A Quick Introduction to Why-Because Analysis

Peter B. Ladkin

1 The Background

Artifacts such as washing machines, computer systems, cars and aircraft are constructed with a purpose in mind, in order to fulfil a goal [16]. Much recent work, especially in computing science, has concentrated on how to state and fulfil such goals. There is much less work on the analysis of failure to fulfil a goal. Why-Because Analysis (WBA) puts these formal methods to work in failure analysis of complex systems. It incorporates an explanatory method, the WB-Graph Method, along with a formal verification method used to prove rigorously that a particular WB-Graph gives a satisfactory causal explanation of the incident being analysed.

In traditional engineering, failure is often signified by a persistent failure *state*: something broke, and what's broke observably stays broke. Observing failure is no longer so simple with systems designed for complex behavior, such as may be found in modern transport aircraft. As computer scientists have observed, such a system may exhibit unwanted behavior, may fail, even though nothing 'breaks'. Consider a trivial illustration: putting money into a vending machine, obtaining the desired item, but failing to receive change. Each individual action is appropriate in and of itself, but giving no change is not appropriate behavior *given* the sequence of actions and states before it.

When the machine is a computer standing by itself, then these misbehaviors can mostly be traced to their origins in the design of the system, software or hardware. But suppose the system is complex, not isolated, and includes or communicates with other digital systems, as well as human operators and physical systems, in a changeable environment. In an aircraft, say. It's now much less clear where faults can be traced: not only to individual systems, and to the operator, and to the influence of the environment, but also to the interactions between all these. Failure analysis becomes technically complex and intellectually difficult.

Suppose there has been a failure, and suppose that all the salient events and contributing system states have been researched and more-or-less discovered. There still remains the task of analysing the causality: which event was a causal factor in the occurrence of which other events and states?

A crucial observation is that the techniques developed within computer science, to specify exactly what is expected of a digital system, can also be applied successfully to the specification of continuous physical systems, human operator procedures, and even the operating environment (such as aviation regulations). WBA extends these techniques to the causal analysis of systems, which is new.

Using WBA allows causal aspects of the faulty behavior of the complex system to be specified and analysed in a uniform manner. (Although important, whether it is simple or complex to do so is a separate issue.)

2 Application: Analysis of Aircraft Accidents

Aircraft accidents are amongst the most carefully researched failures in all of engineering. One would expect the reports to be exemplary. While looking carefully at recent accident reports involving complex and often computerised aircraft, we found what appeared to us to be reasoning discrepancies: significant factors described in the body of the report, and that we judged to be causal, did not appear in the final list of causes ('*probable cause*' and '*contributory factors*') [1] But was it possible to show objectively that logic mistakes had in fact been made? We set out to develop a rigorous method of reasoning that would allow objective evaluation of events and states as causal factors. First came the criterion for *causal factor* and the Why-Because Graph (WB-Graph, or WBG) Method. (The name was suggested by Everett Palmer of NASA Ames Research Center in California [13]). The task of verifying the WB-Graph led to the full WBA. We illustrate results of the WBG Method below by discussing our analysis of the Cali and Warsaw accident reports in Sections 2.3 and 2.4 respectively. First, a few comments in Sections 2.1 and 2.2 concerning new issues arising out of the increasing use of digital systems in aviation operations.

2.1 Example: Problems of Digital Cockpits

Some problems associated with piloting in modern computerised aircraft are [14]:

- too heavy a workload associated with reprogramming flight management systems;
- too much heads-down time in the cockpit attending to the systems (pilots are encouraged to keep heads-up and focussed on what is going on outside the airplane as much as possible);
- deterioration of flying skills because of over-reliance on automation;
- increasing complacency, lack of vigilance, and boredom;
- lack of 'situational awareness' when automated systems fail, making it difficult to identify and correct problems;

[1] The observation that in causal explanation not just one 'probable cause', but normally *many* causal factors explain the occurrence of an event, and that one cannot distinguish between 'more necessary' and 'less necessary' factors, is often attributed to John Stuart Mill; for example, as quoted by Steward [18, p214]:

> It is usually between a consequent and the sum of several antecedents; the concurrence of them all being requisite to produce, that is, to be certain of being followed by the consequent. In such cases it is very common to single out only one of the antecedents under the denomination of Cause, calling the others merely Conditions....The real Cause is the whole of these antecedents; and we have, philosophically speaking, no right to give the name of causes to one of them exclusively of the others. [11, p214].

- reluctance to take over from automated systems even in the face of compelling evidence that something is wrong.

2.2 Other Dependencies on Digital Systems in Aviation

Besides these potential cockpit problems, there are wider issues concerning digital automation in the air traffic system:

- Air traffic control is increasingly dependent on sophisticated digital systems. These systems are becoming very complex (the NERC system in the UK has ca. 2M lines of code), with the goal of increasing traffic density while maintaining safety. However, backup procedures in case of failure must be equally sophisticated. Reversion to manual procedures with the higher traffic densities may no longer be as safe an option.
- Air-ground communications are increasingly digital. Some, such as ACARS messages, are broadcast over radio frequencies in the clear and are thus subject to both interception and spoofing. ACARS messages may contain dispatching information, which is safety-relevant (faulty dispatching communication contributed to the famous dead-stick landing of a B767).
- Navigation systems are increasingly digital. The US Air Traffic Association recently commissioned a study from the Johns Hopkins University Applied Physics Laboratory on whether GPS could be used as the sole navigation aid for US aviation operations [4].

As all these aviation systems become more complex, often through digitalisation, so too will the analyses of their failures.

2.3 The American Airlines B757 Accident near Cali

A recent accident which exhibited many of the cockpit problems listed above occurred to American Airlines Flight 965 into Cali, Colombia, on 20 December 1995 [2]. We analysed the Cali accident from the final report [1] using the WB-Graph Method [3]. There were approximately sixty important events and states mentioned in the report, and these appear in the WB-Graph. The 'Findings' section of the report mentions roughly a quarter of these. Ten of the 'Findings' reported causally-relevant states or events. Six of these ten correspond one-on-one with graph nodes (that is, with events or states in the body of the report). Four others correspond with multiple graph nodes - twelve of them. One *Finding* is important: it directly contradicts what the WB-Graph Method showed was a causal factor! (This factor was addressed in the US NTSB Recommendations to the US FAA.) We discuss this situation in detail in our report [3].

A partial reason for occasional mistakes in reasoning becomes clear: an accident explanation often forms a complex structure. The roughly sixty salient events and states mentioned in the Cali report have 73 direct causal connections between them! Using the WB-Graph Method, supported by internal checks with the tool *DATR-Check* [2] ensures a degree of logical rigor when dealing with such large numbers of causal connections.

[2]This accident was particularly poignant for computer scientists because Paris Kanellakis of Brown University was killed in the crash with his family.

2.4 The Lufthansa A320 Accident in Warsaw

On 14 September 1993, a Lufthansa Airbus A320 landed at Warsaw airport in a thunderstorm. Upon landing, none of the braking systems (air brakes, thrust reverse, wheel brakes) functioned for about nine seconds; the wheel brakes only started to function after about thirteen seconds. The aircraft ran off the end of the runway, collided with an earth bank and started to burn. Primarily because of the superb behavior of the crew, only two people died: one pilot, who died when the aircraft hit the bank, and one passenger, who was unconscious in the front corner and unnoticed in the evacuation as the cabin filled with smoke, and was asphyxiated. It became clear that the logic of the braking systems was indeed a reason why the braking systems hadn't functioned as expected. However, many commentators focused upon this factor as *the* main cause of the accident, which is probably incorrect. There were many other necessary causal factors. The final report [7, The A320 Accident in Warsaw] itself ascribed pilot decisions and behavior as 'probable cause'. But what criteria are being used to determine this?

We analysed this report using the WB-Graph Method [19]. One can immediately observe from Figure 2 of [19], the upper portion of the WB-Graph, that node *3.1.2: earth bank in overrun path* is a causally-necessary node: hitting the bank was a cause of the damage and fire; the hit directly killed one person and rendered the other unconscious and therefore unable to participate in the evacuation. Furthermore, this node itself is not caused by any other event or state in the sequence. It is therefore to be counted amongst the 'original causes' of the accident, according to the WB-Graph Method. However, it does not appear amongst the 'probable cause' or 'contributing factors' of the final report. We have therefore found a reasoning mistake in the report. (It is not the only node which indicates such a mistake.)

What is the consequence of this mistake? Once we have identified the position of the earth bank as an original causal factor, we know that had the bank not been where it is, the accident that happened would not have happened. (It is, of course, possible that the aircraft could have broken up and burned for some *other* reason - whether that was likely can be left to the experts to decide, but it's certainly not as likely as in the case where there's something there to hit!) Therefore, one could recommend repositioning the bank in order to avoid a repeat. However, this was not considered or recommended in the report, we suppose because the position of the bank was mistakenly not considered to be a causally-essential feature.

So, even though press and media opinion may focus on the automated systems of the accident aircraft, or on pilot '*error*', it is also true that had the aircraft had a free 'overrun area' at the end of the runway in which to slow down, the accident could have been a mere incident that would have been unfortunate but not deadly. In a valid account of the accident, positioning of this mundane terrain feature must also be noted as a causal factor along with pilot and airplane behavior.

Thus can the rigorous causal analysis proposed by WBA help the future of air travel. In the future, we will all be travellers on complex, computerised aircraft. But the computers are hardly ever the *only* cause of an accident. Now, what is WBA?

3 Why-Because Analysis

Why-Because Analysis (WBA) is a method we developed for the failure analysis of complex, *open*, *heterogeneous* systems. The adjective '*open*' means that the behavior of the system is highly affected by its environment. Aviation operations are significantly affected by the weather through which aircraft fly, for example, and landing risks are significantly affected by obstacles and their clearance on the approach and go-around paths. The adjective '*heterogeous*' means that the system has components of different types that are all supposed to work together: digital, physical, human and procedural components, and combinations of some or all of these. Modern aviation operations have all of these components. Aviation operations thus form a complex, open, heterogeneous system.

3.1 The First Step - the WB-Graph Method

The WB-Graph (or WBG) Method develops the WB-Graph of the failure scenario. The WB-Graph is a complete statement of the causal relations between all the events and system states of significance for causally explaining the failure scenario. The WBG Method consists broadly speaking of two steps:

list: make a list of all the events and states of significance;

determine causal relations: determine the causal relations between all these, exhaustively, using the semantical test for '*causal factor*' explained in Section 4.

The first step is just what an investigator does 'in the field', and is not new. The second, however, is unique to the WBG Method. It is a rigorous, objective assessment of the facts which produces a causal explanation which is objectively justifiable. The approach used builds on seminal studies in the logic of causal explanation.

Use of the WBG Method on the accident reports of the 1993 Warsaw and 1995 Cali accidents, taking the salient facts to be just those alluded to in some section of the report, has shown that the logic of these reports is faulty - both failed to identify certain demonstrably causally salient events or states in the statement of causes.

3.2 The Second Step - Verification

But how does one know that one has not himherself made errors when using the WBG Method? The full WBA includes

formal proof: a formal proof method which enables one rigorously to prove

1. that the causal relations asserted by the WB-Graph are correct, and
2. that sufficiently many factors have been identified to provide a sufficient causal explanation of each identified fact that is not a root cause.

What's the purpose of formal proofs? In system verification, experience shows that attempting a formal proof leads an experienced analyst very quickly to problems and mistakes in the analysis. In computer science contexts, attempting formal proof has exposed mistakes in widely-used telecommunications protocols, in the design of digital processors used in avionics, in clock-sychronisation algorithms for digital flight control systems, in the security of transactions in distributed databases, and in the design of cache-coherence protocols for multiprocessors. A formal proof in WBA provides an objective assessment that the analysis performed in the WBG phase is correct and relatively sufficient.

The formal-proof step of WBA can, and should, be omitted until analysts and investigators have familiarised themselves with and feel comfortable using the WBG Method. WBA specialists in formal proof can always be brought in to help if particularly sensitive investigations require the level of certainty attained by formal proof and local expertise is insufficient.

The WBA formal proof method is based upon a leading proof method used in the specification and verification of digital systems, the hierarchical proof method of Lamport [8]. All formal proof methods are technically involved and mathematical, require some specialist training to use, are best left to technical documentation [8, 6], and therefore we forego more detailed comment on the WBA proof method here.

3.3 Finding all the Facts - the Method of Difference

Finally, how does one know that one has identified all the facts salient for an explanation? One cannot be completely sure by use of logic alone – unfortunately, one cannot rule out using pure logic the (rather silly!) possibility that a fuel-tank explosion was caused by little green men playing inside it with matches. So some judgement has to be used (Artificial Intelligence specialists confronted with the same issue call such judgement a '*closed world assumption*'). Nevertheless, most investigators will know of cases in which something has been missed. For example, had the nearly-intact FMS not been found at the site of the Cali accident, significant explanatory facts would not be known, and the explanation of the accident would be unsatisfactory and incomplete.

In the Cali case, before the discovery of the FMS, one knew that there were things about the navigational behavior of the aircraft that required explaining but which the available facts were insufficient to explain. But how about the cases in which we don't know that we don't know? WBA promotes use of a version of Mill's *Method of Difference* [11], which is roughly the following.

MD: To find a causal explanation of a significant fact F, ask how the system behavior would have been different, had that fact not pertained; say behavior B. Compare behavior B with the actual behavior (which includes F). Ask: where is the first significant place they diverge from each other? A thorough description D of that place (event or state) contains a causal factor of F. Try to identify it (say G). Repeat now with G, and so on.

Use of *MD* requires a certain skill. Astute readers will have observed that the description of *MD* above is somewhat unspecific. True, but it is methodical, and we observe that many competent investigators use it implicitly anyway. Readers may prefer to consider it just an explicit formulation of good investigatory practice. We would claim that there is virtue in making such practice explicit; what is explicit is also methodical.

3.4 Irreducible Uncertainty - State Diagrams

What does an investigator do, when faced with a situation in which there is uncertainty about what happened that cannot be eliminated? In a situation in which one does not have enough facts to determine uniquely what happened? WBA provides

alternative-description: a method for delineating and representing the alternative possibilities as a *Predicate-Action Diagram* (PAD).

A PAD does not provide certainty about what happened – by assumption, we're dealing with a situation in which that is not to be had – but it does provide a precise delineation of the problem area, and so enables prophylactic measures for future avoidance of the problem area to be designed as reasonably and as precisely as possible. Use of the PAD to describe areas of uncertainty is somewhat technical and is explained in [6].

3.5 Human Actions - The PARDIA Classification

To analyse human operator behavior, WBA uses an information processing classification [12] which we call PARDIA [6, Ch. 7: The PARDIA Classification]. PARDIA stands for the sequence of stages in effective (or ineffective) situational response:

Perception – Attention – Reasoning – Decision – Intention – Action

As with any other component of the heterogeneous system, humans receive input from other components (normally through their senses) and, as 'output', influence other components of the system through action. The PARDIA classification classifies failure in the human component into one of the above components of the PARDIA sequence. Why are there six components to PARDIA (others use four, for example)? We have found that improvements to the system that may be suggested to prevent a recurrence vary significantly according to the category of error. We have found that to each of the six categories, particular different sorts of system improvement can be recommended. These are listed in [6, Ch. 7: The PARDIA Classification]. We call PARDIA a *classification* rather than a *model* because we don't claim that humans behave *as if* they were PARDIA automata – they may or they may not – but we do claim that the classification helps significantly in determining which sort of prophylactic measures for avoiding repeat incidents would be successful.

3.6 Formal Specification of Procedures and Regulations

The final feature of WBA is that the relevant procedures for human operators, and the regulations which govern them, may be formally specified in exactly the same manner as the behavior of the physical or digital components of the system. Again, this is not new – work in the 1980's in Artifical Intelligence showed that significant portions of law could be satisfactorily represented as a logic program. However, we are the first to apply such techniques to aviation procedures and regulations. The procedures and regulations, as well as the PARDIA classification, are specified all in the same formal language that the WBA proof system uses, thereby obtaining the benefits of using a '*wide-spectrum*' language for this task.

3.7 Tools

WBA is supported by a number of freeware computer and design tools.

WB-Script is a language for describing a WB-Graph in written form. *WB-Script* was developed by Michael Höhl, Peter Ladkin and Karsten Loer.

wb2dot is a tool, available for direct use over the WWW, that takes input in *WB-Script* and automatically draws the WB-Graph in *Postscript*. *wb2dot* invokes a parser for *WB-Script* which produces input for the *dot* graph-drawing tool from AT&T Research, passes this input to *dot*, and then files the *Postscript* output for downloading by the *wb2dot*-user. *wb2dot* was developed in the *PERL* multiplatform programming language by Michael Höhl.

DATR-Check is a method of answering general queries concerning the WB-Graph from a specification of each individual node of the WB-Graph and its causal relations with other nodes, without having to inspect the drawing of the WB-Graph itself. Such a query might be, for example: what are the root causes of this incident identified by the WBG Method? *DATR-Check* uses the linguistic programming language *DATR*, and was developed by Thorsten Gerdsmeier.

Emacs pf-mode is a mode for the editor *Emacs* which enables entire source contexts in a formal proof to be generated automatically by a single keystroke during editing. Writing proof source code is a lot like programming (it *is* in fact a form of programming, but in which the output is a document). *pf-mode* eliminates the overhead caused by rewriting each context individually, and guarantees that each context is correctly generated while indicating to the writer the proof level on which the context has been generated. *pf-mode* saves well over 50% of the input work involved in writing a proof, and completely eliminates situations in which the writer inadvertently creates 'unbalanced' proof levels, which is very hard to debug and unfortunately common even amongst experienced writers. *pf-mode* was written by Michael Höhl in consultation with Dirk Henkel.

pf.sty is a style package for writing proof documents using the document preparation package LaTeX. *pf.sty* was written by Leslie Lamport of Compaq/DEC Systems Research Center.

LaTeX is a document-preparation package for the typesetting language TeX, is completely portable, secure and flexible, and is the document preparation system of choice for the scientific community. LaTeXwas written by Leslie Lamport of Compaq/DEC Systems Research Center.

TeX is the typesetting language of choice for the scientific community. TeX was written by Donald Knuth of Stanford University. TeX and supporting tools are freeware for Unix and Linux systems, and are available in inexpensive packages for Windows and Mac systems.

3.8 Summary

Use of WBA, then, involves first the WB-Graph Method – building the WB-Graph from the list of facts. To help the investigation assemble the required facts, WBA promotes the use of *MD*, a version of Mill's Method of Difference. To handle resulting irreducible uncertainties, WBA proposes a method using PADs. To classify human behavior, WBA uses PARDIA. Finally, an attempt at formal proof that the explanation pictured in the WB-Graph is correct, based on a description of the incident in a wide-spectrum formal language, will expose hitherto unnoticed hiatuses in the explanation, and successful proof finally provides an objective criterion for determining the correctness of the causal explanation proposed. To our knowledge, WBA is the only failure-analysis method which provides all these features. Tools are available to support various aspects of WBA.

3.9 Analyses Performed To Date

Seven WBAs have been performed to date and one is in progress. Five of these analyses used the WB-Graph Method alone. Three include a formal proof. The status of these analyses is shown in the bar-chart Figure 1.

4 The Logical Semantics of Causal Explanation

The WB-Graph Method is based on a formal semantics for causality introduced by the philosophical logician David Lewis of Princeton University [9, 10].

Roughly speaking, the semantics of Lewis for the assertion that *A is a causal factor of B*, in which *A*, respectively *B*, is either an event or state, is that *in the nearest possible world in which A did not happen, neither did B*. This relies on a notion from formal semantics of 'possible world', best illustrated by example. Suppose my office door is open. But it *could have been* shut. A semanticist can now say: in another possible world, it *is* shut. A possible world is a way of talking about things that could happen, but didn't. But what about 'near' possible worlds? The 'nearest' possible world in which my door is shut is one in which my door is shut, air currents around it behave appropriately, sound through it is muffled as it should be, but broadly speaking everything else remains the same. A further-away world would be one in which someone else who is not me is sitting here typing, and an even further-away world is one in which this whole environment is situated in Ghana rather than Germany.

Now, suppose my door shuts. What caused it to shut? I was pushing it shut. The air was still, there was no draft, the only thing moving was the door and it was moving because I was pushing it shut. Intuitively, my actions caused the door to shut. How do I know this from the formal semantics? In the nearest possible world in which I didn't push the door, did the door shut? We have already supposed that nothing else was moving, no air drafts, no other person in the vicinity, so in the nearest world these would also be the case. It could be that all the molecules in the door moved the same way at the same time, so the door spontaneously shut - but this situation is so highly improbable as to be almost unthinkable, so could it be really the *nearest* such world? No. In the nearest world, everything behaved the same way, except that I didn't push the door. So it didn't shut. So according to my formal semantics, my action caused the door to shut.

This formal semantical test is particularly important in circumstances in which many causal factors conjoin to make something happen, which is by far the most usual case. The simple semantics asks a question of two events, or states, at a time, and by asking the question systematically of all pairs, pair by pair, a complex WB-graph may be systematically built.

References

[1] Aeronautica Civil of The Republic of Columbia. Aircraft Accident Report: Controlled Flight Into Terrain, American Airlines Flight 965, Boeing 757-223, N651AA, Near Cali, Colombia, December 20 1995, 1996. Available electronically in [7].

[2] T. Gerdsmeier. A tool for building and analysing WB-Graphs. Technical Report RVS-RR-97-02, RVS Group, University of Bielefeld, January 1997. Available from `www.rvs.uni-bielefeld.de`.

[3] T. Gerdsmeier, P. B. Ladkin, and K. Loer. Analysing the Cali accident with a WB-Graph. Technical Report RVS-RR-97-01, RVS Group, University of Bielefeld, January 1997. Included in this volume. Also in [5], and available from `www.rvs.uni-bielefeld.de`.

[4] Johns Hopkins University Applied Physics Laboratory. GPS risk assessment study final report. Technical Report VS-99-007, M8A01 Revised, Johns Hopkins University, Baltimore, MD, January 1999. Available from `www.jhuapl.edu`.

[5] Chris Johnson, ed.,. Workshop on human error and systems development. Technical Report GAAG TR-97-2, Glasgow Accident Analysis Group, Department of Computing Science, University of Glasgow, March 1997.

[6] Peter B. Ladkin and Karsten Loer. *Why-Because Analysis: Formal Reasoning About Incidents*, volume RVS-Bk-01. RVS Group, University of Bielefeld, September 1998. Available from `www.rvs.uni-bielefeld.de`.

[7] Peter B. Ladkin, ed.,. *Computer-Related Incidents with Commercial Aircraft*. Number RVS-Comp-01. RVS Group, University of Bielefeld/Causalis Limited, 1995-to date. A compendium of digitised accident reports, commentary and discussion, with updates, available from `www.rvs.uni-bielefeld.de` and on CD-ROM.

[8] Leslie Lamport, ed.,. The TLA Home Page. Available at *www.research.digital.com/SRC/tla/*.

[9] David Lewis. Causation. *Journal of Philosophy*, 70:556–567, 1973. Also in [17].

[10] David Lewis. Causal explanation. In *Philosophical Papers, ii*, pages 214–240. Oxford University Press, 1986. Also in [15].

[11] John Stuart Mill. *A System of Logic*. Longmans, London, 1843, 8th edn, 1873. Quoted in [18].

[12] Donald A. Norman. *The Psychology of Everyday Things*. Basic Books, New York, 1988.

[13] E. A. Palmer and P. B. Ladkin. Analysing an 'Oops' incident. Working Paper, 1997.

[14] Gene L. Rochlin. *Trapped in the Net*. Princeton University Press, New Jersey, 1997.

[15] David-Hillel Ruben, ed. *Explanation*. Oxford Readings in Philosophy. Oxford University Press, 1993.

[16] John R. Searle. *The Construction of Social Reality*. Penguin Books, London, 1996. Published in the USA by Simon and Schuster, 1995.

[17] Ernest Sosa and Michael Tooley. *Causation*. Oxford Readings in Philosophy. Oxford University Press, 1993.

[18] Helen Steward. *The Ontology of Mind: Events, States and Processes*. Clarendon Press, Oxford, 1997.

[19] Michael Höhl und Peter B. Ladkin. Analysing the 1993 Warsaw accident with a WB-Graph. Technical Report RVS-Occ-97-09, RVS Group, University of Bielefeld, September 1997. Included in this volume. Also available from `www.rvs.uni-bielefeld.de`.

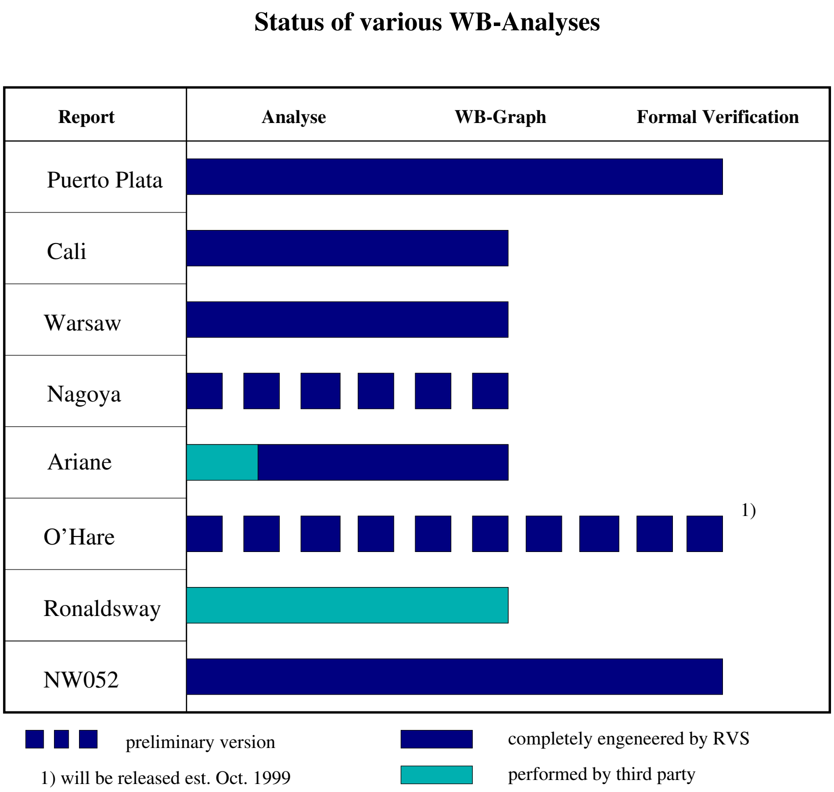

Figure 1: Status of Various WB-Analyses