Modelling Temporal Behaviour in Complex Socio-Technical Systems

A. Burns
Department of Computer Science
University of York, UK.

I.J. Hayes
School of Information Technology and Electrical Engineering
University of Queensland, Australia.

G. Baxter
Department of Psychology
University of York, UK.

and

C.J. Fidge
School of Software Engineering and Data Communications
Queensland University of Technology, Australia.

Abstract

This report is concerned with the role time plays within any structural description of a complex computer-based system. Indeed this report explores the benefits of placing time at the centre of any description of system structure. To exploit the unique properties of time, with the aim of producing more dependable computer-based systems it is desirable to explicitly identify distinct *time bands* in which the system is situated. Such a framework enables the temporal properties and associated dynamic behaviour of existing systems to be described and the requirement for new or modified systems to be specified. A system model based on a finite set of distinct time bands is developed in the report.

Acknowledgements This work has been undertaken as part of the DIRC project, funded by the EPSRC in the UK. Ian Hayes is supported by the Australian Research Council (ARC) Centre for Complex Systems and Colin Fidge is supported by ARC Discovery grant DP0449773.

1 Introduction

The construction of large socio-technical systems imposes a number of significant challenges, both technical and organisational. Their complexity makes all stages of their development (requirements analysis, specification, design, implementation, deployment and maintenance/evolution) subject to failure and costly re-working. Even the production of an unambiguous behavioural description of an existing system is far from straightforward.

One characteristic of these computer-based systems is that they are required to function at many different time scales (from microseconds or less to hours or more). Time is clearly a crucial notion in the specification (or behavioural description) of computer-based systems, but it is usually represented, in modelling schemes for example, as a single flat physical phenomenon. Such an abstraction fails to support the structural properties of the system, forces different temporal notions on to the same flat description, and fails to support the separation of concerns that the different time scales of the system facilitate. Just as the functional properties of a system can be modelled at different levels of abstraction or detail, so too should its temporal properties be representable in different, but provably consistent, time scales.

To make better use of 'time', with the aim of producing more dependable computer-based systems, we propose a framework that explicitly identifies a number of distinct *time bands* in which the system under study is situated. The framework enables the temporal properties of existing systems to be described and the requirement for new or modified systems to be specified. The concept of time band comes from the work of Newell [16] in his attempts to describe human cognition. Newell focuses on hierarchical structures within the brain and notes that different time scales are relevant to the different layers of his hierarchy. By contrast, we put the notion of a time band at the centre of our framework. It can then be used within any organisational scheme or architectural form — for they all lead to systems that exhibit a wide variety of dynamic behaviours.

In this report we first give an informal description of the framework and its time bands. We then take the key properties of the framework and produce a formal model. The motivation for doing this is two-fold. First to clarify the semantics of the framework; and second, to define the analysis necessary to prove that the bands, when taken together for a given system, are consistent (i.e., a description of some action in one band does not contradict a description of the same, or a related, action in a different band). Note that the number of bands required and their actual granularity is system-specific; but the relationships between bands, we contend, exhibit important invariant properties. The report concludes by providing a precise notation for the model. Future work will extend this into a formal logic by adding indications of how the model can be used to expresses properties of behaviours and mappings that are consistent within the framework. First, however, we give more details on Newell's notion of time bands.

Newell's Notions of Time Bands

Newell [16] starts from the viewpoint that intelligent systems are necessarily comprised of multiple levels of systems. He maintains that human cognitive architecture must also be structured in the same way, as a hierarchy of system levels with distinct time scales.

Each system level comprises a collection of components that are connected and interact to produce behaviour at that level. Where a system has multiple levels, the components at one level may be realised by systems at the next lower level. Each

Scale	Time Units	System	World(theory)
10^7	Months		
10^6	Weeks		Social Band
10^5	Days		
10^4	Hours	Task	
10^3	10 Minutes	Task	Rational Band
10^2	Minutes	Task	
10^1	10 Seconds	Unit task	
10^0	Second	Operations	Cognitive Band
10^{-1}	100ms	Deliberate act	
10^{-2}	10ms	Neural circuit	
10^{-3}	ms	Neuron	Biological Band
10^{-4}	100 microseconds	Organelle	

Table 1: Newell's Time Scales of Human Action

level is an abstraction that hides some of the detail of the next lower level. Levels can be stronger or weaker, depending on how well the behaviour at that level can be predicted or explained by the structure of the system at that level. Strong levels are state determined, in that future behaviour is determined by the current state at that level only. For weak levels, future behaviour is at least partly determined by considerations from lower levels.

Newell suggests that a factor of very roughly 10 (he uses the notation $\sim\sim$10) is required to produce a new level. In other words, the number of components and the component times at adjacent levels differ by $\sim\sim$10. As one moves up the hierarchy the size of the components increases (as a geometric progression), and the time taken to produce an output increases (also as a geometric progression). Table 1 gives the time scales, in seconds, identified by Newell.

Each of the time bands comprises a number of time scales. The different bands are characterised by different phenomena, as shown by the right hand column of the table, and are explained by different theories.

Each level is $\sim\sim$10 above its components. In the Biological Band, this is based on empirical evidence, and offers support for Newell's levels analysis. In the Cognitive Band, it is taken more as a prediction.

In the Biological Band, the systems for the different time scales have names already (Organelle, Neuron, and Neural Circuit); in the Cognitive Band, Newell suggests the systems are Deliberate Acts, Operations and Unit Tasks, and focuses most of his attention on establishing the cognitive levels, using available empirical evidence.

The neural system does not have enough time available to produce fully cognitive behaviour, which begins to become observable in the order of seconds ($\sim\sim$1 s). Newell attributes this to there being a real-time constraint on cognition, which was previously noted by other people:

> Real-time constraint on cognition: The principle is that there are available only about 100 operation times (two minimum system levels) to attain cognitive behavior out of neural circuit technology. (Newell [16, p. 130])

Although Newell uses the terms time scale and time band in his descriptions, we use the single notion of a **band** to represent a distinct temporal level in any system description. We also differ from Newell in that it is not assumed that all layers can be completely described with a single physical measure (such as seconds).

2 Informal Description

A large computer-based system exhibits dynamic behaviour on many different levels. The computational components have circuits that have nanosecond speeds, faster electronic subcomponents and slower functional units. Communication on a fast bus is at the microsecond level but may be tens of milliseconds on slow or wide-area media. Human time scales as described above move from the 1ms neuron firing time to simple cognitive actions that range from 100ms to 10 seconds or more. Higher rational actions take minutes and even hours. Indeed it takes on the order of 1000 hours to become an expert at a skilled task, such as flying a plane [18] and the development of highly skilful behaviour may take many years. At the organisational and social level, time scales range from a few minutes, through days, months and even years. Perhaps for some environmentally sensitive systems, consequences of failure may endure for centuries. To move from nanoseconds to centuries requires a framework with considerable descriptive and analytical power.

Most formulations that attempt to identify time granularity do so by mapping all activities to the finest granularity in the system. This results in cumbersome formulae, and fails to recognise the distinct role time is taking in the structuring of the system. An exception is the work of Corsetti *et al*[6, 4]; they identify "*a finite set of disjoint and differently grained temporal domains*". Their framework is not as extensive as the one developed here, but they do show how the notion of temporal domains can be embedded into a logical specification language.

2.1 Definition of a Band

A band is represented by a granularity (expressed as a unit of time that has meaning within the band) and a precision that is a measure of the accuracy of the time frame defined by the band. System activities are placed in some band B if they engage in significant events at the time scale represented by B. They have dynamics that give rise to changes that are observable or meaningful in band B's granularity. So, for example, at the 10 millisecond band, neural circuits are firing, significant computational functions are completing and an amount of data communication will occur. At the five minute band, work shifts are changing, meetings are starting, etc. For any system there will be a highest and lowest band that gives a temporal system boundary — although there will always be the potential for larger and smaller bands. Note that at higher bands the physical system boundary may well be extended to include wider (and slower) entities such as legislative constraints or supply chain changes.

Time has both discrete and continuous characteristics within the framework. Both are needed to capture the essential properties of complex systems; the term *hybrid system* is often used to indicate this dual need. A time band defines a temporal frame of reference (e.g., a clock that *ticks* at the granularity of the band) into which discrete actions can easily be placed. But continuous entities would also be placed in this band if they exhibit significant observable events on this time scale. For these entities time is continuous but significant events occur at a frequency of no more than (but close to) once per 'tick' of the band's abstract clock.

By definition, all activities within band B have similar dynamics and it may be easy to identify components with input/output interactions or precedence relationships. Within a band, *activities* have duration whilst *events* are instantaneous — “take no time in the band of interest”. Many activities will have a repetitive cyclic behaviour with either a fixed periodicity or a varying pace. Other activities will be event-triggered. Activities are performed by agents (human or technical). In some bands all agents will be artificial, at others all human, and at others both will be evident. The relationship between the human agent and the time band will obviously depend on the band and will bring in studies from areas such as the psychology of time [7, 8, 17] and the sociology of time [14].

In the specification of a system, an event may cause a response ‘immediately’ — meaning that at this band the response is within the granularity of the band. This helps eliminate the problem of over specifying requirements that is known to lead to implementation difficulties [10]. For example, the requirement ‘when the fridge door opens the light must come on immediately’ apparently give no scope for an implementation to incorporate the necessary delays of switches, circuitry and the light’s own latency. By making the term ‘immediate’ band specific, it enables a finer granularity band to include the necessary delays, latencies and processing time that are needed to support the immediate behaviour at the higher band.

Events that are instantaneous at band B map to activities that have duration at some lower band with a finer granularity — we will denote this lower band as C. A key property of a band is the precision it defines for its time scale. This allows two events to be simultaneous (“at the same time”) in band B even if they are separated in time in band C. This definition of precision enables the framework to be used effectively for requirements specification. A temporal requirement such as a deadline is band-specific; similarly the definition of a timing failure. For example, being one second late may be a crucial failure in a computing device, whereas on a human scale being one second late for a meeting is meaningless. The duration of an activity is also ‘imprecise’ (within the band). Stating that a job will take three months is assumed to mean plus or minus a couple of days. Of course the precision of band B can only be explored in a lower band.

From a focus on band B two adjacent bands are identified. The slower (broader) band (A) can be taken to be unchanging (constant) for most issues of concern to B (or at least any activity in band A will only exhibit a single state change during any activity within band B). At the other extreme, behaviours in (the finer) band C are assumed to be instantaneous. The actual differences in granularity between A, B and C are not precisely defined (and indeed may depend on the bands themselves) but will typically be in the range 1/10th to 1/100th. When bands map on to hierarchies (structural or control) then activities in band A can be seen to constrain the dynamics of band B, whereas those at C enable B to proceed in a timely fashion. The ability to relate behaviour at different time bands is one of the main properties of the framework.

As well as the system itself manifesting behaviour at many different time bands, the environment will exhibit dynamic behaviour at many different granularities. The bands are therefore linked to the environment at the level determined by these dynamics. In many system abstractions it is useful to assume the environment is in some form of steady state. But this assumption is clearly false as environments evolve, perhaps as a result of the deployment of the embedded system under development. By mapping the rate of this evolutionary change to an appropriate (relatively slow) time band one can gain the advantage of the steady state abstraction whilst not ignoring slower dynamics.

2.2 Behaviour Within a Band

Most of the detailed behaviour of the system will be specified or described within bands. Issues of concurrency, resource usage, scheduling and planning, response time (duration) prediction, temporal validity of data, control and knowledge validity (agreement) may be relevant at any band. Indeed the transfer of techniques from one band to another is one of the motivations for the framework. However, the focus of this paper is on the bands themselves and the relationships between bands, and hence we will not consider in detail these important issues.

We do note however that with human agents (and potentially with artificial learning agents) time itself within a band will play a central role. Time is not just a parameter of a band but a resource to be used/abused within the band. Users will interpret system behaviour from temporal triggers. In particular the duration of an activity will be a source of knowledge and possibly misconceptions; and may be used to give validity (or not) to information, or to infer failure. This use of temporal information to infer knowledge is termed *temporal affordance* [5]. For some bands, agreement (distributed consensus) may depend heavily on such affordances. Plans, schedules or even just routines may give rise to these affordances. Affordances provide robustness; they may be defined into the system but are often developed informally over time by the users of the system. They may be extremely subtle and difficult to identify. Nevertheless the movement of an activity from one band to another (usually a quicker one) may undermine existing affordances and be a source of significant decreased dependability.

Linked to the notion of affordances is that of *context*. A ten minute delay may be a crisis in one context or an opportunity within another. Context will be an issue in all bands but will place a particularly crucial role at the human-centered levels. Context will also play a role in scheduling and planning.

Within a band, a coherent set of activities and events will be observed or planned, usually with insufficient agents and other resources. Robustness and other forms of fault tolerance will also play a crucial role in the description/specification of the behaviour within a band. The specification of some behaviours will require a functional view of time that places ‘time’ at the centre of the design process. To support this process a range of visualisation, modelling and analysis techniques are available including, timed sequence charts, control theory, scheduling analysis, constraint satisfaction, queueing theory, simulation, temporal and real-time logics, timed automata, timed Petri nets, hybrid automata, model checking and FMEA (failure modes and effects analysis).

In all bands, a common set of temporal phenomena and patterns of behaviour are likely to be exhibited by the system itself or its environment. For example, periodic (or regular or cyclic) activities, event handling (responding to an event by a deadline), temporal reasoning (planning and scheduling), interleaving and multi-tasking (and other aspects of concurrency), pausing (or delaying), analysis of response (or completion) time, deadline driven activities, and various aspect of dynamic behaviour such as rates of change. Whilst evident in all bands, these phenomena are not identified using the same terminology in the various time bands of interest (i.e., in the technical, psychological and sociological literature). The development of an agreed collection of guide words within the framework would therefore help link temporal issues with other significant phenomena within a specific band (e.g., terms such as temporal memory, event perception etc., within a ‘psychological’ band).

We also note that the vocabulary usually associated with temporal issues (e.g., late, too soon, on time, simultaneous, instantaneous, immediate, before, never, having enough time, running out of time, plenty of time, etc.) can be given quite specific meanings if they are made band specific. For example, in a human-centred band an electronic spreadsheet responds immediately. Of course at a much lower level band considerable activities are needed to furnish this behaviour. Making the vocabulary of requirements in systems explicitly band-specific will remove some of the misconceptions found with regard to timing issues in such documents. Such temporal keywords are already used as prompts to help elicit time-dependent failures in risk analysis techniques such as Hazard and Operability Studies (Hazops) [12]. Indeed Hazops could be made more effective by making it band-specific.

Finally, we emphasize that the framework is not reductionist. Lower bands contain more detail about individual events. Higher bands contain information about the relationships between activities in a more accessible form. Emergent properties will be observed within a band. The motivation for the framework is to be able to describe these properties, and where necessary link them to more primitive actions at a lower band.

2.3 Behaviour Between Bands

To check the coherence of a description, or the consistence of a specification, for a complex socio-technical system, requires behaviours between bands to be examined. This involves two issues:

1. the relationship between the bands themselves, and
2. the mapping of activities and events between bands.

The link between any two bands is expressed in terms of each band's granularity and precision. Usually the finer of the two bands can be used to express these two measures for the broader band. Where physical time units are used for both bands these relations are straightforward. For example a band that is defined to have a granularity of an hour with a precision of 5 minutes is easily linked to a band with a granularity of 10 seconds and precision of half a second. The granularity relation is a link from one time unit (1 hour) in the higher band to 360 units in the lower band. The precision of 5 minutes means that a time reference at the higher band (e.g., 3 o'clock) will map down to the lower band to imply a time reference (interval) between 2.55 and 3.05.

Granularity can however give rise to a more complex link. In particular, the duration of activities in the lower band may not be the same for all corresponding activities in the higher one. For example, a band with a granularity of 'one month' when linked to a band with a granularity of 'one day' can give rise to a granularity of 28, 29, 30 or 31 days. Here precision is exact, both bands may have the same notion of accuracy about the time reference.

The mapping of actions between high and low bands is restricted to: event to event, or, event to activity relations. So an event in some band can be identified as being coupled to (implemented by) an event or activity in some lower band. A specific named activity exists in one, and only one, band. But for all activities there are events within the same band that are defined to denote the start and end of an activity — these events can be mapped to finer bands. Moreover the whole activity may be seen as an event in a broader band. Figure 1 illustrates three bands (A, B and C) with an event E in band A being mapped to activity X in band B. The start and end events of this activity can then associated with activities in band C.

To exercise these concepts, consider the planning of a university curriculum. When planning courses on a term-by-term basis, a lecture is an event. When planning room

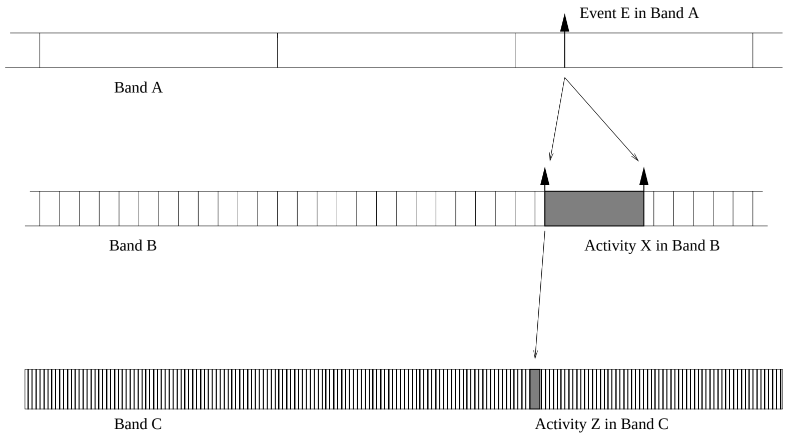

Figure 1: Time Band Example

allocations, a lecture becomes an activity of duration one or two hours (with precision 5 minutes). When planning a lecture, each slide is an event (with an implicit order). When giving a lecture each slide is an activity with duration. This description could be given in terms of a number of bands and mappings of events to activities in finer bands. Note when focusing on the band in which slides have duration it is not possible or appropriate to consider the activities in higher bands that represent whole courses or semesters. The time bands therefore correctly help separate concerns. Students may learn that the time spent on a slide implies importance (at least in terms of the likelihood of the topic turning up in an exam). This is an example of a temporal affordance. Also illustrated by this situation is the difference between planned behaviour (as one moves down the time bands) and emergent properties that enable students to structure the knowledge and understanding they have obtained in many different ways during their progression through their degree course.

To return to the crucial issue of coherence and consistency between bands, the proposed framework facilitates this by making explicit the vertical temporal relationships between bands. Specifically, it becomes possible to check that the temporal mapping between event E in band A with activity X in band B is consistent with the bounds on the relationship identified between bands A and B. Moreover this consistency check can be extended to ordered events and causality (see next section). So, to give a simple example; a lecture of 11 slides each with duration 5 minutes (precision ± 1 minute) cannot be guaranteed to implement the lecture event (as this was mapped to an activity with duration one hour and precision ± 5 minutes).

2.4 Precedence Relations, Temporal Order and Causality

For the time bands associated with computational activity there is usually a strong notion of time and (adequately accurate) physical clocks that will aid scheduling and coordination. This is also increasingly the case with the bands of human experience as external sources of time and temporal triggers abound. But there are contexts in which *order* is a more natural way of describing behaviour [2, 9] (X was before Y, e.g., “before the end of the shift”, “after the plane took off”, “before the flood”, “after the thread has completed”, “before the gate has fired”). The framework must therefore represent both precedence relations and temporal frames of reference. A frame of reference defines an abstract clock that counts *ticks* of the band's granularity and can be used to give a time stamp to events and activities. A band may have more than one such abstract clock but they progress at the same rate. For example the day band will have a different clock in each distinct geographical time zone.

There is of course a strong link between temporal order (i.e., time stamped events and activities) and precedence relations. However, in this framework, we do not impose an equivalence between time and precedence. Due to issues of precision, time cannot be used to infer precedence unless the time interval between two events is sufficiently large in the band of interest.

We develop a consistent model of time by representing certain moments in the dynamics of a band as "clock tick" events, which are modelled just like any other event. When necessary, an event can be situated in absolute time (within the context of a defined band and clock) by stating a precedence relationship between the event and one or more clock ticks.

Precedence gives rise to potential causality. If P is before Q then information could flow between them, indeed P may be the cause of Q. In the use of the framework for specification we will need to use the stronger notion of precedence to imply causality. For example, "when the fridge door opens the light must come on". As noted earlier within the band of human experience this can be taken to be 'immediate' and modelled as an event. At a lower band a number of electromechanical activities will be needed to be described that will sense when the door is open and enable power to flow to the light. Importantly, no causality relationship can be inferred (without explicit precedence) for two events occurring at the same time within their particular band. In effect they are logically concurrent and may occur in sequence or overlapped in time when mapped to a lower band.

Where bands are, at least partially, ordered by granularity, then order and hence potential causality is preserved as one moves from the finer to the coarser bands. However, as noted above, order and causality are not necessarily maintained as one moves down through the bands. This is a key property of the framework, and implies that where order is important then proof must be obtained by examining the inter-band relationship (as discussed above).

2.5 Summary

Rather than have a single notion of time, the proposed framework allows a number of distinct time bands to be used in the specification or behavioural description of a system. System activities are always relative to (defined within) a band. A (non-event) activity has duration of one or more ticks of the band's granularity. Events in a band take no time in that band, but will have a correspondence with activities within a lower band. It follows that a number of events can take place "at the same time" within the context of a specified band. Similarly responses can be "immediate" within a band.

Precedence relations between activities and events are an important part of the framework and allow causal relations to be defined without recourse to explicit references to time. Moreover they can be used to define clock tick events within a band, and hence link other events to the absolute time of the band.

We require all time bands to be related but do not require a strict perfect mapping. Each band, other than the lowest, will have a precision that defines (in a lower band) the tolerance of the band. However within these constraints we do need to be able to show that system descriptions at different bands are consistent. For this to be possible a formal description is required.

3 Time Band Model

In this section we provide a more precise definition of some of the concepts introduced above. This model forms the basis of later work to define a complete logic for the time bands which can then lead to the production of tools to support the use of banded time. There are 6 central notions in the model:

- Bands
- Activities
- Events
- Precedence Relations
- Clocks
- Mappings
- Behaviours

Each of these will be discussed in turn, but note that other entities would be required if the model were to be expanded into a complete system modelling framework, for example: Resources, Agents, and State Predicates.

Bands. A band is defined by its granularity. This establishes a unit of time for the band. Bands are related to one another by the relationship between their granularities; this relates the 'unit' in one (the higher) band to an integer number of 'units' in the lower band. A system is assumed to consist of a partially ordered finite set of bands.

Activities. An activity is an item of work undertaken by some agent. All state changes and effects on the system environment occur within activities. Each activity is bound to one band and has duration in that band.

Events. An event is an activity with zero duration. The start and end of any activity is denoted by an event.

Precedence Relations. Two events from the same band have a precedence relation if one is defined to occur before the other.

Clocks. A band may have one or more abstract clocks that define temporal frames of reference within the band. Each such clock counts in *ticks* and measures the passing of time in the units of time of the band.

Mappings. A mapping is the means of relating behaviours in one band to those in another. Specifically a mapping associates an event in one band to an activity in a lower band. The mapping of a clock tick's start event in one band to an activity with duration in another band leads to the definition of the clock's *precision*. It is precisely the duration of the associated activity (hence precision is a property of the relationship between two bands).

Behaviours. A behaviour is a set of activities and events (within the same band), partially ordered by precedence, giving rise to composition of behaviours.

3.1 Formalising the model

While the time band model captures the essential properties of the time band framework, we need a representation which is more amenable to mechanical manipulation for a formal analysis. The definition is expressed in the Z notation [20] and takes a number of concepts from RTL [11].

Timebands and activities. Let $\mathcal{A}$ be the set of all possible activity instances and $\mathcal{B}$ be the set of time band identifiers. Each activity is associated with a unique time band:

$$band : \mathcal{A} \to \mathcal{B}$$

and each band has associated with it a nonempty set of activities:

$$activities : \mathcal{B} \to \mathbb{P}_1 \mathcal{A}$$
$$\forall b : \mathcal{B} \bullet activities(b) = \{A : \mathcal{A} \mid band(A) = b\}$$

Durations and events. An activity, A, has a duration, $\#A$. The duration of an event is zero and a non-event non-zero. We define $\mathcal{E}$ to be the set of all events.

$$\# : \mathcal{A} \to \mathbb{N}$$
$$\mathcal{E} : \mathbb{P}\,\mathcal{A}$$
$$\mathcal{E} = \{E : \mathcal{A} \mid \#E = 0\}$$

For a band, b, $events(b)$ defines the set of all events in that band.

$$events : \mathcal{B} \to \mathbb{P}\,\mathcal{E}$$
$$\forall b : \mathcal{B} \bullet events(b) = \{E : \mathcal{E} \mid band(E) = b\}$$

Precedence. A precedence relation defines an ordering on the events. Only events in the same time band are related by the precedence relation, and the precedence relation is reflexive and transitive (a preorder). We use the operator $\prec$ for strict precedence and $\equiv$ for equivalence.

$$\_ \preceq \_ : \mathcal{E} \leftrightarrow \mathcal{E}$$
$$\_ \prec \_ : \mathcal{E} \leftrightarrow \mathcal{E}$$
$$\_ \equiv \_ : \mathcal{E} \leftrightarrow \mathcal{E}$$
$$(\forall E, F : \mathcal{E} \bullet E \preceq F \Rightarrow band(E) = band(F)) \land$$
$$(\forall E, F, G : \mathcal{E} \bullet E \preceq E \land (E \preceq F \land F \preceq G \Rightarrow E \preceq G)) \land$$
$$(\forall E, F : \mathcal{E} \bullet E \prec F \Leftrightarrow E \preceq F \land \neg (F \preceq E)) \land$$
$$(\forall E, F : \mathcal{E} \bullet E \equiv F \Leftrightarrow E \preceq F \land F \preceq E)$$

Note that we don't insist that all pairs of events are related one way or the other, and if both $E \preceq F$ and $F \preceq E$, we don't insist that $E = F$, but use the equivalence $E \equiv F$.

Start and end events. Any activity, A, has start and end events, $\uparrow A$ and $\downarrow A$, that are events in the same band as A. The start event of an activity precedes its end event. Events are activities which are their own start and end events. An activity C that exactly encompasses an activity A immediately followed by an activity B has a duration which is the sum of their durations.

$$
\begin{array}{l}
\uparrow \_ : \mathcal{A} \rightarrow \mathcal{E} \\
\downarrow \_ : \mathcal{A} \rightarrow \mathcal{E} \\
\hline
(\forall A : \mathcal{A} \bullet band(\uparrow A) = band(A) = band(\downarrow A) \wedge \uparrow A \preceq \downarrow A) \wedge \\
(\forall E : \mathcal{E} \bullet \uparrow E = E = \downarrow E) \wedge \\
(\forall A, B, C : \mathcal{A} \bullet \\
\quad (\uparrow C \equiv \uparrow A \wedge \downarrow A \equiv \uparrow B \wedge \downarrow B \equiv \downarrow C \Rightarrow \#C = \#A + \#B))
\end{array}
$$

An activity, B, occurs *within* an activity, A, if the start of A precedes the start of B and the end of B precedes the end of A. An activity B that is enclosed within an activity A has a duration no longer than that of A.

$$
\begin{array}{l}
\_ \mathsf{within} \_ : \mathcal{A} \leftrightarrow \mathcal{A} \\
\hline
\forall A, B : \mathcal{A} \bullet \\
\quad (B \mathsf{\ within\ } A \Leftrightarrow \uparrow A \preceq \uparrow B \wedge \downarrow B \preceq \downarrow A) \wedge \\
\quad (B \mathsf{\ within\ } A \Rightarrow \#B \leq \#A)
\end{array}
$$

Clocks. A clock may be associated with a time band. A clock is represented by a nonempty, possibly infinite sequence of *step* activities, each of which is of duration one within that time band. The end event of one step is the start event of the next.

$$
\begin{array}{l}
Clock : \mathbb{P}(\mathrm{seq}_1^{\omega}(\mathcal{A})) \\
\hline
Clock = \{ step : \mathrm{seq}_1^{\omega}(\mathcal{A}) \mid \exists b : \mathcal{B} \bullet step \in \mathrm{seq}_1^{\omega}(activities(b)) \wedge \\
\quad (\forall i : \mathrm{dom}(step) \bullet \#step(i) = 1 \wedge \\
\quad\quad (i + 1 \in \mathrm{dom}(step) \Rightarrow \uparrow step(i+1) = \downarrow step(i))) \wedge \\
\quad (\forall E \in events(b) \bullet \exists i : \mathrm{dom}(step) \bullet \\
\quad\quad \uparrow step(i) \preceq E \wedge E \prec \downarrow step(i)) \\
\}
\end{array}
$$

All events within a time band are related to some clock step's start and end events, and hence all events can be given a time of occurrence according to a clock.

$$
\begin{array}{l}
@ : Clock \rightarrow (\mathcal{E} \rightarrow \mathbb{N}) \\
\hline
\forall step : Clock \bullet \\
\quad @_{step} = \{ E : \mathcal{E};\ n : \mathbb{N} \mid \uparrow step(n) \preceq E \wedge E \prec \downarrow step(n) \bullet E \mapsto n \}
\end{array}
$$

Note that the relationship to $\uparrow step(n)$ is $\preceq$ while that to $\downarrow step(n)$ is $\prec$.

We assume that there is a primary clock for each time band.

$$
\begin{array}{l}
clock : \mathcal{B} \rightarrow Clock \\
\hline
\forall b : \mathcal{B} \bullet clock(b) \in \mathrm{seq}_1^{\omega}(activities(b))
\end{array}
$$

We overload @ so that if a clock is not specified the primary (default) clock of the band is used.

$$
\begin{array}{l}
@ : \mathcal{E} \rightarrow \mathbb{N} \\
\hline
\forall E : \mathcal{E} \bullet @E = @_{clock(band(E))} E
\end{array}
$$

3.2 Mapping between time bands

We assume there is an ordering on time bands (a partial order).

$$\begin{array}{|l}
_ \sqsubseteq _ : \mathcal{B} \leftrightarrow \mathcal{B} \\
\hline
\forall b1, b2, b3 : \mathcal{B} \bullet \\
\quad b1 \sqsubseteq b1 \land \\
\quad (b1 \sqsubseteq b2 \land b2 \sqsubseteq b3 \Rightarrow b1 \sqsubseteq b3) \land \\
\quad (b1 \sqsubseteq b2 \land b2 \sqsubseteq b1 \Rightarrow b1 = b2)
\end{array}$$

A mapping from a time band to a lower time band maps (not necessarily all) events in the upper band to activities in the lower band. The mapping preserves the precedence relation between two high-level events E and F by requiring that the corresponding activity for F cannot finish before the low-level activity for E has started.

$$\begin{array}{|l}
Mapping : \mathbb{P}(\mathcal{E} \rightarrowtail\!\!\!\!\!\rightarrow \mathcal{A}) \\
\hline
Mapping = \{m : \mathcal{E} \rightarrowtail\!\!\!\!\!\rightarrow \mathcal{A} \mid m \neq \{\} \land \\
\quad \exists hi, lo : \mathcal{B} \bullet lo \sqsubseteq hi \land m \in events(hi) \rightarrowtail\!\!\!\!\!\rightarrow activities(lo) \land \\
\quad (\forall E, F : \mathrm{dom}(m) \bullet E \preceq F \Rightarrow \uparrow m(E) \preceq \downarrow m(F))\}
\end{array}$$

The above allows a mapping from a band to itself. This may be useful for something like mapping from one timezone to another.

Events in one time band may map to activities in another lower time band (but not the other way around). The duration of the activity that an event is mapped to gives the *precision* of the event with respect to the lower band.

Any mapping has unique *from* and *to* bands.

$$\begin{array}{|l}
fromband, toband : Mapping \rightarrow \mathcal{B} \\
\hline
\forall m : Mapping;\ b : \mathcal{B} \bullet \\
\quad (fromband(m) = b \Leftrightarrow \mathrm{dom}(m) \subseteq events(b)) \land \\
\quad (toband(m) = b \Leftrightarrow \mathrm{ran}(m) \subseteq activities(b))
\end{array}$$

Granularity. Given two time bands, *hi* and *lo*, such that $lo \sqsubseteq hi$, the granularity of the higher time band with respect to the lower time band is a set of durations (in the time units of the lower time band) that correspond to activities of duration one in the higher time band. Section 2.3 introduced the example of a *MonthBand* and a *DayBand* for which the granularity can be represented as follows:

$$granularity(MonthBand, DayBand) = \{28, 29, 30, 31\}.$$

$$\begin{array}{|l}
granularity : \mathcal{B} \times \mathcal{B} \nrightarrow \mathbb{P}_1 \mathbb{N}_1 \\
\hline
\mathrm{dom}(granularity) = \{hi, lo : \mathcal{B} \mid lo \sqsubseteq hi\}
\end{array}$$

A mapping, *m*, is consistent with the granularity between the bands it maps, if all activities of unit duration map their start and end events to activities in the lower band that are separated by an element of the granularity.

$$
\begin{array}{l}
ConsistentMap : \mathbb{P}\, Mapping \\
\hline
ConsistentMap = \{m : Mapping \mid \\
\quad \forall A : \mathcal{A} \bullet \# A = 1 \wedge \{\uparrow A, \downarrow A\} \subseteq \mathrm{dom}(m) \Rightarrow \\
\quad\quad \exists\, G : granularity(fromband(m), toband(m));\ s, f : \mathcal{A} \bullet \\
\quad\quad\quad s = m(\uparrow A) \wedge f = m(\downarrow A) \wedge \\
\quad\quad\quad @\!\uparrow f - @\!\downarrow s \leq G \leq @\!\downarrow f - @\!\uparrow s\}
\end{array}
$$

For any unit duration activity, A, in the higher band, whose start and end events are mapped by m to activities s and f, respectively, there must exist an element of the granularity relation between the bands that is between the maximum time interval for A (in the lower band) and the minimum time interval for A.

3.3 Extensions to the basic underlying Model

Behaviours. A behaviour consists of a nonempty set of activities within the same band. The start and end events of any activities are also in the behaviour.

$$
\begin{array}{l}
\underline{\ Behaviour\ } \\
band : \mathcal{B} \\
act : \mathbb{P}_1\, \mathcal{A} \\
ev : \mathbb{P}_1\, \mathcal{E} \\
\hline
act \subseteq activities(band) \wedge \\
ev = act \cap \mathcal{E} \wedge \\
(\forall A : act \bullet \uparrow A \in ev \wedge \downarrow A \in ev)
\end{array}
$$

Given a mapping between bands, one may map a behaviour.

$$
\begin{array}{l}
map : Mapping \rightarrow (Behaviour \leftrightarrow Behaviour) \\
\hline
\forall\, m : Mapping;\ hi, lo : Behaviour \bullet \\
\quad ((hi \mapsto lo) \in map(m) \Leftrightarrow \\
\quad\quad hi.band = fromband(m) \wedge lo.band = toband(m) \wedge \\
\quad\quad m(\!|\ hi.ev\ |\!) \subseteq lo.act)
\end{array}
$$

Combining behaviours. Two behaviours within a band may be combined to give a composite of their combined activities and events.

$$
\begin{array}{l}
\_ \cup \_ : Behaviour \times Behaviour \rightarrow Behaviour \\
\hline
((h1, h2) \in \mathrm{dom}(\_ \cup \_) \Leftrightarrow h1.band = h2.band) \wedge \\
(\forall\, h1, h2, h : Behaviour \bullet h = h1 \cup h2 \Leftrightarrow \\
\quad h.act = h1.act \cup h2.act)
\end{array}
$$

The behaviours may be overlapping in time (parallel composition) or the last event of one behaviour may precede the first event of the other (sequential composition).

4 Use of the Model and Framework

Using the temporal framework described in this report to give structure to a system has the immediate advantage that the dynamic aspects of the system's behaviour are partitioned into bands that exhibit similar dynamics. This directly supports the separation of concerns that is at the heart of good system structuring. Within a band, actions with compatible timing properties can be modelled together with the necessary attention being given to issues of precedence, causality, temporal affordance, resource usage and timely progress. Between bands the use of a formal model will allow consistency to be asserted between different (temporal) descriptions of the system. As well as these general usages of the framework, other specific issues can be addressed. In particular:

- the consequences of failure,
- the impact of change, and
- the analysis of responsiveness.

In the first of these the consequences of late (or early) events in one band can be evaluated in terms of the impact on activities in the next higher band. Other structural means can then be employed to contain the consequences of errors flowing up through the bands. Similarly the result of changes in higher bands on the required performance of activities in lower bands can be evaluated within the framework. This can translate, for example, onto workload issues for human operators. In general, the analysis of responsiveness will make it possible to determine whether the system will be able to respond in time for its outputs to be useful. It should also allow the designer to observe at which band problems are occurring, which may lead to redesign of that band of the system.

In human cognition the time bands in which the brain functions are fixed (although on-going research may change our understanding of the role of observed activities). Moore's law indicates that the technical components of systems are unlikely to stay in the same time bands during system upgrades. With technical systems there is also often a trade-off between time (the speed of an activity) and other non-functional attributes such as power consumption, heat production or space (silicon layout). During design, various system behaviours can be evaluated by moving agents between bands. In some dynamic systems such movements may even be made during operation; for example to lower power consumption during a 'quiet' period.

In the process of upgrading a system, or automating parts of an existing manual system, significant changes to the temporal behaviour are likely. These may lead to unanticipated negative consequences such as the undermining of developed affordances or in the breaking of an implicit precedence relationship. The time band framework will enable many of these consequences to be investigated during modelling rather than deployment.

5 Conclusion

In this paper we have argued that complex systems exhibit behaviour at many different time levels and that a useful aid in describing and specifying such behaviour is to use time bands. Viewing a system as a collection of activities within a finite set of bands is an effective means of separating concerns and identifying inconsistencies between different 'layers' of the system. Time bands are not mapped on to a single notion of physical time. Within a system there will always be a relation between bands but the bands need not be tightly synchronised. There is always some level of imprecision between any two adjacent bands. Indeed the imprecision may be large in social systems and be a source of dependability (robustness).

References

[1] A.F.Monk. Lightweight techniques to encourage innovative user interface design. In L. Wood, editor, *User interface design: bridging the gap between user requirements and design*, pages 109–129. CRC Press, 1998.

[2] J. Allen. Towards a general theory of actions and time. *Artificial Intelligence*, 23:123–154, 1984.

[3] G.D Baxter, A.F. Monk, K. Tan, P.R.F. Dear, and S.J. Newell. Using cognitive task analysis to facilitate the integration of decision support systems into the neonatal intensive care unit. In *Artificial Intelligence in Medicine*, 2005.

[4] E. Ciapessoni, E. Corsetti, A. Montanari, and P. San Pietro. Embedding time granularity in a logical specification language for synchronous real-time systems. *Science of Computer Programming*, 20:141–171, 1993.

[5] A.P. Conn. Time affordances: the time factor in diagnostic usability heuristics. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 186–193, 1995.

[6] E. Corsetti, A. Montanari, and E. Ratto. Dealing with different time granularities in formal specifications of real-time systems. *Journal of Real-Time Systems*, 3(2), 1991.

[7] P. Fraisse. *The psychology of time*. New York: Harper and Row, 1963.

[8] W. Friedman. *About time: Inventing the fourth dimension*. Cambridge, MA: MIT Press, 1990.

[9] E. Hollnagel. *Human Reliability Analysis: Context and Control*. Academic Press, 1993.

[10] S.G. Hutchesson and N. Hayes. Technology transfer and certification issues in safety critical real-time systems. In *Digest of the IEE Colloquium on Real-Time Systems*, volume 98/306, April 1998.

[11] F. Jahanian and A.K. Mok. Safety analysis of timing properties in real-time systems. *Transactions on Software Engineering*, SE-12(9), 1986.

[12] T. A. Ketz. *Hazop and Hazan: Identifying and Assessing Process Industry Hazards*. Institution of Chemical Engineers, 1999.

[13] G. Klein, R. Calderwood, and D. MacGregor. Critical decision method for eliciting knowledge. *IEEE Transactions on System, Man, and Cybernetics*, 19:462–472, 1989.

[14] R. Levine. *A geography of time*. New York: Guilford Press, 1997.

[15] S. Miksch, A. Seyfang, W. Horn, and C. Popow. Abstracting steady qualitative descriptions over time from noisy, high-frequency data. In *Artificial Intelligence in Medicine. Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making, AIMDM'99, Aalborg, Denmark*, pages 281–290. Springer: Lecture Notes in Computer Science (1620), 1999.

[16] A. Newell. *Unified theories of cognition*. Harvard University Press, Cambridge, MA, 1990.

[17] J.E. Roeckelein. *The concept of time in psychology: A resource book and annotated bibliography*. CT: Greenwood Press, 2000.

[18] W. Schneider. Training high-performance skills: Fallacies and guidelines. *Human Factors*, 27(3):285–300, 1985.

[19] J.M. Schraagen, S.F. Chipman, and V.L. Shalin. *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.

[20] J. M. Spivey. *The Z Notation: A Reference Manual*. Prentice Hall International, second edition, 1992.

Appendix — A Case Study

In order to illustrate the descriptive use of the framework, time bands will be employed to help describe the dynamic characteristics of the Neonatal Intensive Care Unit (NICU) at St James' University Hospital in Leeds, UK. This Unit has recently been the subject of an intensive study [3] using a Cognitive Task Analysis (CTA) [19] to collect information about all aspects of task performance. The CTA was comprised of:

- Lightweight rich pictures [1] to describe the physical and social work context and identify the roles and responsibilities of the various system stakeholders.
- The Critical Decision Method [13] to analyse the processes used by staff in deciding on changes that need to be made to the ventilator settings.
- Naturalistic observation of how staff use the ventilator in situ.

Premature babies often suffer from problems that are associated with being underdeveloped at the time of birth. The lungs of babies that are born prior to 28 weeks gestation are often incapable of producing enough of the surfactant that is required to allow gaseous exchange to take place in the lungs. This problem, which is called Respiratory Distress Syndrome (RDS), is a self-regulating disease which peaks about 72 hours after birth and normally disappears within 5 to 7 days. Treatment of RDS usually involves a combination of drugs and the use of mechanical ventilation to control the partial blood gas pressures of the baby.

In general, once the ventilator has been configured, changes to the settings are only made in response to acute situations. Typically, this will be in response to an alarm, or when the staff notice that the baby's condition is deteriorating towards a state where an alarm will be raised by the monitoring equipment. Senior House Officers (SHOs — junior doctors) and nurses provide the first line of care for the babies. Normally the nurses do not change the ventilator settings, except for the inspired oxygen level and possibly the breathing rate. If the SHOs or nurses decide that a particular problem is too difficult for them to deal with they can call for the assistance of a registrar; in more complex cases, one of the consultants may be called in.

An analysis of the timing issues and system dynamics within the NICU leads to the identification of a number of distinct time bands. Not only are these bands situated at different granularities, they also use 'time' in quite diverse ways. Within a system boundary that excludes the internal operations of the computer equipment which is in use within the NICU, we identify four time bands. In the following brief descriptions we note; the granularity of each band, key/typical activities and events, and the implied precision of each band. Note all these observations are approximate as the original Cognitive Task Analysis did not frame its questions in the context of time bands. The use of the framework to specify a system would give greater attention to these notions.

P – Future Planning Granularity: a week. Activities: planning the introduction of new procedures (perhaps in response to changes required by the regulatory authority), clinical trials of new equipment or new drugs. These could last weeks or months. Events: setting up a trial, evaluating a trial, ordering equipment. Precision: a day.

W – Ward Organisation Granularity: half hour. Activities: shifts (typically 8 or 12 hours), daily ward round (2 hours), time for X-ray to be available (30 minutes), setting up ventilator (30 minutes), stay of baby on ward (5 days to 12 or more weeks). Events: observations every hour. Precision: 10 minutes.

C – Clinical Procedures Granularity: five minutes. Activities: calling in the Registrar or Consultant (5 minutes), medical interventions of various forms (5 or more minutes), response of baby to change of ventilator setting (20 to 30 minutes). Clinical aspects of admitting a baby (5 to 10 minutes). Events: responding to an alarm, observing movements of a baby, putting baby on ventilator, take a blood sample. Precision: one minute.

B – Baby Dynamics Granularity: one second. Activity: breathing cycle, regular heart beat, sampling of ventilator and Neotrend [1] (every second), response to alarm (30 seconds to 1 minute), response of blood gas levels (several seconds). Events: heart beat, a single sample. Precision: ten milliseconds.

As indicated earlier, other time bands could be included if the system boundary is extended. For example, modeling of the baby's internal breathing cycle, or the computer system scheduling or micro code execution, or the nurses' cognitive behaviour would all need much finer granularities. However, for the study of the NICU, such descriptions are unnecessary and instantaneous events (e.g., taking a blood sample) within the finest band of interest will suffice. Also above the higher band of the system described here are significant temporal issues. If the baby's brain is not supplied with enough oxygen, there can be brain damage. The full effects of this may not be known until appropriate tests can be performed, and this is generally 2 or preferably 3 years after being born. Some of the more subtle lesions may not show up until as late as 7 years old. If the baby is supplied with too much oxygen, it can affect their sight. This is usually checked at 6 weeks. Again we choose not to include these bands in our description of the NICU.

The four bands, which span granularities, from one second to a week, use 'time' in very different ways. Band **P** is mainly concerned with durations (e.g., a trial will take two months), whereas **W** uses time as defined by standard clocks to coordinate and help manage the ward's operations. For example, an observation round takes place every hour on the hour; this exact timing is not necessary but is a useful convenience.

In band **C** precedence relations are more important; nurses follow procedures. As long as there is sufficient time to complete these procedures, time does not play any explicit role in their actions. And there is sufficient time if the staffing levels and skill/experience are appropriate — this issue is normally addressed by various forms of work flow analysis.

Band **C** also exhibits a number of delays that are significant (e.g., 10 minutes to call a Registrar, and 20 to 30 minutes for a baby to respond to treatment). In addition to the humans (nurses, doctors etc., and the babies), the technical system (the Neotrend) also inhabits this band. This band is therefore the most important one on which to focus if the dependability of the NICU is to be assessed. The final band, **B**, has its time granularity set by the dynamics of the baby (heart rate, breathing rate). Within the system the baby represents the controlled object and there will always be a time band within any system that matches the external dynamics of any such controlled object. By external we imply the useful measurable variable that is sensed by the system. The granularity of **B** was given as a second as this is a reasonable approximation to a baby's key rates (breathing and heart).

[1] The Neotrend is an indwelling arterial sensor that is used to continuously monitor the pH and partial pressures of oxygen and carbon dioxide in the baby's blood.

Between the four bands there are a number of potential consistency issues. These usually arise when an event in one band maps to an activity in another. For example, 'setting up a trial' in **P** will lead to activities within **W** (and possibly **C** that will need planning as there may not be sufficient resources available, e.g., nurses). Similarly a new mode of operation of the ventilator may require closer monitoring of the baby by a senior nurse; this will again have an impact on resources at band **C**. Another example is the event 'responding to an alarm' in **C**; here the dynamics and safety of the baby will determine the duration of the corresponding activity in **B**. A more detailed incident that identifies a number of inter- and intra- band relationships is given below.

This case study also illustrates a typical relationship between bands. Band **P** lays out the strategy for what should happen at band **W**; band **W** lays out the strategy for band **C**; and band **C** lays out the strategy for band **B**. Then, going in reverse, band **B** determines the tactics at band **C**; band **C** determines the tactics at band **W** and so on. Strategy is basically concerned with planning, whereas tactics is what happens in response to things that change locally within that time band.

A further example from the Case Study

This example is based on an incident described by one of the experts interviewed during the case study. The incident was described from the point at which a registrar called in one of the consultants by telephone (band **C**). The expert attempted a diagnosis over the telephone, and when that appeared not to work, went in to the NICU. When he arrived, things were fairly busy around the baby's cot, and a few alarms were sounding.

An X-ray was taken to check the position of the Neotrend indwelling device that is used to continuously monitor the baby's blood gases (band **W**). Once the results of the X-ray confirmed that the Neotrend was correctly positioned, it was decided to change the ventilator settings (band **C**). The alarms continued to sound, and the trend displays were not showing the desired response to the changes, so further changes to the ventilator settings were made to increase the baby's oxygenation. A nurse spotted that the ventilator alarm was sounding, indicating a leak in the ventilator tubing circuit. No leak was found, so it was decided to replace the endo-tracheal tube (ETT) with one of a larger diameter. After the change had been made, the trend displays indicated that the blood gas levels were starting to move in the right direction.. The consultant decided to take a step back and went off to make tea for everyone. This was to allow the changes to take effect, and to reflect on the case. The ventilator alarm persisted, so the ETT tube was replaced again with one of an even larger diameter. Once the trend displays indicated that the baby's blood gases were moving back to more appropriate levels, the changes that were previously made to the ventilator were undone. This and all the above activities were contained within band **C**.

This example illustrates two important points. The first relates to the taking of an X-ray, which is a shift to the next higher band (from the **C** band up to the **W** band). The human factors literature on controlling complex systems often refers to the need for the operators (doctors, pilots, process operators and so on) to maintain the big picture of how the system is working. In this example, the taking of the X-ray acts as a validation check on the activities that are being performed in band **C**. If the Neotrend is not correctly positioned, the readings it gives do not accurately reflect the baby's blood gases.

The second point is that there is a natural pace of activities within a band. It can sometimes be tempting to react too quickly as new data keeps coming in, rather than waiting for the appropriate time, particularly when things appear to be going wrong.

This can be particularly important in the NICU, where the data are inherently noisy [15]. The decision by the consultant to go off and make tea allowed time for the changes to start taking effect, and for real trends in the data to start to emerge.