2. Geometry Theorem Proving

This chapter first surveys the early development of mechanical theorem proving in geometry. GTP was initially viewed as an artificial intelligence problem that many believed would be easily tackled by machines. However, the initial optimism soon vanished as various difficulties associated with the domain emerged and no significant results could be proved. The developments over the last twenty years or so though have led to a revolution in the field. We survey several of the powerful methods proposed and analyse our own choice for formalization in Isabelle. Concepts that are important to deal with the geometry of the *Principia* are also introduced.

2.1 Historical Background

In 1899, Hilbert proposed five groups of axioms in his *Grundlagen der Geometrie* (Foundations of Geometry) [46]. In this classic work, Hilbert showed the consistency and independence of the sets of axioms and from them derived the various properties of plane (Euclidean) geometry. Hilbert's geometry consists of points, lines, and planes as primitives and of relations between these for angle congruence and incidence amongst others. The relationships between the primitives are completely determined by the axioms. Hilbert's insistence that no geometrical intuition was needed to prove any results – he suggested that the primitives could be replaced by chairs, tables, and beer mugs, as long as these satisfied the axioms – marked a clear departure from the geometry of Euclid. Geometry of the Ancient Greeks was meant as an axiomatization of concepts that were intuitively obvious, while Hilbert abstracted geometry away from any concrete interpretation.

The role of Hilbert's *Grundlagen* in relation to mechanical GTP is quite an important one. Indeed, the formalization given by Hilbert made clear for the first time the possibility of mechanizing elementary geometry. This was realized by Poincaré who, with great prescience, argued the following in his review of the *Grundlagen* (1902) [70]:

> Thus Hilbert has, so to speak, tried to put the axioms in such a form that they could be applied by someone who did not understand their meaning because he had never seen a point, a straight line, or a plane. Reasoning should, according to him, be capable of being carried out

> according to purely mechanical rules, and for doing geometry it suffices to apply these rules to the axioms slavishly without knowing what they mean. In this way one could build up all of geometry, I will not say without understanding anything at all since one must grasp the logical sequence of the propositions, but at least without perceiving anything. One could give the axioms to a logic machine, for example the *logical piano* of Stanley Jevons, and one would see all of geometry emerge from it.

Hilbert's geometry, despite its great influence, has been criticized though for giving the same status to points, lines, and planes. It can be contrasted with Tarski's theoretical contribution published in 1926 which is an axiom system for Euclidean geometry. In Tarski's geometry, the universe contains only points with two primitive relations on them: betweenness and equidistance. A direct consequence was that much more work had to be done to prove results in the Tarski system.

As far as actual machine geometry is concerned, the first geometry theorem prover (the Geometry Machine) was developed by Gelernter in 1959; it was then extended and used to prove a number of theorems from high-school textbooks of the time [37]. Gelernter's Geometry Machine included specific heuristic knowledge about the geometry domain and had a backward chaining search strategy. The main heuristic built into the machine was to use the diagram accompanying the statement of the geometry problem to reject false goal statements.

More work was also done using the same approach by Gilmore [38], Nevins [63] and Elcock [29], with additions such as forward chaining for example. The axiomatic or synthetic approaches used by Gelernter and the others above were, however, not very successful in proving or discovering any non-trivial theorems. The main problem, despite the numerous search strategies and heuristics that have been tried, is the high inefficiency; this is due to factors such as the huge search space of geometry rule applications. Koedinger argues that traditional geometry problem solving is hard [54] and outlines how the number of inferences that can be made rapidly increases at each layer in the proof (from seven at the beginning of the proof of a typical problem to over 100 000 at the third layer where a minimum of six layers are required). This makes it essential to add sophisticated search strategies and heuristic knowledge to GTP systems for them to have any chance of proving anything in the geometry domain. Koedinger further argues that the lack of success of these approaches lies in the fact that the underlying problem representation on which most of them are built has remained the same; namely one that has the formal geometry rules as operators and requires a search in the problem space for the rules that can be applied.

In 1969, Cerutti and Davis, rather ahead of their time, used symbolic manipulation in a system called FORMAC to prove theorems in elementary analytic geometry [15]. They used Descartes' method, that is an essentially algebraic approach that assigns coordinates to points, to prove Pappus' theorem. In the same paper, the authors also outlined how they obtained two