Part IV

Formal Proof of WBA Correctness

(with Karsten Loer)

Chapter 18

The PARDIA Classification

In order formally to demonstate the correctness and relative sufficiency of our WBA, a formal proof needs to be constructed. This means that all components of the explanation must be formally axiomatised. We consider first the human-tasking classification component. We need some scheme for classifying human behavior, because there are certain human processes which play a role in this incident. Our aim in formalising the human tasking is to classify the kinds of errors that can be made in human tasking so that potential corrective measures may be researched. We do not aim to provide a model of human functioning or of human agency. The PARDIA classification scheme leads to the correlation between errors and what we might call 'domains of correction' in Figure 18.3. We are sure that finer classifications are possible, but we have found that PARDIA yields clear insights in the cases in which we have used it.

18.1 Analysis of Pilot Behavior

The PARDIA scheme for classifying human actions is an extended '*information-processing model*', in which for a given system state, a pilot's interaction with the system is considered to form a sequence:

perception-attention-reasoning-decision-intention-action

Perception	$Perceive(crw, feature)$
Attention	$Attend(crw, feature)$
Reasoning	$Reason(crw, conclusion)$
Decision	$Decide(crw, conclusion)$
Intention	$Intend(crw, conclusion)$
Action	$Act(crw, action)$

Figure 18.1: Notation for the PARDIA attitudes

This sequence reads as follows.

- *perception*: An annunciation of the system state is presented to the pilot's perceptual systems;
- *attention*: the pilot notices the annunciation;
- *reasoning*: figures out what are the possible actions to take;
- *decision*: decides on an action;
- *intention*: forms the intention to carry it through;
- *action*: and finally carries it out.

Other 'information-processing models' have fewer states, but we claim that a decomposition of pilot behavior with at least so fine a grain is needed for incident narratives. PARDIA is a classification scheme and we therefore prefer not to call it a 'model'. The granularity of PARDIA is justified through considering that failures can occur and have occurred at any stage in this sequence. Examples are:

perception failure: During the A330 flight test accident in Toulouse in 1994, an annunciation of the autopilot mode change was not displayed to the pilots, because the angle of attack of the aircraft was higher than 25deg. This was cited as a contributing factor in the DGA report.

attention failure: In the incident analysed in [Pal95], the pilot flying (PF) failed to notice that the altitude capture mode was no longer armed.

reasoning failure: In the B757 accident off Puerto Plata, Dominican Republic in 1996, the captain chose to switch on the center autopilot, after concluding that his air data was faulty. The center autopilot obtains its air data from the captain's air data system.

decision failure: During the B757 accident off Lima, Peru in 1996, the pilots had lost all effective air data, presumably related to fact that the left-side static ports were covered with masking tape which had not been removed as the aircraft was returned to service after cleaning. During the incident, the pilot asked for altitude data from Lima Tower, who reported indicating 9,000ft. The PF's AI was apparently reading similarly. He took a calculated risk to begin a descent, and impacted the ocean since his true altitude was a few feet above sea level. His AI read 9,500ft on impact.

intention failure: In GLOC (G-induced loss of consciousness) incidents, pilots who regain consciousness are reportedly unable to form the intention to recover an aircraft obviously heading for ground impact.

action failure: The test pilot of the A330 let the departure from normal flight develop, presumably to obtain test data, and initiated recovery too late to avoid ground impact; in the B757 Puerto Plata accident, the crew were unable to take effective action during stick-shaker warnings, allowed the aircraft to stall, and could not recover the stall.

To some readers, it may seem intuitively odd that the first type of failure is classified as a human-involved failure, because it seems as if the remedy for this type of failure cannot lie with the pilots. Indeed, that is correct - the remedy cannot lie with anything to do with the pilots. Our perspective is as follows. The flight crew is part of a complex, heterogeneous, distributed system. A perception failure is a communication failure between the rest of the system and the human part of the system. This involves the humans - they sit at one end of the communication mechanism - therefore it is a human-related failure. It is the unique part of the PARDIA classification in which no potential remedy can have anything to do with the pilots.

The intent of the PARDIA model is not to explain actions. That is better left to psychologists. It is intended to provide important 'system' states for human actors in a system. A system state is 'important' for failure reasoning if it reflects a match or a mismatch of state knowledge or actions between various system parts – in this case between the human agent and another system part. So pilot beliefs and misperceptions are not explicitly handled because logically they are local state predicates of the human agent alone, not predicates involving the state of other system parts as well. A mistaken perception or belief will affect the reasoning process of a pilot; but because it is mistaken, it has no correlate in the world. To put it intuitively using old metaphors, the only place this mistaken perception or belief is 'located' is the pilot's head. It may be expected to affect the reasoning processes, and therefore will show up eventually as a 'reasoning failure' in the PARDIA classification. Finer classifications may try to analyse reasoning further, to identify for example the perceptual or epistemological situation of the pilots at various stages.

18.2 A Tricky Example

As an illustration, we discuss what is perhaps the subtlest case of PARDIA use in this example, namely to show

$$\vdash (\mathit{Hypotheses} \land \mathit{Procedures}) \;\Box\!\!\Rightarrow \Diamond[11]$$

where

$$\mathit{Hypotheses} \triangleq \begin{array}{l} \land\ [111] \\ \land\ \langle 112\rangle \\ \land\ \langle 113\rangle \end{array}$$

and

$$\textit{Procedures} \triangleq \text{Standard Operating Procedures}$$

and

[11] /* CRW opts to continue landing */
[111] /* CRW realizes they are landing at the wrong airport */
⟨112⟩ /* CRW has safety reasons for continuing landing */
⟨113⟩ /* Standard Operating Procedures */

An intuitive analysis of these circumstances would proceed roughly as follows. Normal procedure (that is, except in case of emergency, or when weather is below minimums) is that one does not land at an airport that is not the airport of landing in the flight plan. Therefore when the flight crew realised that the airport was not the airport of planned landing, they have to evaluate this procedure: the airport is not the airport of planned landing; there is no emergency; therefore the conditions for not landing are fulfilled; but also there is some risk involved in breaking off a landing procedure at a late stage; a decision is therefore necessary either to continue or to abort. Had they not realised that the airport of landing was not the destination airport, there would have been no procedural conflict, for them, needing decisive resolution.

We can describe the procedures and conditions formally, as we do in Section 21 in Module *Landing_Norms* (see *GeneralSafetyRule*, *Landing_Procedures*, and *Normal-Progress*). The concept 'realisation' must be translated into the PARDIA model. What the pilots realise is the truth of an assertion that is not a direct physical perception (e.g., '*there is another aircraft in the vicinity*') but the product of ratiocination. Therefore 'realise' is best translated as *Reason*. In PARDIA notation, we are informed then by the source that the crew realised the airport of landing was not the destination:

$$Reason(CRW, APT \neq destAPT)$$

and also that they have safety reasons for continuing the landing; using logical notation:

$$Reason(CRW, \neg\Diamond endanger(CRW, TFC) \Rightarrow \Box(AC) in\_landing\_phase)$$

From the former, and the fact that the other conditions of *Normal-Progress* are satisfied, we can infer that this procedure entails that they shouldn't land there:

$$O(\neg\Diamond(AC) lands\_at(APT))$$

($O(...)$ stands for '*ought*') and from the second consideration, we can infer that they should continue

$$O(\Box(AC) in\_landing\_phase)$$

and following this obligation will entail (via *Landing_Norms.Landing_Procedures*) that the aircraft should land there. We note finally that in PARDIA, the existence of a procedural obligation entails that everyone has reason to follow it:

$$O(A) \Rightarrow \mathit{Reason}(X, A)$$

So this states the situation more formally. How are we to explain why the crew opted ('*Decide*' in PARDIA) to continue? The obligations entail that the crew has reason to follow them:

$$\mathit{Reason}(\mathit{CRW}, \neg\Diamond(AC)\mathit{lands\_at}(APT))$$

$$\mathit{Reason}(\mathit{CRW}, \Box(AC)\mathit{in\_landing\_phase})$$

and this latter, $\Box(AC)\mathit{in\_landing\_phase}$, given that the aircraft is near BRU, entails via *Landing_Norms.Landing_Procedures* that the aircraft will eventually land there, and therefore the crew has reason for this eventuality (another PARDIA Rule is that having reasons is closed under modus ponens), namely:

$$\mathit{Reason}(\mathit{CRW}, \Diamond(AC)\mathit{lands\_at}(APT))$$

so the crew has reasons both for and against landing. In such a case, one wants to say, a decision is called for. This is directly expressed by the axiom

$$\mathit{Reason}(X, A) \wedge \mathit{Reason}(X, \neg A) \Rightarrow O(\mathit{Decide}(X, A) \vee \mathit{Decide}(X, \neg A))$$

which enables one to conclude, with a bit of logic, that

$$O(\mathit{Decide}(\mathit{CRW}, \Box(AC)\mathit{in\_landing\_phase}) \vee \mathit{Decide}(\mathit{CRW}, \neg\Box(AC)\mathit{in\_landing\_phase}))$$

(this latter will entail a breakoff, of course).

So we have shown that a decision ought to be made. And it was – the crew opted to continue the landing. The final step concerns the reasons for this decision. We take the reasons simply to be that a decision was called for, and indeed it was taken. That is, schematically,

$$O(\mathit{Decide}(X, A) \vee \mathit{Decide}(X, \neg A)) \wedge \mathit{Decide}(X, A) \Rrightarrow \mathit{Decide}(X, A) \qquad (18.1)$$

Although this seems to be a general property, we didn't include it in the actual PARDIA classification for two reasons:

- it appears to be part of a normalised human behavior, and therefore part of a model, rather than a classification scheme;
- it involves the deontic modal operator ('*ought*') essentially.

These are reasons concerning the formalism, though, and not 'actual' reasons based on what is true and false about the world. However, PARDIA is a set of temporal logic axioms defining certain primitives, and as such is relatively modular and thereby flexible, for example those who wish to try to base explanations purely on temporal reasoning can use PARDIA without modification; also, for various reasons, automated reasoning tools are easier to build and use when the logical parts are highly modular. The introduction of the deontic modality could also lead to further questions. For example, belief has also been handled as a modal logic (the doxastic modality), and surely what a crew believes is also important for their behavior? If deontics were to be part of PARDIA, why not then also epistemics? We argue against the attempted inclusion of epistemics in PARDIA in Section 18.4, namely that it's not clean. The overriding reason against these inclusions is again that PARDIA is a classification scheme, not a model; therefore one need not include everything and formal considerations play an enhanced role in deciding what to include.

Thus we choose to use the required instance of Rule 18.1 as an assumption in the correctness proof, rather than to adopt Rule 18.1 itself as a PARDIA axiom.

This illustrates the use of PARDIA in a somewhat tricky case in which somewhat complex properties of human behavior need also to be handled. In particular, we have shown how to use a classification system that is not intended to be a model where modelling appears to be called for. Namely, after a certain amount of analysis, one can include required modes of behavior as individual assumptions, whose justification will then simply be their truth in the individual instance. One thereby avoids the common trap of modelling, in which one generalises phenomena too broadly, when all that is really needed is an instance or two.

18.3 Perception and Attention

Reflecting the view above that mistaken perceptions and beliefs are part of what we call human reasoning processes, we use the predicates $Perceive(CRW, A)$ and $Attend(CRW, A)$ to reflect veridical physical perception and veridical mental attention (attention to a phenomenon means simply that one is somehow consciously aware of it, cognisant of it). Hence the two PARDIA axioms:

$$PARDIA \vdash Perceive(CRW, A) \Rightarrow A \tag{18.2}$$

$$PARDIA \vdash Attend(CRW, A) \Rightarrow A \tag{18.3}$$

Thus *Perceive* is the predicate of *veridical perception* and *Attend* the predicate of *veridical awareness* or attention. It is not intended to be possible in the PARDIA model as we conceive it here to state that a pilot misperceived some situation. We regard such a statement as one about pilot beliefs (construed in the broad sense as those hypotheses from which a pilot conducts reasoning). The justification for this position is that the pilot attitude to these non-facts is the same as to facts: this is often expressed by saying the pilot 'believes' that these non-facts are 'facts'. So we should probably say a couple of words about epistemics, and why we don't include them here.

18.4 Why Epistemics Are Not Included

Unlike [Joh97], we do not attempt to explicate the role of belief in pilot behavior, for broadly two reasons:

- we did not find it necessary to include epistemics in order satisfactorily to explain any of the incidents we considered;
- there is considerable disagreement in the philosophical and logical communities concerning exactly what the properties of belief are.

Let us consider this latter point more deeply.

18.4.1 Some Apparent Paradoxes of Belief

The first puzzle is as follows. Were belief to be closed under Modus Ponens for every actor (that is, if X believes A and X believes $(A \Rightarrow B)$ then X believes B) then it would follow that all actors would believe all the validities of propositional logic; and if an actor believed one propositional logic statement that is in fact not valid, that actor would believe all statements of propositional logic, including $(P \& \neg P)$. If this were really the case, then the only consistent beliefs that anyone could have would constitute precisely all the extensions of the set of tautologies up to but not including the set of all propositional logic formulas; and the only inconsistent beliefs possible would be the set of all propositional formulas. Thus the possible sets of beliefs would be precisely these.

We hope it is as obvious to the reader as it is to the first author, a sometime teacher of propositional logic, that this conclusion is patently false about what people actually believe - it is not hard to construct a tautology which will be judged not so by some average logic students. Therefore belief is not necessarily closed under Modus Ponens. This means that if X believes A and X believes $(A \Rightarrow B)$ then X does not necessarily believe B. However, any X who believed A and $A \Rightarrow B$ but did not believe B would be clearly irrational: after all, Modus Ponens is one of the basic rules of logic! So much for the first puzzle.

A second problem with the properties of belief could be called the Preface Paradox. Suppose we have written a lengthy book about accident analysis. Naturally we have made a serious attempt to ensure that everything we say in the book is correct, is true. However, we are like many authors also socially modest, and in the preface we say that we believe that we will have made at least one mistake and that responsibility for this mistake rests with us. Thus, apparently we believe each individual statement in the book, and we believe that one of these individual statements is in fact false. Suppose that belief is closed under conjunction: that if X believes A and X believes B then X believes $(A\&B)$. Then we believe both that the conjunction of all the statements in the book is true (using closure under conjunction, over all the individual statements), and we believe that this conjunction is false (which follows directly from what we say in the preface). Thus we have inconsistent beliefs: let the conjunction be A; we believe both A and $\neg A$!

So we must conclude either that it is not only possible but sometimes reasonable to hold inconsistent beliefs; or that belief is not closed under conjunction: if X believes A and X believes B, then X does not necessarily believe $(A\&B)$. But such an X would again be pretty clearly irrational. So maybe it is reasonable to hold inconsistent beliefs: but then if belief is closed under Modus Ponens, the only inconsistent belief set would be the set of all formulas, including $(P\&\neg P)$, which would be irrational to hold. We therefore conclude that if belief is closed under conjunction, it cannot be closed under Modus Ponens. But we have pointed out that someone holding a set of beliefs not closed under Modus Ponens appears to be irrational.

We conclude that any possible belief set appears to have features that would render its bearer irrational.

18.4.2 Allowing for Belief

When paradoxes abound, there are only two courses of action for those wishing to develop a traditional formal logic using such an operator: to use weak enough axioms and inference rules that the paradoxes cannot be derived; or to accept and accommodate the paradoxes. The topic of belief can be a minefield for the unwary. We do not know an appropriate way to accomodate the paradoxes in Section 18.4.1: we are thus loathe to build part of a logical failure analysis method on top of these potentially wobbly foundations. We are also loathe to use only principles of belief that are weak enough to avoid derivation of the paradoxes, because this is not a purely theoretical application: the point of using belief in this application is to explain real human behavior in the system, and weak axioms lead to handicapped methods.

We believe, however, that it is important for WBA and PARDIA to allow inclusion of epistemic considerations if desired: we shall remain consistent with (reasonable) views of epistemics. Johnson's observations and suggestions about human beliefs can be placed in an "*epistemics*" module and included with the PARDIA modules if required. Epistemic rules should not just include rules concerning belief alone, but should also include rules for interaction with the PARDIA primitives. For example, two epistemic interaction rules which adherents might like to consider for an epistemics module are

$$Believe(X, Perceive(Y, A)) \Rightarrow O(Believe(X, A)) \tag{18.4}$$

$$Believe(X, Attend(Y, A)) \Rightarrow O(Believe(X, A)) \tag{18.5}$$

because if X believes that Y veridically perceives A, then X believes A is veridical (corresponding to a PARDIA axiom, below); mutatis mutandis for *Attend*.

We note, finally, that this situation with belief is unique amongst the intensional operators we have considered. It does not pertain to the other operators. Although there are well-known deontic 'paradoxes', these are not so much paradoxes as conflicts with certain intuitions about obligations [MWD78]. Giving up those intuitions does not significantly weaken the application of deontics to system analysis, as argued strongly in [MWD78]. Similarly, although the understanding of counterfactuals has induced lively discussion amongst philosophical logicians, the semantics we use is technically clear and the logic has been proven to be sound and complete for the intended semantics: thus there are no technical problems of balancing inconsistency against weakness, as in the case of belief. Finally, although there has been some discussion in the late twentieth century about modal contexts in general and quantifying into them in particular, and about persistence of objects through time and across possible worlds, the modal semantics is technically unimpeachable and the corresponding logics are complete (or relatively complete in the case of TLA). We therefore have no qualms about using them.

18.5 PARDIA Axioms as a Module

The PARDIA axioms are part of the PARDIA model of human agent behavior in a system. WBA can, however, be based on various different models of cognition and action, as for example discussed in [Dav80, Sea83, Mel97, Hor97, Ste97]. Thus we do not at this point wish to include PARDIA axioms as axioms of EL: we include them as a module. This entails that each clarification of human behavior in a WBA prooof will need explicitly to include PARDIA axioms as a hypothesis. The *PARDIA Axioms* Module is shown in Figures 18.4 and 18.5. We do not claim that this is the complete set of PARDIA axioms which we need to use, but merely a sufficient set for current purposes. The reader should expect further axioms to be included as experience builds with the method. In this module, and in the following *PARDIA Norms* module in Figure 18.6, we use the notation for the six PARDIA attitudes given in Figure 18.

The fact that the PARDIA primitives are attitudes may lead some to question whether EL is being stretched to accomodate higher-order logic: the second argument to a PARDIA primitive appears to be a sentence defining an action or a state predicate. However, we don't intend this to be so – this is pure syntax. We suggest using the definitional facilities of TLA to define short names for any complex formula that needs to occur in a PARDIA primitive in a proof, and then using this short name in the primitive. Because of this definitional facility, we need explicitly to include axioms that make the PARDIA attitudes 'referentially transparent': invariant under logical equivalence. Although some might complain that if I intend to do A, I do not necessarily intend to do $(B \wedge A)$, where B is some complicated logical truth that I can barely read, let alone understand. This is not, however, our intended meaning for the PARDIA sort of intention. I intend to do something – to perform an act, to accomplish a state of the world in the near future – and this *something* is what is meant. Means of referring to it in a PARDIA primitive must thus be referentially transparent. The same act is denoted under either description A or $(B \wedge A)$, and it is *this* act, under any description, which I intend to perform; *this* state of the world, under any description, which I intend to achieve. We see no problems in requiring and using the referentially transparent senses of these attitudes. It is more usual in philosophy and logic to call the PARDIA attitudes '*propositional attitudes*', but we avoid this terminology to avoid unintended conflation with other, traditional, non-referentially-transparent, notions.

18.6 PARDIA Norms

Besides the PARDIA axioms, there are also certain principles which are *intended* to hold; *ought* to hold in a 'normal' behavior. For example, a pilot ought to perceive things that the aircraft system designer intended she should perceive; be aware of things she perceived; reason from those facts which she correctly perceives; intend to do those things which she concludes; act to fulfil her intentions. But as we have remarked above, any of these steps may fail, and this failure can be part of an incident history. Hence, unlike the PARDIA axioms (which constitute, if you like, part of the meaning of the PARDIA attitudes), these properties of the PARDIA model are defeasible. We place them separately into a module *PARDIA-Norms* in Figure 18.6.

18.6.1 Intentions and Deontics

If CRW intends to carry out some action A, that does not necessarily mean that this action ought to be carried out. Similarly, if CRW intends to omit some action, it does not necessarily mean this action ought be omitted under some circumstances. Therefore nothing substantial concerning obligations follows from intentions. For example:

$$\nvdash \left(\begin{array}{ll} \wedge & Intend(crw, A) \\ \wedge & \neg Intend(crw, \neg A) \end{array} \right) \Rightarrow O(Act(crw, A))$$

18.7 "Human Subsystems" - PARDIA

18.7.1 The Classification Scheme

The levels of the PARDIA classification are:

P*erception* **A***ttention* **R***easoning* **D***ecision* **I***ntention* **A***ction*

(...thus the name PARDIA.)
Probably the easiest way to explain the different levels is to characterize the typical behaviour of an operator in a typical working situation – e.g. the behaviour of a pilot in an operational aviation situation. An operator's interaction with the system is considered to form a sequence as demonstrated in Figure 18.2 [GLL97a].

Perception:	An annunciation of the system state is presented to the pilot;
Attention:	the pilot notices the annunciation;
Reasoning:	figures out what are the possible actions to take;
Decision:	decides on an action;
Intention:	forms the intention to carry it through;
Action:	and finally carries it out.

Figure 18.2: The PARDIA information processing model

A similar model, but with only four categories model is used by the american National Aeronautics and Space Administration (NASA) [NAS]:

> "The human operator senses the environment to acquire data, filters that data through mechanisms of perception and attention to turn it into information, processes that information to select decisions and actions, and then communicates intent and effects behavioral interaction with the environment, technologies, or other people".

We use six categories, because as we have noted there are examples of failures lying in each of these separate categories. When the source of the error is classified, we can derive particular improvements and possible points of departure for further investigations, as shown in Figure 18.3. The areas which have to be addressed to improve a particular PARDIA error differ considerably from category to category. Thus is the classification justified - but it also demonstrates that a finer classification of errors could be more useful, by narrowing down the potential areas of improvement, were it to be possible.

Error class	Classification of possible improvements
Perception:	instrument/cockpit layout
Attention:	user interface design human cognition
Reasoning:	human cognition operator training revision of handbooks
Decision:	operator training revision of handbooks Procedures
Intention:	psychological investigations physiological investigations
Action:	operator training modification of control elements user interface design

Figure 18.3: Possibly necessary improvements derivable from PARDIA errors

18.7.2 Specifying the PARDIA Model

Technically, we represent this part of human behaviour as a state machine[1]. We introduce two modules: one of axioms – *meaning postulates*, as it were – that hold invariably in every use of PARDIA; and one of *norms*, which are intended to be standard but defeasible behavior patterns. Errors are classified by noting which norm or norms is or are defeated in a given situation.

PARDIA-Axioms

Module *PARDIA-Axioms* (figures 18.4 and 18.5) defines a set of PARDIA axioms. stating the relationships between the attitudes of the model. There may be more axioms that we haven't yet found a need to use - this module constitutes a part of the meaning of the PARDIA attitudes, but is not guaranteed to be complete. Each attitude is defined as a state predicate taking two arguments:

1. An agent X (represented by a variable), as well as
2. an object A (represented by an action or state predicate) she acts on.

To express, e.g. that the Crew attends an open door, we would define:

$$Attend(CRW, open(door))$$

[1] Of course we know that our cognitive-linguistic and psychology colleagues would not accept this simple model of the human behaviour. Neither do we. Again, we do not intend to explain, but rather to classify, human error.

One can think of these primitive state predicates as intuitively describing the "state" the human process is in. The transitions between these states are also defined as axioms.

Some of these axioms require distinguishing whether the object of an attitude is an action or a state predicate. Second, the predicate ENABLED A is defined by taking the primed variables in A, replacing them with new variables that do not occur in A, and existentially quantifying them [Lam94c, Lam], e.g. consider

$$A \triangleq \begin{array}{l} \land\, x = 1 \\ \land\, x' = 3 \end{array}$$

then

$$\textsc{Enabled}\, A \triangleq (\exists q : x = 1 \land q = 3)$$

which is logically equivalent to $(x = 1 \land \exists q : q = 3)$, which, since its second conjunct is provable in ZF set theory, is equivalent to $(x = 1)$. If B is a state predicate, then B contains no primed variables and thus the operation generating the predicate ENABLED B is a no-op, leaving B unchanged. Conversely, even though there might be syntactic changes, if the predicate ENABLED B is equivalent to B, then B is logically equivalent to a state predicate, namely ENABLED B itself. We use this syntactic criterion to determine whether a predicate is semantically a state predicate or an action.

PARDIA-Norms

We claim that the principles elucidated in the norms module *ought* to hold in a normal behaviour: at some time the Crew should pay attention to something they perceived; reasoning about the attended circumstance should lead to a decision. The decision itself should cause the Crew to create some intention. Finally, an intention should lead to an action. However, the properties of *PARDIA-Norms* are defeasible. Any of these steps may fail, and this failure may be part of an incident history.

module *PARDIA-Axioms (Part 1)*

DECLARATIONS

CONSTANTS X, A, B

DEFINITIONS

$IsAction(A) \triangleq (\text{ENABLED } A \not\equiv A)$

$IsStatePredicate(A) \triangleq (\text{ENABLED } A \equiv A)$

AXIOMS

$A1 \triangleq Perceive(X, A) \land (A \equiv B) \Rightarrow Perceive(X, B)$

$A2 \triangleq Attend(X, A) \land (A \equiv B) \Rightarrow Attend(X, B)$

$A3 \triangleq Reason(X, A) \land (A \equiv B) \Rightarrow Reason(X, B)$

$A4 \triangleq Decide(X, A) \land (A \equiv B) \Rightarrow Decide(X, B)$

$A5 \triangleq Intend(X, A) \land (A \equiv B) \Rightarrow Intend(X, B)$

$A6 \triangleq Act(X, A) \land (A \equiv B) \Rightarrow Act(X, B)$

$A7 \triangleq O(A) \Rightarrow Reason(X, A)$

$A8 \triangleq Attend(X, A) \Rightarrow Reason(X, A)$

$$A9 \triangleq \begin{pmatrix} \land & Act(X, A) \\ \land & IsAction(A) \end{pmatrix} \Rightarrow A$$

$$A10 \triangleq \begin{pmatrix} \land & Act(X, A) \\ \land & IsStatePredicate(A) \end{pmatrix} \Rightarrow A'$$

Figure 18.4: PARDIA Axioms (Part1)

module *PARDIA-Axioms (continued)*

$$A11 \triangleq \begin{pmatrix} \land & Perceive(X, A) \\ \land & IsStatePredicate(A) \end{pmatrix} \Rightarrow A$$

$$A12 \triangleq \begin{pmatrix} \land & Attend(X, A) \\ \land & IsStatePredicate(A) \end{pmatrix} \Rightarrow A$$

$$A13 \triangleq \begin{pmatrix} \land & Decide(X, A) \\ \land & \neg Decide(X, \neg A) \end{pmatrix} \Rightarrow O(Intend(X, A))$$

$$A14 \triangleq \begin{pmatrix} \land & Reason(X, A) \\ \land & \neg Reason(X, \neg A) \end{pmatrix} \Rightarrow O(Decide(X, A))$$

$$A15 \triangleq \begin{pmatrix} \land & Perceive(X, A) \\ \land & A \Rightarrow B \end{pmatrix} \Rightarrow Perceive(X, B))$$

$$A16 \triangleq \begin{pmatrix} \land & Reason(X, A) \\ \land & A \Rightarrow B \end{pmatrix} \Rightarrow Reason(X, B))$$

$$A17 \triangleq \begin{pmatrix} \land & Decide(X, A) \\ \land & A \Rightarrow B \end{pmatrix} \Rightarrow Decide(X, B))$$

$$A18 \triangleq \begin{pmatrix} \land & Act(X, A) \\ \land & A \Rightarrow B \end{pmatrix} \Rightarrow Act(X, B))$$

DEFINITION

$$Spec \triangleq \land\ A1 \land A2 \land A3 \land A4 \land A5 \land A6 \land A7 \land A8 \land A9$$
$$\land\ A10 \land A11 \land A12 \land A13 \land A14 \land A15 \land A16 \land A17 \land A18$$

Figure 18.5: PARDIA Axioms (Part 2)

module *PARDIA-Norms*

DECLARATIONS

CONSTANTS CRW, A

NORMS

$$N1 \triangleq Perceive(CRW, A) \leadsto Attend(CRW, A)$$

$$N2 \triangleq Reason(CRW, A) \leadsto Decide(CRW, A)$$

$$N3 \triangleq Decide(CRW, A) \leadsto Intend(CRW, A)$$

$$N4 \triangleq Intend(CRW, A) \leadsto Act(CRW, A)$$

DEFINITION

$$Spec \triangleq N1 \land N2 \land N3 \land N4$$

Figure 18.6: The PARDIA Norms