AGATHA: An Integrated Expert System to Test and Diagnose Complex PC Boards

D. Allred, Y. Liechtenstein, C. Preist,
M. Bennett, A. Gupta
Information Management Laboratory
HP Laboratories Bristol
HPL-91-65
April, 1991

expert system,
diagnosis,
manufacturing,
integrated printed
circuit boards,
Prolog, causal
reasoning,
knowledge based,
test

Agatha, an expert system for testing and diagnosing HP's PA-RISC processor boards during manufacture, is a joint development of HP Labs and ICBD. It has been successfully deployed at three sites within HP, where it is in routine use, and is giving significant benefits. It has been successfully upgraded to deal with new families of PA-RISC boards. The system is implemented in Prolog and C, and is integrated with the PRISM tester. Because different tests need different styles of inference to process them, the tester consists of suites of mini-expert systems. These cooperate with each other and communicate through the diagnose manager. This paper describes the problem Agatha addressed and the results that were achieved. It also gives an overview of the Agatha architecture, and further details about the diagnose manager, the cachebus slice, and the DIP slice.

Internal Accession Date Only

1 Testing PA-RISC Processor Boards

PA-RISC is Hewlett-Packard's (HP) reduced instruction set computer (RISC) architecture that is used in its high performance computer systems.[1] Implementations of this architecture have produced some of the most complex processor boards that HP makes.[2] [3] They may contain as many as 8 VLSI chips — most of them custom, from CPUs to bus controllers to floating point processors — several high speed RAM arrays, one or more high speed busses with over 100 lines, and many other components. Due in large part to this complexity, testing of PA-RISC processor boards became a bottleneck, resulting in an undesirable backlog of undiagnosed boards, growing at a rate of 10% per month.

1.1 The Process

Part of a typical production flow for a PA-RISC processor board is diagrammed in Fig. 1, emphasizing board test.

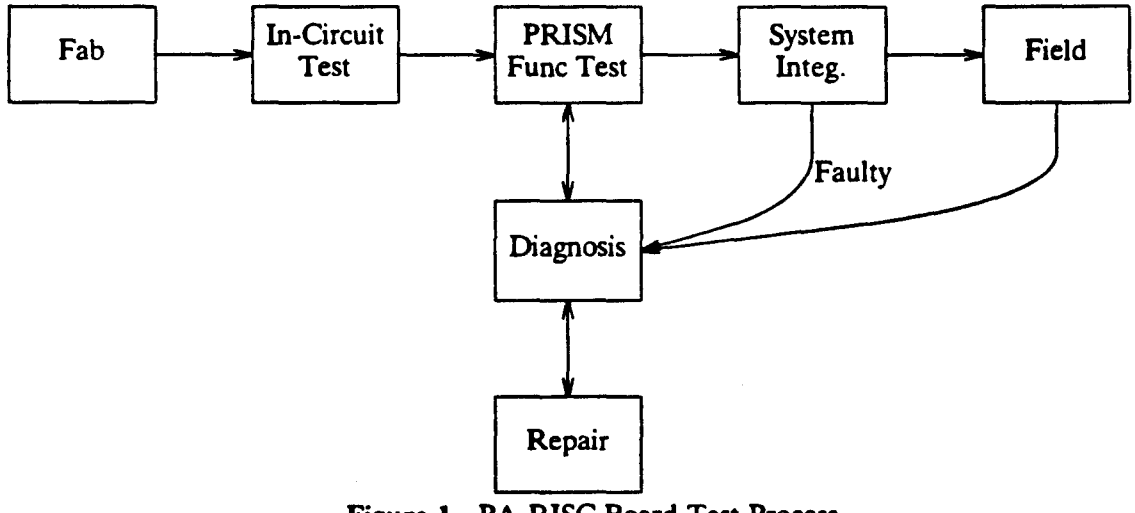

Figure 1. PA-RISC Board-Test Process

After the board is fabricated, it is run through an in-circuit or *bed of nails* test. This consists of testing the components and connections on the board individually and catches misloaded or missing parts and most of the opens and shorts.

Next, the board is tested on a PRISM tester,[4] [5] a Hewlett-Packard proprietary test system which integrates scan-based testing with functional testing. Functional testing involves testing

the behavior of various subsystems on the board, verifying that they behave properly in a system environment. Scan-based testing takes advantage of HP's proprietary scan methodology implemented on its VLSI chips in order to poke around their internal registers and test individual blocks of the chips or busses between them. All of these tests can be run at various voltages and frequencies on the PRISM tester.

At any point in the production line, or even out in the field, a faulty board can be returned for diagnosis and repair. Diagnosis is normally done at a PRISM test station where the technician has access to a battery of diagnostic tests that can help in localizing the problem.

1.2 The Problems

Because of the board's complexity, technicians found it very difficult to diagnose failed boards. Consequently, manufacturing began to encounter several problems which became quite severe:

- The PRISM test station became a bottleneck in the production process. The difficulties in diagnosing failing boards, along with long test times, began to interfere with the flow of boards on the production line.
- An unacceptable backlog of undiagnosed boards accumulated, growing at a rate of 10% per month. This was due to the time and difficulty of diagnosing certain failure modes.
- Thorough diagnosis of boards was a time-consuming and tedious process. The technicians would thus take short cuts to save time. However, this sometimes led to an incorrect part being replaced, resulting in further repair cycles being needed.
- It was very difficult to effectively train new technicians to diagnose PRISM failures. The learning curve was large. Furthermore, new manufacturing sites were being opened worldwide, exacerbating the problem.

The difficulties of diagnosing faulty processor boards in manufacturing were attributable to the following:

- As the processor board is complex, it has many subsystems with quite different functions. PRISM addresses these various subsystems with different tests, yielding a diversity of test output. A few tests try to suggest a failing component or subsystem; others just report a failure mode or code which only indirectly identifies a fault. Still others dump a lot of internal processor state information that can only be interpreted in light of the processor architecture and the test strategy. The technician has to become familiar with the output of these diverse tests in order to effectively diagnose boards.
- There are lots of *special case* failure modes which are difficult for the technician to keep in mind. Some of these failure modes are identified by unique patterns in the data, often across several test results. Because these *special cases* are less frequent than the normal failure modes, they are easily forgotten.
- Some diagnostics produce reams of data. It is difficult and time consuming for the technician to deal with such diagnostics and important information can be overlooked in the large volume of data.
- Some test results are low level — hex numbers and bit patterns — that must be manually manipulated, indexed into tables, and cross referenced in order to map to a component. This is tedious and error prone work.

- There is some uncertainty in some of the test results. For example, a test might suggest replacement of a certain chip, but in fact that chip is the actual fault only part of the time. Additionally, some faults exhibit intermittent behavior such that a discriminating test might not always detect it.

2 Problem Assessment

The nature of the problem described above suggested to us that automation of the board diagnosis process by expert system technology would be of great benefit. The expertise was available, but in short supply, and time consuming to pass on to new individuals and sites. The diagnostic process was understandable, though laborious to carry out, due to the large amounts of data and possible failure modes which needed to be addressed. The problem was significant, with a large potential benefit to be gained from removing the bottleneck in the production process.

However, two issues seemed to go beyond the standard solutions for such problems:

1. The functional diversity of the board, and as a result the diversity of the testing and diagnosing techniques, demanded several different inference strategies and knowledge-bases. It appeared to be impractical to expect a single expert system to diagnose all the different tests.
2. The difficulties technicians and operators suffered in interacting with the tester suggested that the expert system should replace the current front-end of PRISM and become not only a diagnostic inference machine, but also a high-level human interface.

The different tests were analyzed along several dimensions in order to start designing and implementing a diagnostic expert system. These dimensions were:

1. The quality of output. Some tests produce internal information which can be used in diagnosis. Others output only pass/fail data.
2. The amount of data — from a single line to hundreds of lines of hexadecimal data.
3. The degree of interaction necessary with the user. Some tests require no interaction whereas others require complex manual tests in order to perform diagnosis.
4. The sophistication and *depth* of the knowledge used by the expert — from simple heuristic matching to a deep understanding of causality within a subsystem.

3 System Design

To solve the above design issues, we decided to implement a suite of mini expert systems, called slices. Each slice diagnoses the results of a single test. The inference process for each slice was tailored according to how the test measured according to the above 4 dimensions. Where large amounts of data and causal rules are available to a slice, it uses a generation/elimination process with many data abstraction predicates, and no user interaction. Other slices need to interact with the user to gather more data before recommending further tests so they use a table lookup process.

All slices cooperate with each other and communicate through a diagnose manager, a further slice which is responsible for coordinating the overall diagnostic process and interfacing to the tester. The entire system, named Agatha, is fully integrated with the PRISM tester; furthermore, with a user interface, it forms the new front-end of the tester.

3.1 The Slice Architecture

The overall architecture of Agatha (Figure 2) consists of 9 different slices, with 27 associated knowledge and data bases. The 9 slices need only 6 different inference engines between them; sharing of inference machinery is allowed between slices with similar inference strategies. The diagnose manager slice is responsible for passing control among the other slices, depending on previous test results and diagnostic hypotheses. It is also responsible for feeding results from the slices to the user and information from the user to the relevant slice.

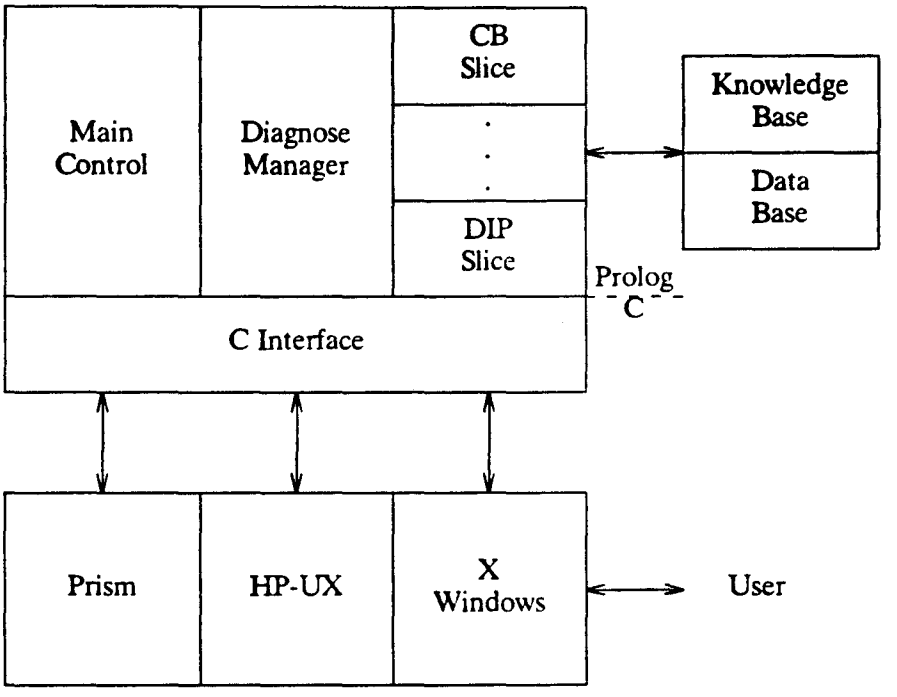

Figure 2. Agatha Architecture

To show the different inference methods involved in the slices, we focus on three. These are the diagnose manager, the cache bus slice, and the DIP (Diagnostic Interface Port) slice. We choose the latter two because they are quite different from each other in terms of the dimensions described in section 2. The cache bus slice is an example of a slice with large amounts of data but little user interaction. It analyzes the data using causal rules derived from the expert. The DIP slice, on the other hand, is an example of a slice with sparse data which uses heuristics and close interaction with the user in performing further tests to yield the final diagnosis.

3.1.1 The Diagnose Manager

The diagnose manager has overall responsibility to coordinate the running of tests, invoking slices to interpret them, reporting of suspected faulty components, and servicing requests to repair them. It delegates these tasks to three separate sub-managers:

1. The **slice manager** invokes the proper slice for a failed test, coordinates the display of suspects, and directs other requests by the slice, such as to run a test, to the diagnose manager.

2. The **repair manager** services requests to perform repairs, usually after the slices have finished analyzing the test data. After repair, it reruns the original test to confirm the problem on the board has been fixed.
3. The **test manager** services all requests for tests to be run on the tester, providing a common interface to PRISM.

The most interesting of these is the test manager, which is described here. For each test request, it must perform the following 5 tasks:

1. Select a *test-point* at which to run the test. The test-point is the power supply voltage and clock frequency to be supplied to the board under test by the PRISM tester.
2. Run the test.
3. Determine the failure mode of the test (e.g. is this test intermittent?).
4. Retry the test if necessary, and
5. Recommend to the diagnose manager what to do next.

First, the test manager selects a test-point at which to run the test. It may select one of a set of predefined test-points or an entirely new one in an attempt to get a test to fail more reliably.

The test manager then runs the test at the computed test-point, parses the test output, and makes a summary decision about whether it passed or failed. No attempt is made to otherwise interpret the test results at this point.

Next, based on the pass/fail results of the test, a new failure mode is computed. The test manager supports the following failure modes:

- **Hard.** Failing tests fail at all test-points with little or no variation in their output (i.e. they don't differ dramatically from test-point to test-point).
- **Dependent/Repeatable.** Dependencies on voltage or frequency have been determined (i.e. it only fails at certain voltages/frequencies), and it's repeatable.
- **Dependent/Non-repeatable.** Dependencies on voltage or frequency have been determined but it's **NOT** repeatable (it only fails sometimes at those voltages/frequencies).
- **Intermittent.** It's not failing hard, but no dependency can be determined (e.g. random failures).

The failure mode is a symbolic representation of uncertainty in the test results — accounting for possible intermittent test failures. Heuristic knowledge about tests, along with data about test results, is used to infer the failure mode. For example:

```
IF    this board is a field return AND
      the board test passed
THEN the failure condition is intermittent
BECAUSE this board has failed before (in the field),
      but won't fail now
```

With a new failure mode computed, the test manager next determines whether the test should be rerun, due to suspected uncertainties in the test results. Heuristics are used to control this; for example:

```
IF   the test passed AND
     the failure mode is either dependent-non-repeatable or
      intermittent
THEN the test should be rerun up to 3 times until it fails
```

Finally, the test manager is ready to recommend to the diagnose manager what to do next. If the test failed, the associated slice is determined and it is recommended for invocation. If the test passed, then the previous slice is reinvoked with this new information; as a result, the slice may recommend further tests.

Thus the test manager uses heuristic knowledge embedded in procedural control to run tests and manage uncertainty in their results, removing this burden from the slices.

3.1.2 The Cache Bus Slice

The cache bus subsystem on the assembled circuit board consists of several large VLSI chips which communicate via a bus of around 150 lines. Each chip is connected to a subset of these lines, and is able to transmit onto the bus by driving values onto the lines. These can in turn be read by the other chips.

Failures which can occur in this system include shorts between lines, opens on lines, and chips failing to transmit or receive properly. A fault may be intermittent; i.e., sometimes a test will miss it, while at other times, it will show up. Also, multiple faults can occur. For example, several shorts can be caused by solder being splashed across the board.

The PRISM tester tests the cache bus by getting each chip to drive a series of binary values onto the bus, and getting all chips to read each value back. This process is done automatically, and produces a very large amount of data. Discrepancies between the expected values and observed values in this data are used by the cache bus slice to diagnose the fault.

The cache bus slice has the responsibility of diagnosing failures in the cache bus test. The design of this slice (see Figure 3) allows it to handle both intermittent faults and the majority of multiple faults. Rather than dealing with the bus system as a whole, it subdivides it into semi-independent subsystems, namely, each line.

The data from the test is divided into *batches*, one associated with each line which is exhibiting bad behavior. Each batch contains only that data which was received incorrectly off its associated line. This is then used to deduce the fault on this particular line. Hence, rather than assuming that the cache bus as a whole has only a single fault, the system can treat each line independently, and assume that there is at most one fault per line. The single fault assumption is replaced with the single fault per line assumption.

A template generates a list of hypotheses for each line. Where a multiple fault on a single line is to be considered, this is explicitly entered in the template. Elimination rules are then used to remove as many of these as possible. The knowledge in these rules is derived from the causal rules of the expert, by taking it's contrapositive — if a hypothesis makes a prediction, and that prediction is found to be false, the hypothesis can be eliminated. Elimination rules take the form:

```
IF        a '1' is observed on the line
       when a '0' was expected
ELIMINATE short to ground and
       short to another cache bus line
BECAUSE   these faults can only pull a line to '0'.
```

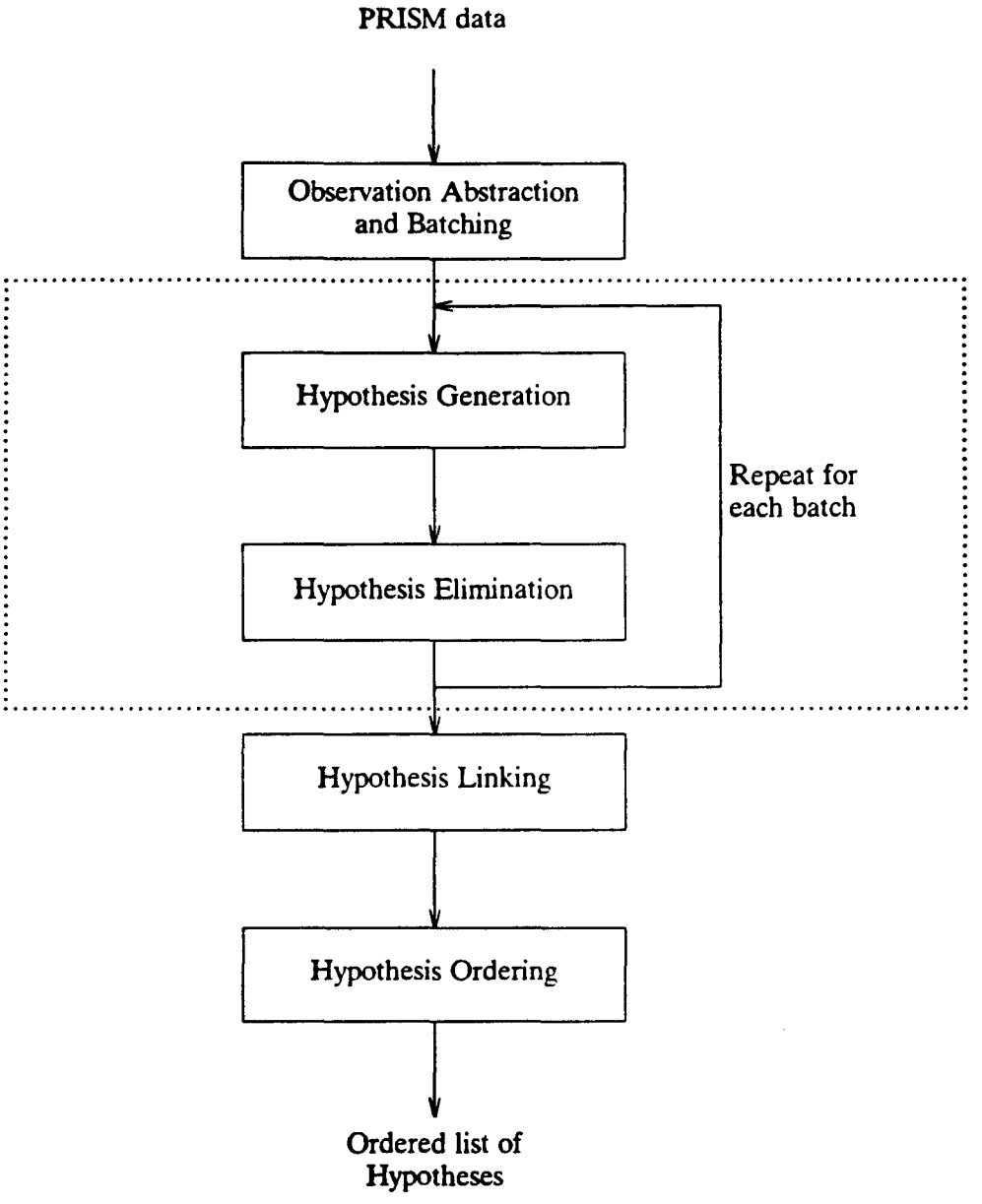

Figure 3. Cache Bus Slice Architecture

After elimination, each line will have a small number of hypotheses associated with it. However, a line is only semi-independent. Many hypotheses, such as "short between lines", and "bad VLSI", will manifest their behavior on several lines. Hence, linking rules are used. These take hypotheses associated with different lines, and combine them, where appropriate, into single hypotheses explaining the bad behavior of several lines. They have the following form:

```
LINK: Short to cache bus on Line 1
WITH: Short to cache bus on Line 2
IF:   Line 1 and Line 2 are adjacent somewhere on the board
TO GIVE HYPOTHESIS: Short between Line 1 and Line 2.
```

Finally, the remaining hypotheses are ordered, using heuristic knowledge, according to their likelihood. This list is returned to the diagnose manager for presentation to the user.

The cache bus slice thus combines causal-based rules with heuristic knowledge. The causal rules are used to deduce which hypotheses are possible and which are impossible. The heuristic knowledge is then used to determine the relative likelihood of the hypotheses which remain.

Full details of the cache bus slice are being published.[6]

3.1.3 The DIP Slice

Before scan-based tests can be performed on a VLSI chip, the Diagnostic Interface Port (DIP) on the chip, which is the serial scan port, must be tested to verify that the tester is able to communicate properly with the chip and that other electronic subsystems are working reasonably well (i.e. the power, clocking, and reset subsystems). This test, the DIP test, tests the DIP port on all VLSI chips and the DIP slice diagnoses any failures.

Unlike the cache bus slice, the DIP slice works with very simple pass/fail and test-point information from the DIP test and interacts with the user to give the final diagnosis. Because of the sparsity of data, it is unable, alone, to deal with intermittent and multiple faults, relying instead on the user to explore these possibilities.

The DIP slice performs its diagnosis in two stages. First, it proposes which subsystems it considers are the main and secondary suspects. Second, it aids the user in performing further manual tests on a particular subsystem to determine if it is indeed faulty, and if so, exactly where.

The first stage, generating suspects, uses a heuristic mapping from the symptoms to the possible causes. The symptoms are the pass/fail data of the DIP test. The possible causes, divided into main and secondary suspects, are the candidate subsystems and connections of which at least one (possibly more) is faulty. There are twenty such mapping rules; the following is an example:

```
IF      all chips failed the DIP-test
        AND the System-Test passed
THEN    suspect the MDA as a main suspect
        AND the Reset, Clock, Power and System-bus-connector
            as secondary suspects.
```

The main/secondary distinction is a very simple form of probability handling. More complex schemes were rejected, as the expert couldn't substantiate finer separation of fault likelihood.

The second stage of reasoning, that of guiding the user in manual tests, is an iterative process. The user chooses to concentrate on a particular subsystem and tries to find out which of its components (if any) are faulty. The decision of which subsystem to focus on is left up to the user.

Each subsystem is composed of a set of components and a list of tests to test them, both automatic and manual. A table (Figure 4) then represents the knowledge that tests would give about the components as follows: if a test fails, at least one of the components with an "F" entry in the table will be faulty; if a test passes, all the components with a "P" entry in the table

Components:

1	Tester
2	System-bus (signals PON/NPFW)
3	Reset-buffer
4	SIU
5	Cache-Bus (signals NRS0 and NRS1)
6	VLSI chips (other than SIU)

Tests:

1	Probe NRS0/1 on Cache Bus
2	Scope NRS0/1 on Cache Bus
3	Ohm out NRS0/1 on Cache Bus
4	Probe PON/NPFW on System Bus
5	Scope PON/NPFW on System Bus
6	Ohm out PON/NPFW on System Bus
7	Probe and scope PON/NPFW on both sides of buffer
8	Inspect System Bus connector
9	Test rest path through tester subsystem

F/P Table:

1 Tester	2 Sysbus	3 Buffer	4 SIU	5 Cache Bus	6 VLSI	Tests to perform
F	F	F	F	F		1
			F/P	F/P		2
				F/P		3
F	F	F				4
F/P	F/P	F/P				5
	F/P					6
		F/P				7
	F/P					8
F/P						9

Figure 4. DIP Slice Knowledge for the Reset Subsystem

will be functional (not faulty).

Using this knowledge, together with an approximate "cost" of performing each test, the DIP slice presents the user with an ordered list of tests which are worth carrying out. The user then chooses which test to perform, receives instructions on how to do it, performs the test, and enters the result (Pass or Fail) into the system. This continues until the DIP slice is able to diagnose a component as faulty, or the user chooses to explore another subsystem.

Hence the user, working with Agatha, is able to explore the different candidates and diagnose exactly which of them is indeed failing. The user is always in control, yet can rely on an ordered set of tests, arranged so as to isolate the fault as fast as possible.

3.2 Tester Integration

As indicated above, in order to solve crucial testing issues, Agatha would not only be an automated diagnostic system, but would provide a user interface and become the new front-end to PRISM.

The challenge here was to integrate Agatha into an old PRISM system whose code had not been touched for a long time. We opted to layer Agatha on top of PRISM, as diagramed in Figure 5. Communication is via the HP-UX interprocess communication (IPC) facilities. With this layering the PRISM code remains virtually untouched, except for minor modifications of the Stream files or scripts.

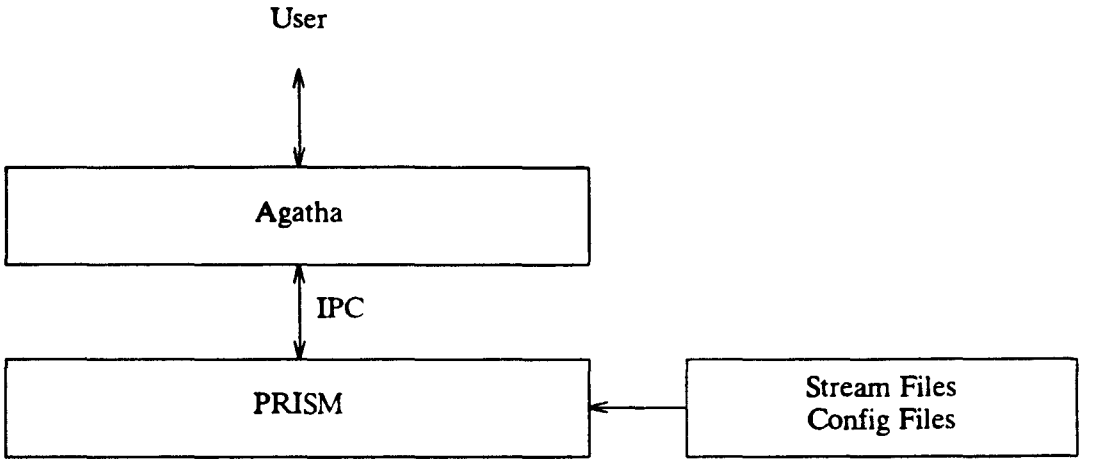

Figure 5. Agatha/PRISM Interface

The layering of Agatha on top of PRISM had several advantages:

1. PRISM code remained unmodified, minimizing its maintenance requirements.
2. A new, friendlier front end was now provided for PRISM by Agatha.
3. Agatha now detects some failures that PRISM could not detect before. For example, there are some messages that are printed by system tests that PRISM never recognized, possibly passing faulty boards. Agatha now detects these and reports them accordingly, improving the reliability of the testing process.
4. Some tests that required a long, tedious sequence of commands to execute are now automated by Agatha, saving time and being more thorough as the technician would otherwise shy away from running the test to completion or even running it at all.

3.3 How Agatha is Used

In manufacturing, *operators* test the boards on the PRISM testers, binning the good and bad boards separately. Later, technicians put the failed boards back on the tester and diagnose them. This was modeled in Agatha, supporting 3 levels of users. These are, in order of increasing capability:

1. **Operator** — the user can only test boards and bin them,
2. **Technician** — the user can also diagnose boards, including running manual tests, etc., as well as access other utilities, such as review/print results of past tests, etc.
3. **Maintainer** — the user can also edit the knowledge/data bases and drop down into Prolog.

This made a simpler interface for the operator to learn, greatly reducing the learning curve. At the same time additional flexibility was available for technicians, making their task simpler. This proved a significant contribution of Agatha — a major benefit to our users.

With this model, Agatha could run several additional diagnostic tests while in operator mode which would save time for the technician who would otherwise have to run them later. Thus a mode is provided, called "automatic test", where Agatha automatically runs all tests it believes necessary, provided no intervention is required from the user. This provides a significant time savings to the technician and effectively reduces the skill level required for this task. This feature can be turned off during times of heavy workload where higher throughput is needed from the operators.

One major decision was whether to provide a diagnostic system, or a diagnostic adviser. Specifically, would it dictate to the user what to repair, or only advise? During knowledge acquisition, we found that which component to repair would vary depending on certain circumstances, including variations in the production process that could cause one failure mode to start appearing more frequently. Hence, Agatha only advises the user on repairs, presenting an ordered list of candidates to the technician who chooses the most suitable. The technician would usually pick the first item from the list, unless aware of extenuating circumstances that might suggest another.

4 Development

Agatha was a joint development effort between the Knowledge-Based Programming Department (KBPD) of HP Labs, Bristol, England, and the Integrated Circuit Business Division (ICBD) in Fort Collins, Colorado. The development process was broken down into the following phases.

1. The Alpha Phase produced a prototype version which was deployed first in ICBD. It consisted of only 3 slices. The goal was to gain experience with it and get user feedback.
2. The Beta Phase reviewed this feedback producing major refinements to the slices and a new Diagnose Manager. More slices were added. This produced the first production version of Agatha, installed at a user's site.
3. The Refinement Phase continued to add slices and make refinements to the knowledge. More users were added, leading to the Manufacturing Release of Agatha.
4. The Maintenance Phase followed where minor enhancements and refinements are ongoing.

4.1 Implementation Language

Prolog was chosen as the principal implementation language for Agatha. (In the diagram of Figure 2 everything above the "C interface" box was written in Prolog, including the knowledge and data bases.) The main reason for choosing a language, rather than an AI shell or toolkit, was the need to be able to code up different inference strategies for the different slices. These strategies do not always fit into classical forward/backward chaining regimes provided by shells, and so would have been awkward and inelegant to code in this way. The disadvantage with this decision is that we lose the support the shell provides — namely a good user interface, and ready-written inference strategies. These had to be coded and maintained, imposing some additional burden on the project, yet could be tailored to the task completely, not restricted by what a shell or toolkit provides.

4.2 Verification

Prior to Agatha all tests run on PRISM were sent directly to the printer. We had lots of printouts to use in the design of Agatha, but no machine readable tests to verify the implementation. Therefore, *scaffolding* was built to capture test results on disk to be used for verification. Using this scaffolding a large suite of verification cases was gathered from the following sources.

- *Poisoned* boards. Faults were caused on an otherwise good board, including shorts, opens, and missing parts.
- Bad VLSI. Bad VLSI (acquired from field returns, etc.) were inserted into a socketed board.
- Later on the verification cases were augmented with actual cases that Agatha encountered while in use on the production line.

These verification cases helped to verify the knowledge and functionality of Agatha, and refine it while in use. They helped assure delivery of a reliable, confident system to our users.

4.3 Maintainability

Maintainability was a foremost consideration throughout the design and implementation of Agatha, for two reasons. First, it was decided from the outset that the original designers of significant parts of the system, HP Labs, would not be responsible for their maintenance. Instead it would be the work of ICBD. Second, the system had to be able to deal with new board types, which were structurally different from the original board, but were tested using the same tester.

This led to the clear separation, in each slice, of the knowledge specific to a certain board type and the knowledge specific to the tester. Hence, the tester rules had to deal with an abstract board, and call the structural knowledge base to gather information specific to a certain board type.

Where the knowledge consisted of simple relations (such as which faults resulted in the failure of which further tests), it was represented directly as relational tables. Rules then access these as necessary. This reduces the number of rules needed, and allows easy and rapid maintenance.

This policy has paid off. The system is now entirely maintained by ICBD. They have successfully updated the system to support three different board families, with a fourth nearly completed. (Some families have multiple board types distinguished by varying cache ram sizes, etc., all of which Agatha has to know about.) This update process has been partly automated. C routines access design data files used by the PRISM tester, and extract structural information which is of relevance to Agatha, constructing files of Prolog clauses. The update process takes only a short amount of time to deal with a new board family.

Other than supporting new tests and board families, most of the maintenance since production release of Agatha has been in adding new features and enhancements, rather than refining the slices or their knowledge. And most of these enhancements were outside the slice architecture proper.

4.4 Deployment

Agatha has been in routine use since January of 1990. It has been successfully deployed at three sites within HP — two are manufacturing facilities, the third is a field repair center. One manufacturing site, for example, uses it 24 hours a day: operators test boards on the production line, often letting Agatha diagnose them to the extent that manual testing isn't required. Technicians then examine failure reports and let Agatha work on boards that may need further diagnosis.

The field repair center receives failed boards from the field and diagnoses them on Agatha. Although they don't see the volume of boards the production sites do, for that very reason Agatha is perhaps even more critical to their operation — it preserves diagnosis and repair knowledge that they might otherwise lose with low volumes. Furthermore, they have a great need for some of the features of Agatha that support uncertainty: since a board returned from the field is suspected as faulty, one can't simply return it if the board-test passes when run just once; it must be thoroughly exercised and checked out. Agatha addresses such intermittencies by running a test many times to try to get it to fail.

5 Results

Agatha has been very favorably received by its users and has proved to have many benefits to them. First of all, it has addressed many of the production problems that were being experienced prior to Agatha.

1. The PRISM test station is much less a bottleneck in production. Though the long raw test times (which Agatha has no control over) are at times cause for congestion in production, the savings in diagnosis time have greatly alleviated this problem.
2. Agatha, in combination with other efforts on the production line, has helped virtually eliminate the backlog of undiagnosed boards.
3. Agatha saves time by running tests automatically, especially when run by an operator. Agatha also automates some tests that used to be painstaking manual tests, saving additional time. Time savings is one of the principal benefits hailed by all Agatha users.
4. Also, costs are saved when Agatha runs tests automatically. By effectively lowering the skill level required it can be done by operators rather than technicians.
5. Agatha makes a more thorough diagnosis, eliminating many common human errors. It also improves test reliability by detecting failures not formerly caught by PRISM that could go through undetected.
6. Agatha has provided an easier to use interface to the PRISM tester. In a recent survey, users gave a top score to Agatha's friendly user interface. This, coupled with automation, has significantly reduced technician and operator training time and greatly improved user satisfaction.

It's difficult to quantify the full impact of Agatha within Hewlett-Packard but some benefits are mentioned below.

- During Agatha development, test time was reduced by 80% on one board, yielding tremendous cost savings. Though there were several factors at work here, Agatha was a principal contributor in this effort.
- The field repair center indicates Agatha has reduced scrap rate, average repair time, training costs, and material costs. This adds up over the life of a product, and could especially

be valuable toward the end of its life when expert knowledge on an aging board could otherwise be scarce.

- One production site related a one fourth reduction in technician training time, with the *ramp-up* time for a new technician dramatically improved. They also reported a 40% reduction in diagnose time and a significant increase in user satisfaction with the friendlier interface.

In addition to the gratifying manufacturing results, the joint development effort between HP Labs and ICBD has proved to be of mutual benefit:

- ICBD gained expertise on the design, development and deployment of expert system technology.
- HP Labs gained knowledge of realistic problems in electronic circuit diagnosis. This has been used to drive a longer term research program in model-based diagnosis.

ACKNOWLEDGMENTS

We wish to recognize the valuable contribution of many others. Rick Butler was the domain expert who consulted on the project. Caroline Knight of HP Labs invested a lot of time in training ICBD on knowledge acquisition techniques. Jason Brown was a summer student who helped code part of Agatha. We express gratitude to the many operators and technicians who gave invaluable time, assistance, and feedback to the Agatha project.

REFERENCES

1. Mahon, Lee, Miller, Huck, and Bryg, *Hewlett-Packard Precision Architecture: The Processor*, **HP Journal,** Vol. 37, No. 8, August 1986, p. 4-21.
2. Robinson, et al, *A Midrange VLSI Hewlett Packard Precision Architecture Computer*, **HP Journal,** Vol. 38, No. 9, Sept. 87, p. 26-34.
3. Gassman, et al, *VLSI-based High-Performance HP Precision Architecture Computers*, **HP Journal,** Vol. 38, No. 9, Sept. 87, p. 38-48.
4. Schuchard and Weiss, *Scan Path Testing of a Multichip Computer*, **1987 IEEE International Solid State Circuits Conference Digest,** (Publisher: Lewis Winner, Coral Gables, FL 33134), p. 230-231.
5. Weiss, *VLSI Test Methodology*, **HP Journal,** Vol. 38, No. 9, Sept. 87, p. 24-25.
6. Preist, Allred, and Gupta, *An expert system to perform functional diagnosis of a bus subsystem*, HP Technical Paper, HP Labs, Bristol, England, to be published.