

Explaining and Recovering from Computer Intrusions: Progress 7/97



Douglas B. Moran
Pauline Berry

**Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park CA 94025**

Overview/Review

- Apply AI Technology to Network-based Intrusions
 - diagnosing
 - explanation
 - reporting
 - repair and recovery: automated security manual

Overview/Review (cont.)

- Tool:
 - manages other tools (esp. OTS) for data collection
 - rule-base to guide/advise user
 - add additional tools as agents
 - add intrusion scenarios
- Approach: start with common cases and build outward and upward

Status

- DERBI: Diagnosis, Explanation and Recovery from Break-Ins
- Single host testbed: full system on removable drive
- Integrating tools and information sources
- Pending: trying to identify a partner to beta-test tool and extensibility

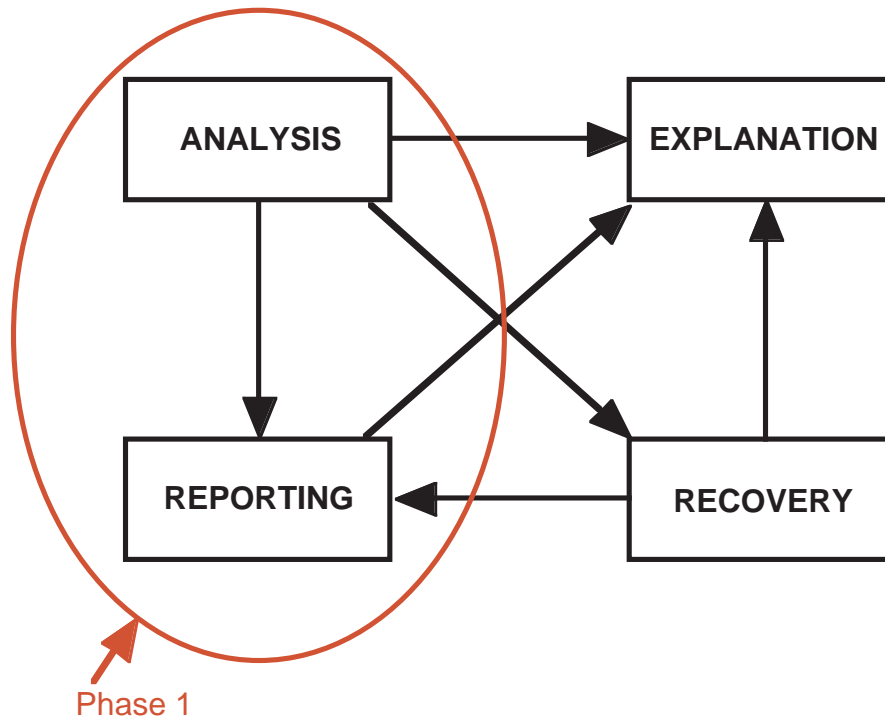
The Role of PRS in DERBI

- DERBI: Objectives and Techniques
- What is PRS
- Where PRS fits
- Initial Development
- Benefits and Problems

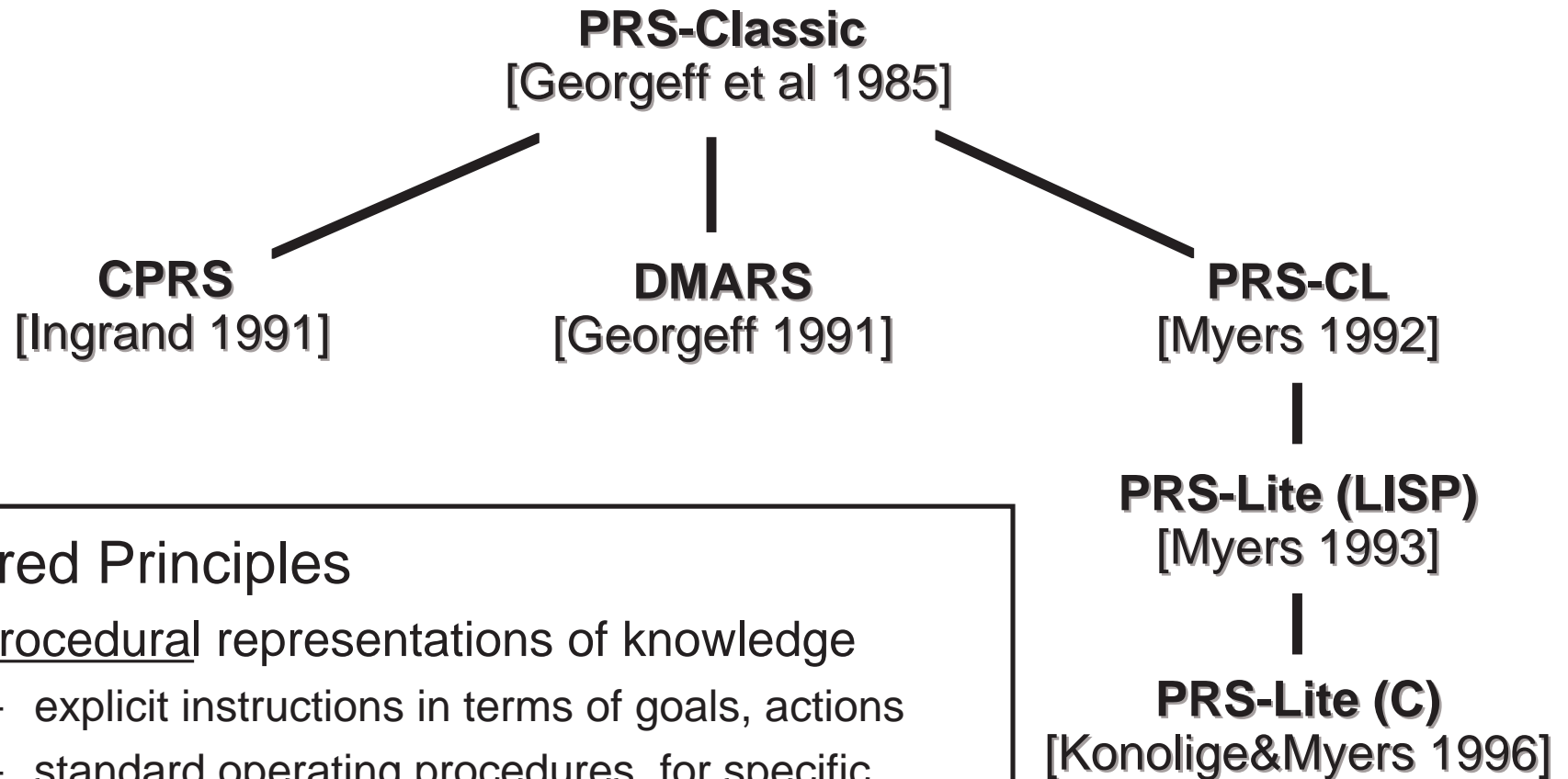
DERBI Analysis

- Confirm and Trace a break-in given suspicious condition
 - Where, What, When, How
- Report Trace and Conclusions
- Existing Sets of useful Tools, Tool Components, Methods and Experts
 - varying costs, side-effects, benefits
- Inexperienced System Administrators
- **NEED:** Overall Coordination of Processes

DERBI design



PRS - (history slide)



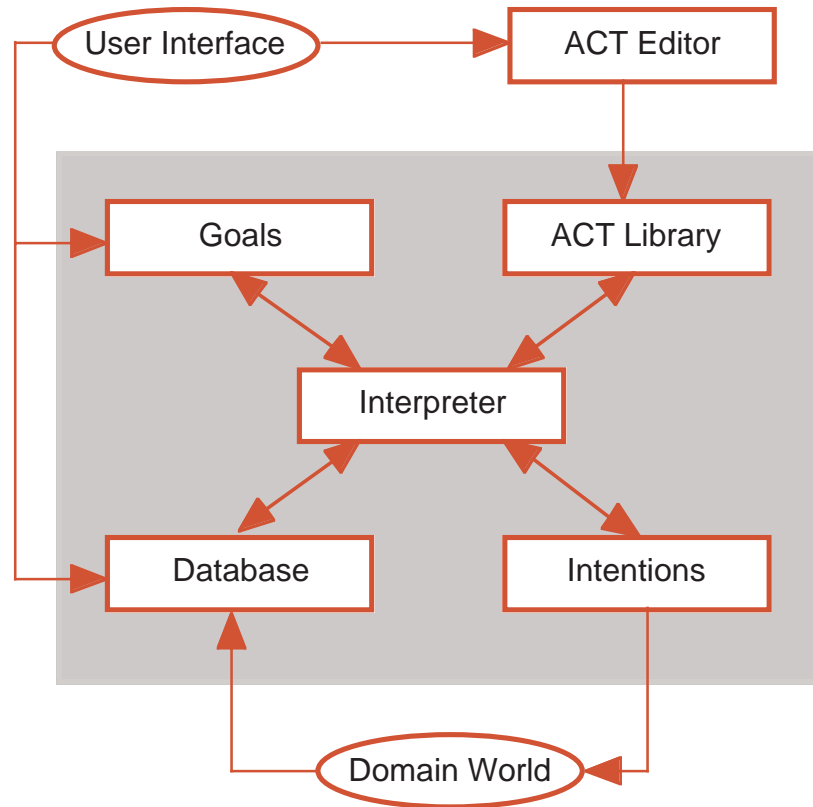
Shared Principles

- procedural representations of knowledge
 - explicit instructions in terms of goals, actions
 - standard operating procedures for specific situations
- combine goal- and event-driven activity

What is PRS?

- What is PRS?
 - a framework for building reactive controllers
 - smoothly integrates goal- and event-based activities
 - Goal-based behavior: *send a robot to target location*
 - Event-based behavior: *respond to unexpected obstacles*
 - supports distributed problem-solving through multiple, communicating agents
 - metalevel reasoning for defining complex control strategies

A PRS Agent



Goals: conditions to be satisfied

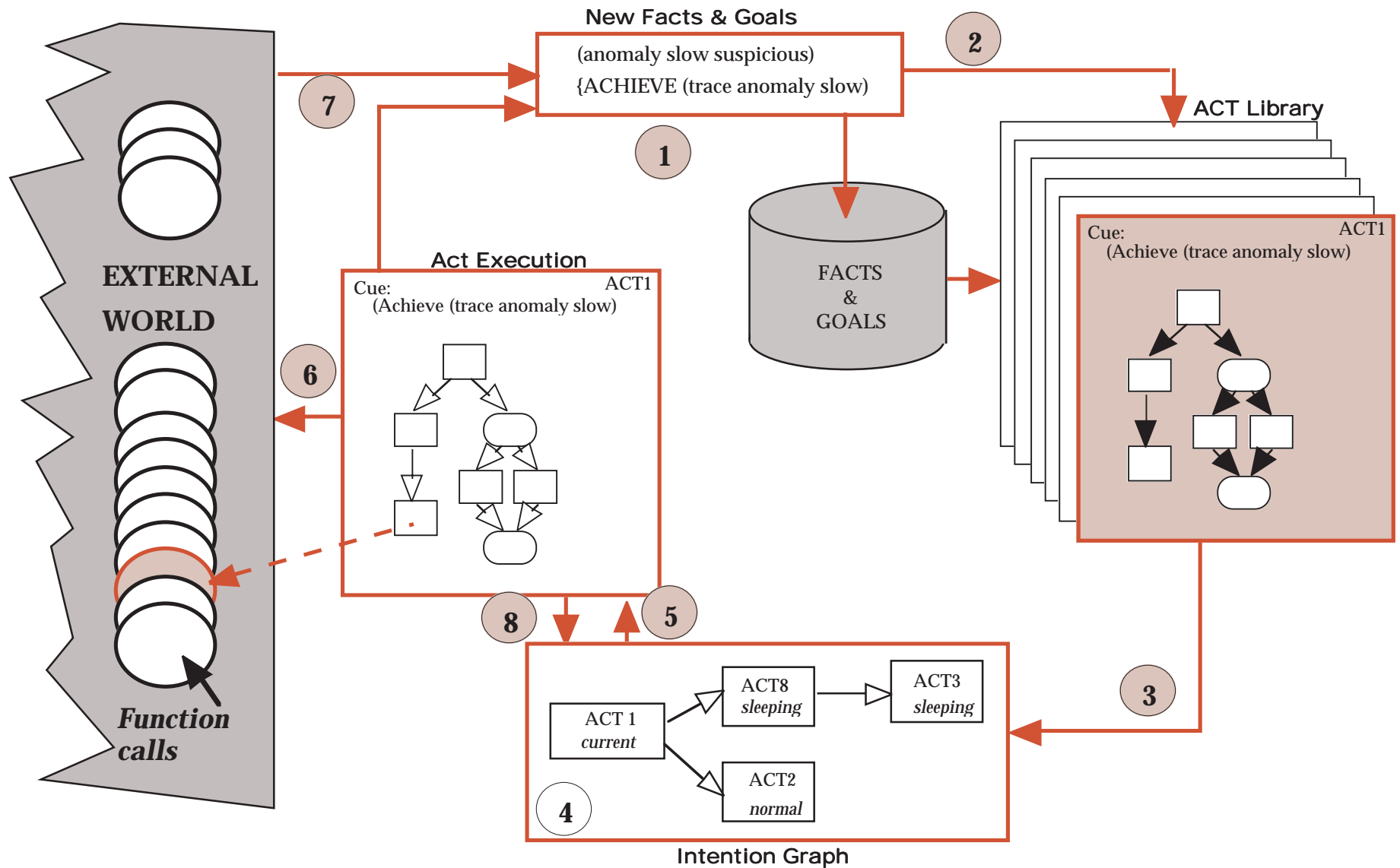
Database: beliefs about the World

Act Library: possible actions/activities

Intentions: tasks to be executed

Interpreter: over-all controller

PRS Architecture in DERBI



Execution Cycle

1. New information arrives that updates facts and goals
2. Acts are triggered by new facts or goals
3. A triggered Act is intended
4. An intended Act is selected
5. That intention is Activated
6. An Action is performed, usually a call to an external function
7. New Facts or Goals are posted
8. Intentions are updated

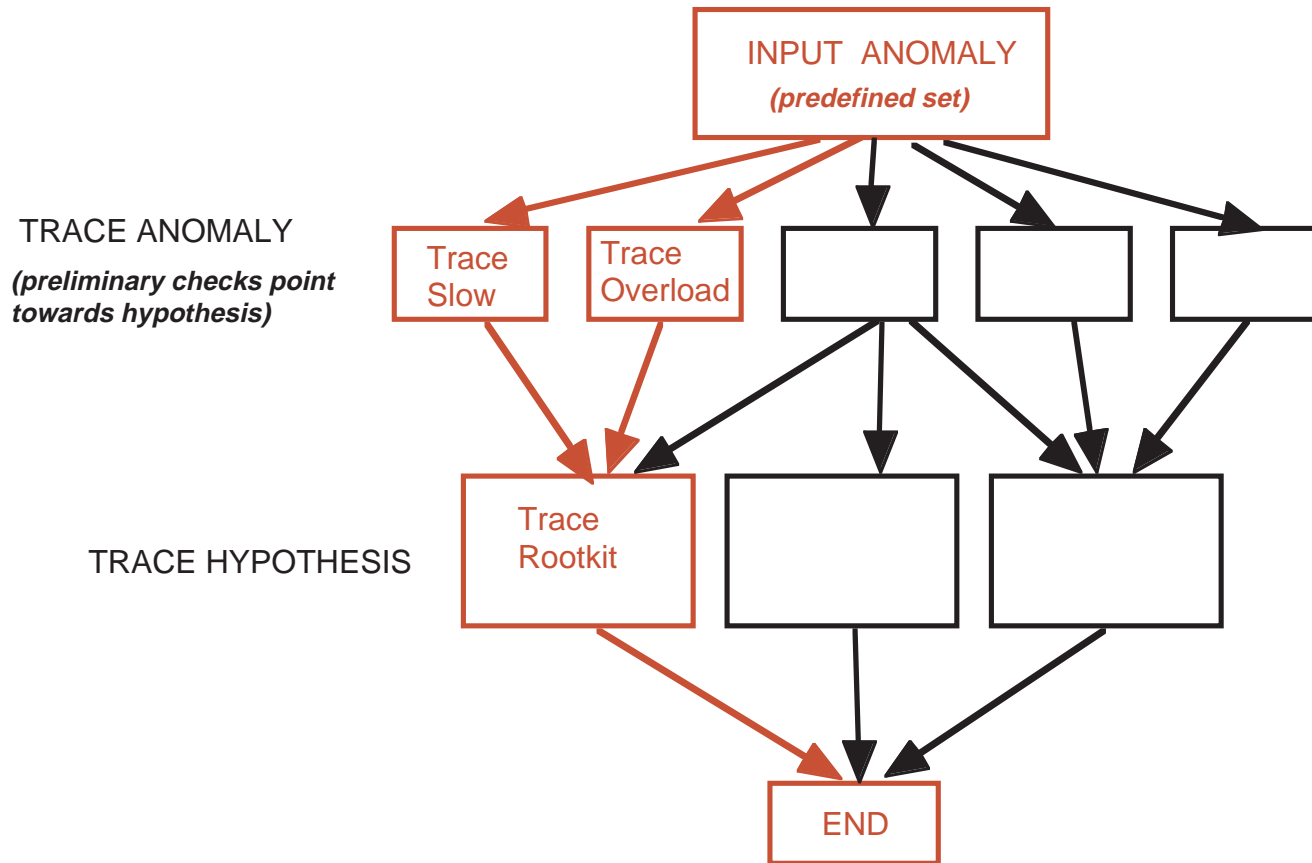
DERBI Design Issues

- Cost vs. Benefits
 - compare system output with trusted copy of ps, ls, du, ifconfig and netstat
 - compare md5 checksums of suspicious copies with known values
- Speed vs. Side-effects
- Coverage vs. Focused Search

Rootkit Scenario

- Initial Anomaly Reports
 - Another site reports suspicion of your site involvement
 - Unexplained disk fill-up
 - Slow-running/hidden processes
 - Certain suspicious files
- Strategies
 - Confirm Unauthorized system/root access
 - Check suspicious files by name
 - Check for hidden files and processes
 - Check for replaced system commands
 -

Anomaly Leads to Hypothesis



Trace Anomaly: Slow

