecharts do not permit, on the grounds of prudence, transitions between states with unknown relative risk levels. As a result, many transitions, including those which can be functionally useful, could be potentially excluded between states lying in sparsely covered areas of the risk graph. Safecharts relieves the effect of this

restriction by providing a default interpretation for relative risk levels between states not covered by the risk ordering relation. This is done through the concept of *risk band*. The default interpretation places a burden on the designer to clarify the risk levels in the event of a disagreement, thus bringing about a refinement of the risk graph to a level of accuracy required by the design.

2 Safecharts

2.1 Statecharts – An Outline

Statecharts forms the basis of Safecharts; see (Harel *et al.*, 1985) (Harel, 1987) for more details on Statecharts. There are three categories of states in Statecharts: AND, OR and BASIC. An AND-state, or a OR-state, consists generally of two or more substates. In diagrams, substates of AND-states and OR-states are distinguished by dashed and solid lines respectively. Being in an AND-state means being in all of its substates simultaneously, while being in an OR-state means being in exactly one of its substate. A BASIC state is a state with no substates. Statecharts thus enables the construction of a hierarchy of states and extends conventional finite state machines by AND/OR decomposition of states. Changes in state are brought about by transitions. In an OR-state, for example, a transition may bring about a distinct change in the state by moving the currently active substate to another substate. Transitions are shown as arrows from one state to another and are labelled. The most general form of labelling is $e[c]/a$, e being an event that triggers the transition, c a condition that guards the transition when e occurs, and a an action that is carried out precisely if and when the transition takes place. Once generated, the action a is broadcast to the whole Statechart, triggering, if applicable, other transitions in the system. A default state, pointed by a short arrow, is a substate of an OR-state to be entered by any transition if its arrow terminates on the boundary of the OR-state concerned.

2.2 Features of Safecharts

As was mentioned in Section 1, Safecharts maintains two separate layers in the representation of any system. The *functional layer* specifies the transformational behaviour of the system purely from a functional point of view using Statecharts in the conventional sense, whereas the *safety layer* is devoted exclusively to safety issues. Central to various features of the safety layer is an ordering of system states according to risks posed by them. Mathematically, it is a relation and is denoted by $\sqsubseteq$. It is also referred to as a

risk graph. Treatment of transitions and default states and representation of failure handling safety mechanisms are all based on this ordering of states. The relation $\sqsubseteq$ may consist of pairs of states which are known to be either of two distinct risk levels or of an identical risk level. This can also be represented mathematically by decomposing $\sqsubseteq$ into two relations: a partial order relation and an equivalence relation, denoted by $\preccurlyeq$ and $\approx$ respectively. The interpretation of this notation is such that, given two distinct states s_1 and s_2,

- $s_1 \sqsubseteq s_2$ – the risk level of s_1 is known to be lower than, or equal to, the risk level of s_2.
- $s_1 \preccurlyeq s_2$ – the risk level of s_1 is known to be strictly lower than that of s_2.
- $s_1 \approx s_2$ – the risk levels of s_1 and s_2 are known to be identical.

Based on the risk graph, Safecharts classifies transitions according to the nature of risks they carry and, accordingly, extends the specification (labelling) of transitions with additional guards and enforcement conditions. Thus, transitions belong to three categories: *safe* (from a higher risk state to lower risk state), *unsafe* (from a lower risk state to higher risk state) and *neutral* (between states of the same risk level). Being an exhaustive classification, this precludes transitions between states the risk levels of which are unknown (i.e. non–comparable by $\sqsubseteq$). The reasoning behind this principle is prudence and it is intended to prompt the designer to resolve, as a matter of discipline, the risk levels of any non-comparable states, if a transition is desired between them.

Transition labelling in Safecharts has the general form $e\,[c]/a\,[l,u)\,\Psi[G]$, with e, c and a remaining the same as in Section 2.1 and certain components being mandatory depending on the risk classification of the transition concerned. $[l,u)$ is a right-open time interval from time l to time u. Ψ is a safety enforcement pattern specified using two alternative symbols: $\upharpoonleft$ and $\upharpoonright$, and $[G]$ is a safety clause. $t \upharpoonleft [G]$ is mandatory for unsafe transitions and means that the transition t is forbidden to execute as long as G holds. $t[l,u) \upharpoonright [G]$ is mandatory for safe transitions and means that the transition t is forced to execute within $[l,u)$ from whenever G begins to hold irrespective of the occurrence of its triggering event.

Turning to failures, each component is represented in the form of an OR-state with two distinguished substates, denoted generically by IN and OUT, meaning respectively that the component is functioning correctly or has failed. The nature of these two states are such that IN $\preccurlyeq$ OUT. Associated with these states are also two generic events: a nondeterministic event ε signifying a failure, and an event μ signifying a maintenance or repair action which returns the component back to service. It follows from the above that ε triggers an unsafe (low-to-high risk) transition, while μ triggers a safe (high-to-low risk) transition. A component may have more than one failure mode, in which case OUT may itself be an OR-state with a distinct substate for each of the failure modes, possibly with further transitions to model failure propagation.

3 Risk bands

3.1 Risk and the Need for a Risk Ordering Relation

Despite our focal interest in risk, below is a brief outline of several facets of risk (Redmill, 1997) of immediate relevance here. Risks are associated with undesirable events, or accidents, capable of causing injury or danger to human life or damage to property or the environment. In the context of this work, such events result primarily from breaches of reliability (system or equipment failures). Risks involved are primarily 'speculative', undertaken by choice in pursuit of a service (reward). Risks consist of two contributory factors: the likelihood (probability) of occurrence of a hazard concerned and scale of damage or consequences, each contributing proportionately to the severity of overall risk. In relation to the risk ordering relation, two particularly important risk management activities are risk analysis (analysis of causes and quantitative or qualitative assessment of consequences) and risk prioritisation (judicious discrimination between relatively close estimates of severity of risk on the basis of social or economic preferences). These two tasks are of crucial importance in putting the design approach advocated in this paper into practice.

Risk management strategies include, amongst others, avoidance, elimination or reduction, and acceptance of risk. In the context of this work, they take the form of several specific categories of safety requirements aimed at reducing the exposure to risk, namely, through a) safe initialisation when the system or equipment is returned to service, b) prohibitive conditions attached to transitions taking the system from a low risk state to a high risk state to deliver a service, c) mandatory timely transitions taking the system from a high risk state to a low risk state after delivering a service, and d) fail–safe and fail–soft patterns of behaviour in response to equipment failures. An implication of such safety requirements is that relative risk levels posed by different states are known, at least in the case of source and target states of transitions, thus justifying the need for a risk ordering relation such as $\sqsubseteq$ introduced earlier.

3.2 The Need for Risk Bands

Role of human competence in risk analysis and risk prioritisation has certain implications, especially in relation to the accuracy and the completeness of the resulting risk graph and the uniformity of coverage of the state space by the risk graph. A point in question is the possibility of extreme cases, where a few states may happen to be non-comparable with a large number of other states by the risk ordering relation, though the states in the two groups

concerned may themselves be mutually comparable with one another. Based on an implicit assumption that the arcs in the graph run upwards, Figure 1(a) gives an example, where the state I happens to be non-comparable with the states D, E, F, G, H and C, and the state B with the states A, D, E, F, G, H and C. State C is also non-comparable with the states D, E, I and B. Consequently, if reliance is placed solely on $\sqsubseteq$, the principle of barring transitions between non-comparable states will exclude transitions in either direction between the above mentioned states. The concept of *risk band* is designed to overcome such situations, if such an extensive exclusion of transitions is undesirable.

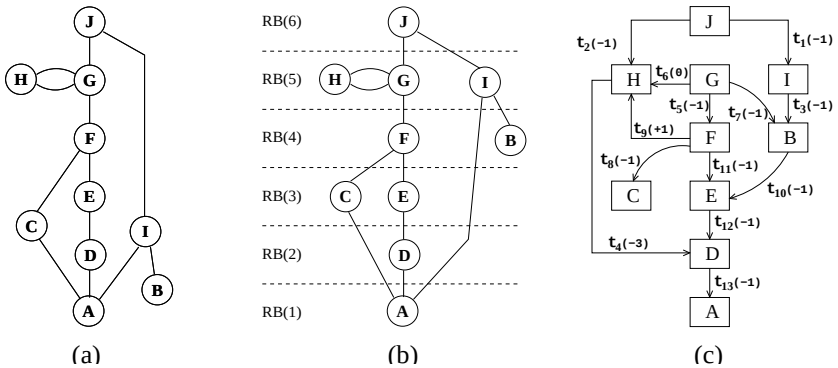

Fig. 1. Non–comparable states in a risk graph, risk bands and risk distances.

3.3 Definition of Risk Bands

Risk bands are an enhancement of the risk ordering relation $\sqsubseteq$. By definition, each state belongs to a unique risk band and every pair of distinct states belonging to the same risk band is there either because the states concerned are comparable by $\approx$ (explicitly stated to have an identical risk level) or non-comparable by $\sqsubseteq$. The intended use of risk bands is to allow all transitions between pairs of states either belonging to different risk bands, or belonging to the same risk band but comparable by $\approx$. Therefore, excluded transitions are only those between states residing in the same risk band but not comparable by $\approx$. Risk bands are thus a default scheme for ranking states according to risk levels when $\sqsubseteq$ is inadequate on its own. Obviously, unacceptability of such a default interpretation of risks should prompt the designer to reassess the risk levels of the states concerned more accurately.

Given the risk ordering relation $\sqsubseteq$, and assuming that risk bands are indexed numerically from 1 to some n, risk bands of states may be defined according to the following set of rules:

(i) States in the highest risk band n consists of exactly a) maximal elements (states) in the partial order relation $\preccurlyeq$ but excluding those elements, if any, which are comparable by $\approx$ with any of the rest of elements in $\preccurlyeq$, and b) elements which are comparable by $\approx$ with those elements defined in (a) above.

(ii) Any state s with just a single immediate (distinct) successor state, which is in risk band i according to $\preccurlyeq$, is in risk band $(i-1)$. However, if a state s has more than one immediate successor state, then it has a risk band one less than the lowest of the risk bands of its immediate successor states.

(iii) For any states s_1 and s_2, if $s_1 \approx s_2$ then both states s_1 and s_2 are in the same risk band.

Figure 1(b) illustrates the above rules using the risk graph given in Figure 1(a). As a result, transitions such as those between state I and F can now be introduced. The number of prohibited transitions can thus be reduced significantly, in this case, to transitions between the pairs of states: I and G, I and H, B and F, and C and E. Position of state B in Figure 1(b) is significant; due to the lack of knowledge about the risk posed by the state B, it is now considered to be of a higher risk nature compared to the states C, E, D and A.

3.4 Uses of Risk Bands

Risk bands allow a novel safety oriented approach to resolution of nondeterminism between simultaneously enabled conflicting transitions. Approaches such as (Pnueli *et al.*, 1991) resolve such nondeterminism on the basis of scope of transitions and (Day, 1993) on the basis of the hierarchy of their source state. As a safety oriented improvement, our previous work (Dammag *et al.*, 1999) suggested prioritisation of transitions based on $\sqsubseteq$ according to the risk level of their target states so that lower the risk level of its target state, higher is the priority enjoyed by a given transition. A deficiency of using just $\sqsubseteq$ is that two conflicting transitions with a common source state would still enjoy the same priority if their target states happen to be non-comparable by $\sqsubseteq$, irrespective of the relative positions of the latter states in the banded risk graph.

In this respect, risk bands result in a new concept called *risk distance*. For a given transition, risk distance is the risk band index of its target state minus that of its source state, the positive and negative signs signifying respectively an increasing and decreasing risk. Any nondeterminism between two or more transitions can now be resolved by giving higher priority to the transition with the shortest risk distance. In the case of transitions with equal risk distances, prioritisation may be based on the cumulative risk distances of future transitions of conflicting transitions, i.e. those that could be triggered in one or more specific number of subsequent steps. In the case of several future transitions, the transition with the shortest risk distance is to be considered for comparison with other competing future transitions. Nondeterminism may still continue to persist even with future transitions, but this kind of nondeterminism is considered a *safe nondeterminism* since all outcomes are identical in terms of the risks involved.

As an illustration, for a set of transitions introduced between states in Figure 1(b), Figure 1(c) shows their risk distances in parentheses. Here, due to its shorter risk distance, the transition t_5 will enjoy higher priority over t_6, if they happen to be in conflict. In the case of t_1 and t_2 being in conflict, it is necessary to consider their future transitions in the next step, i.e. t_3 in the case of t_1 and t_4 in the case of t_2. Since the cumulative risk distances of their future transitions are -2 and -4 respectively, the transition t_2 enjoys higher priority over t_1. However, non-deterministic situations could still occur. This is the case when t_5 and t_7 are in conflict since the cumulative risk distances covering their future transitions (t_8, or t_{11}, for t_5 and t_{10} for t_7) are the same.

The concept of risk band subsumes the concept of risk ordering relation. As a result, the classification of transitions into *safe*, *unsafe* and *neutral* transitions can now be based on risk bands, rather than on the risk ordering relation as in (Dammag *et al.*, 1999). According to the new definitions, safe transitions have negative risk distances, unsafe transitions positive risk distances and neutral transitions zero risk distances.

4 A Case Study Illustrating Safecharts – A Nuclear Reactor

As an illustrative case study, this section considers certain aspects of a nuclear reactor of a power plant. The nuclear reaction, taking place in a bundle of fuel rods called the *core*, is controlled by cadmium *control rods* held above the core by magnetic clamps. In an emergency, control rods are released into the core in an action called *scram* to halt the nuclear reaction. The water circulating through *primary* circuit, passing through the reactor, extracts the heat from the nuclear reaction. It acts as a coolant and prevents fuel rods from overheating. The heat so extracted from the core is subsequently passed on to the water circulating through the *secondary circuit*, located outside the reactor, creating the steam used in the generation of electricity. Maintenance of correct temperature and pressure is critical for the processes involved. For example, loss of coolant in the primary circuit – an accident referred to as LOCA – can lead to extreme temperatures and pressures, damaging the reactor building and causing release of radiation. In the case of such an event, operation of a *relief valve* in the vessel can relieve the pressure, by letting the coolant to flow out into a safe drainage system.

Several of the above components played a critical role in the accident at the Three Mile Island (TMI) nuclear power plant in 1979; see (Bignell *et al.*, 1984) and (Leveson, 1995). A closure of a valve in the secondary circuit stopped the process of heat exchange between the two circuits, leading to an abnormally high temperature first in the coolant of the primary circuit and then in the reactor. As expected, the pressure relief valve opened and the control rods

dropped, bringing the conditions inside the reactor to an acceptable level. At this stage, the pressure relief valve should have closed, but instead malfunctioned and remained stuck-open. This went undetected by the operators because of a misleading sensor, attached to a solenoid operating the valve rather than to the valve itself.

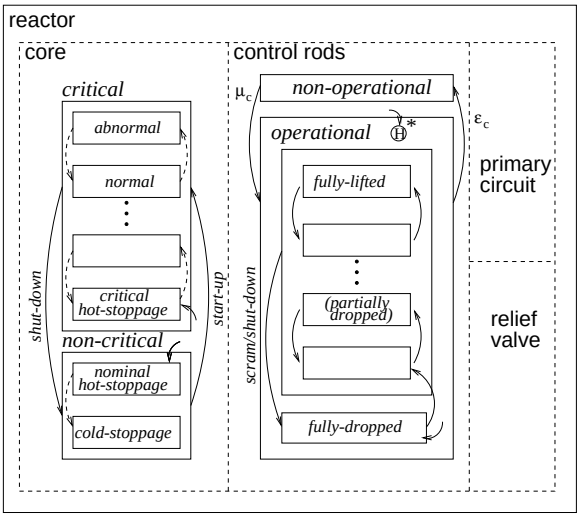

Fig. 2. Safechart model of reactor core

A Safechart representation of the reactor, shown in Figure 2 with certain simplifications and conventions [1], consists of several AND-ed components: core, control rods, relief valve and primary circuit, with the latter two expanded in Figures 3 and 4 respectively. The core consists of two OR-substates: *critical* and *non-critical*, each decomposed into further substates. Transitions triggered by *start-up* and *shut-down*, brought about by the movements of control rods, enable alternations between the *critical* and *non-critical* substates. The state *control rods* consists of two substates: *non-operational* (a failure state) and *operational*. The latter has a safe state for initialisation. The state primary circuit, shown in Figure 3, consists of attributes such as pressure, temperature, coolant content, etc., and not components (valves, pumps, etc.) which bring about changes in those attributes.

In relief valve (PRV), Figure 4(a), the failure state *stuck* consists of two substates: *stuck-closed* and *stuck-opened*, entered depending on the generic event ϵ_v. The sensor, modelled as part of the relief valve, picks up failures of the valve through the static reaction PRV-*stuck* and, similarly, its current status through the static reactions PRV-*working*, PRV-*close*, etc. In tracing the chain

[1] For brevity, diagrams show an integrated view of the functional and safety layers. In any OR-state, higher risk states are placed higher in its diagram compared to lower risk states. Transition labelling is kept to a minimum. Dashed line arrows show internal transitions that can be observed but not controlled, whereas solid line arrows show controllable transitions.

of events that led to the TMI accident in our model, the time when the high temperature in the coolant of the primary circuit triggered the opening of the relief valve corresponds to reaching the state *excessive* in the substate *temperature* (*pressure*) of *primary circuit*. The occurrence of the event *up* (and *down*) is established by the monitoring equipment. The relevant transition generates the actions *open* in PRV and *scram* in *control rods*. The state *opened* in PRV has a lower risk level than the state *closed*. The event *open*, therefore, triggers a safe transition with a mandatory time limit $[l, u)$ and generates the event *drain* in *primary circuit*. As the situation returned to a more acceptable state, the event *down* in primary circuit should have taken place, triggering the event *close* in PRV. However, another possible event is the occurrence of the generic event ϵ_v, an event not bound by the rules on prioritisation of non-deterministic transitions. This event would have resulted in the static reaction PRV-*stuck* and the sensor moving to *prv-out*, letting the operators take an appropriate action to avert, or mitigate, any accident. In the TMI accident, this failure was not detected due to the form of attachment of sensors.

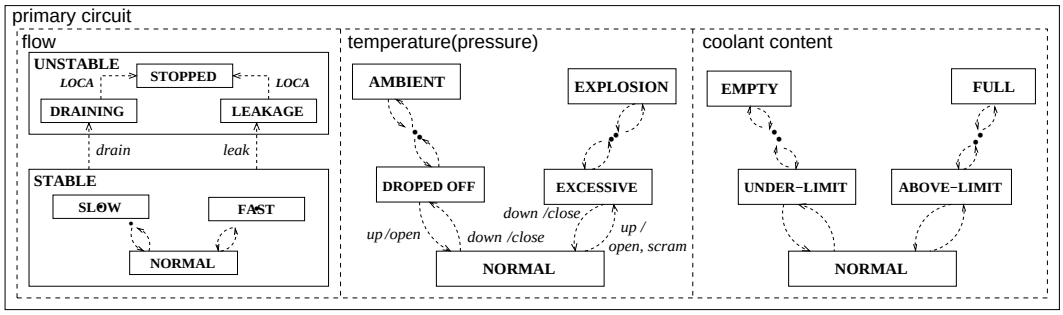

Fig. 3. Safechart model of primary circuit

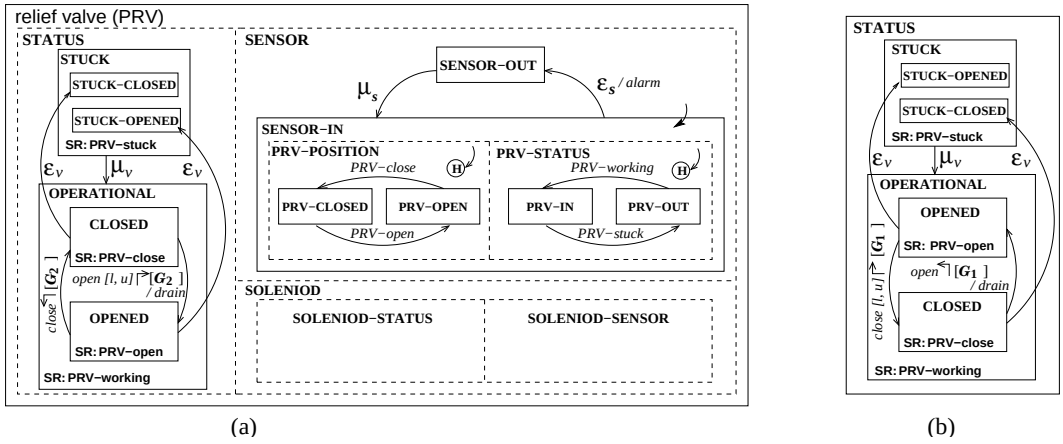

Fig. 4. Safechart model of relief valve

Relative risk levels of the states *opened* and *closed* of PRV are not fixed and alternate from situation to situation. This can be captured by replacing dynamically the state *status* in Figure 4(a) with that in Figure 4(b), which consists of a different risk graph. Dynamic fluctuation of risk levels in this manner is due to *situational events* – an issue under current research.

5 Conclusions

In formalisms intended for safety critical systems design, the capability to explicitly represent risks posed by hazardous states can have a number of benefits. These include greater alertness to safety issues in the design process and effective means to ensure a more comprehensive and a correct capture of safety requirements than what could be achieved otherwise. This has been demonstrated here in Safecharts – a novel variant of Statecharts proposed in (Dammag *et al.*, 1999) for specification and design of safety critical systems. This capability has wider relevance in similar state–based formalisms such as Petri nets. Representation of risk is based on a *risk ordering relation*, enhanced by a concept called *risk band* for compensating conservatively for any gaps in the risk assessment process. The two concepts together provide a formal framework for safety oriented classification of transitions, safety oriented definition of default states, representation of safety mechanisms and safety oriented resolution of any non–determinism between conflicting transitions. Areas under current research include the study of *situational events* - events that dynamically alter the risk ordering relation - and a mathematical definition of the semantics of Safecharts.

References

V. Bignell and J. Fortune. Understanding Systems Failures. Manchester University Press. 1984.

H. Dammag and N. Nissanke. Safecharts for specifying and designing safety critical systems. 18th IEEE Symposium on Reliable Distributed Systems. Lausanne. IEEE. October 1999.

N. Day. A model checker for Statecharts. Technical report. University of British Colombia. Vancouver. Canada. 1993.

D. Harel, J. P. Schmidt, and R. Sherman. On the formal semantics of Statecharts. 2nd IEEE Symposium on Logic in Computer Science. 1985.

D. Harel. Statecharts: a visual formalism for complex systems. Science of Computer Programming. Vol. 8. pp 231–274. North-Holland. 1987.

N. G. Leveson. Safeware – System Safety and Computers. Addison-Wesley Publishing Company. 1995.

N. Nissanke and H. Dammag. Risk Ordering of States in Safecharts. Safecomp, Rotterdam. LNCS. Vol. 1943. Springer-Verlag. 2000.

A. Pnueli and A. Shalev. What is in a step: On the semantics of Statecharts. Symposium on Theoretical Aspects of Computer Software. LNCS. Vol. 526. Springer-Verlag. 1991.

F. Redmill, Practical risk management, in F. Redmill and C. Dale (eds.), Life Cycle Management for Dependability, Springer, 1997.