

# Automated Theorem Proving

By Damian Jakusz-Gostomski

[jakuszd6@cs.man.ac.uk](mailto:jakuszd6@cs.man.ac.uk)

7060823

# What is Automated Theorem Proving

- Automated Theorem Proving deals with the development of computer programs that show that some statement (the *conjecture*) is a *logical consequence* of a set of statements (the *axioms* and *hypotheses*)<sup>1</sup>
- **Axiom** – All men are mortal, Socrates is a man
- **Conjecture** – Socrates is mortal

1) <http://www.cs.miami.edu/~tptp/OverviewOfATP.html> - 30/09/2009

# How it works

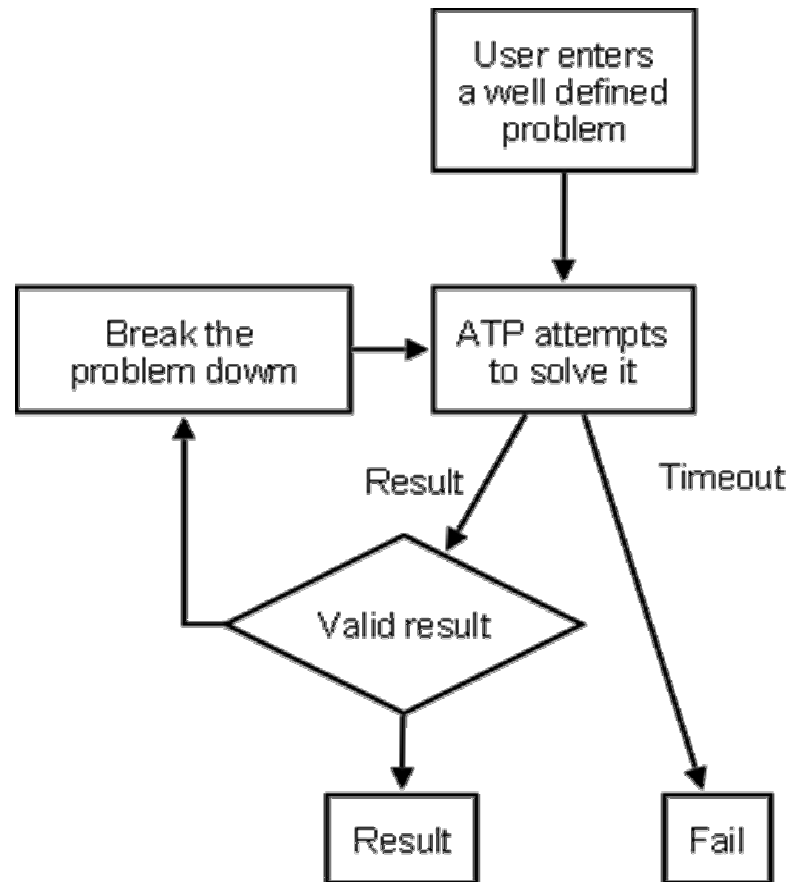


Figure 1: The flow of a ATP

Figure 1 shows a simplified overview of how a ATP works.

If the ATP gives a result, it needs to be checked by a domain expert. If the result is wrong, the problem may need to be given in a different form, or broken down into a set of smaller problems

# How it's used

All of these are tasks that can be performed by an ATP system, given an appropriate formulation of the problem as axioms, hypotheses, and a conjecture.<sup>1</sup>

- A mathematician might prove the conjecture that groups of order two are commutative, from the axioms of group theory
- A management consultant might formulate axioms that describe how organizations grow and interact, and from those axioms prove that organizational death rates decrease with age
- A frustrated teenager might formulate the jumbled faces of a Rubik's cube as a conjecture and prove, from axioms that describe legal changes to the cube's configuration, that the cube can be rearranged to the solution state

1) <http://www.cs.miami.edu/~tptp/OverviewOfATP.html> - 30/09/2009

# The Pentium FDIV bug and ATP

- ATP is used heavily in integrated circuit design and verification
- Since the Pentium FDIV bug, processor manufacturers have used ATP to verify division and other operations work

A number multiplied and then divided by the same number should result in the original number, as shown below

$$4195835 * 3145727 / 3145727 = 4195835$$

The result given from the flawed Pentium

$$4195835 * 3145727 / 3145727 = 4195579^1$$

1) [http://en.wikipedia.org/wiki/Pentium\\_FDIV\\_bug#Affected\\_models](http://en.wikipedia.org/wiki/Pentium_FDIV_bug#Affected_models) - 30/09/2009

# Existing Automated Theorem Provers

Name	Description
<b>Otter</b>	An ATP system for statements in first-order (unsorted) logic with equality. Otter is based on resolution and paramodulation applied to clauses
<b>Vampire</b>	Vampire is an automatic theorem prover for first-order classical logic. Its kernel implements the calculi of ordered binary resolution and superposition for handling equality
<b>Waldmeister</b>	Waldmeister 704 is a system for unit equational deduction. Its theoretical basis is unfailing completion in the sense of [BDP89] with refinements towards ordered completion
<b>Mace2</b>	Mace2 searches for finite models of first-order (including equality) statements. Mace2 iterates through domain sizes, starting with 2. For a given domain size, a propositional satisfiability problem is constructed from the ground instances of the statements, and a DPLL procedure is applied.