Automated Generation of Readable Proofs with Geometric Invariants[†]
II. Theorem Proving With Full-Angles

Shang-Ching Chou, Xiao-Shan Gao[‡] and Jing-Zhong Zhang[§]
Department of Computer Science, The Wichita State University
Wichita KS 67260-0083, USA
e-mail: (chou,gao,zhang)@cs.twsu.edu

Abstract. We present a set of rules based on full-angles as the basis of automated geometry theorem proving. We extend the idea of eliminating variables and points to the idea of eliminating lines. We also discuss how to combine the forward chaining and backward chaining to achieve higher efficiency. The prover based on the full-angle method has been used to produce short and elegant proofs for more than one hundred difficult geometry theorems. The proofs of many of those theorems produced by our previous area method are relatively long.

Keywords. Automated reasoning, automated geometry theorem proving, method based on angle, forward chaining, backward chaining.

1 Introduction

In part II of this series, we continue developing the approach to prove geometry theorems with geometric invariants. We introduce another important geometry notion – the full-angle which is explicitly used for proving geometry theorems for the first time. In terms of methodology, this is a further development of our previous work based on the area method. Our experience shows that, the computer program based on full-angles can produce elegant proofs for many extremely difficult geometry theorems. Furthermore, the proofs produced with full-angles are more like the way people solving geometry theorems.

[†]This work was supported in part by the NSF Grants CCR-9117870, CCR-9420857 and the Chinese NSF.

[‡]On leave from Institute of Systems Sciences, Academia Sinica, Beijing 100080.

[§]Permanent address: Chengdu Institute of Computer Application, Academia Sinica, 610015 Chengdu, China

[4] is a collection of 110 geometry theorems and proofs produced automatically by our prover based on full-angles. It is the case that by combining the area method with the method based on full-angles, our prover can produce short and readable proofs for most of the geometry theorems (not including inequalities) in Euclidean geometry. Examples listed in the appendix of the paper include some from the International Mathematics Olympiad and recently proposed problems in the American Mathematics Monthly (the solutions of some of those problems have not been published yet). Our machine proofs of these theorems are very short and elegant. We are especially pleased with the elegance of the proof of the "five circle" theorem (Example 3.2). The theorem was proposed in the net news group sci.math by Noam D. Elkies of Harvard University as a challenge problem in 1992.

In the method based on full-angles, we extend the idea of eliminating variables and points to the idea of eliminating lines. We will also discuss how to combine the forward chaining and the backward chaining to achieve higher efficiency and greater power for the prover.

Although the method based on full-angles alone is not complete, we present a complete (decision) method for constructive problems involving full-angles. This complete method is a combination of the area method with the full-angle method. So we consider the full-angle method as a complement to the area method. For a constructive geometry statement, our program first tries the full-angle method. If the full-angle method successes then we satisfy with it since proofs based on full-angle are always short; otherwise the program uses the complete method to produce a proof.

2 Rules Based on Full-Angles

2.1 Why Full-Angle

Not all proofs generated by the area method are short. There are two main factors in the description of the geometry statements that affect the length of a machine proof: the *number of points* in the statement and the *type of predicates* needed to describe it. Generally speaking, the number of points in a geometry statement is fixed and reflects the complexity of the statement. On the other hand, the difficulty related to the type of predicates is due to the method. According to our experience, the geometry relations can be listed in ascending order of difficulties as follows: collinear, parallel, ratios, perpendicular, circles, angles. This means that the area method works better for geometry theorems about collinear and parallel and is relatively ineffective for theorems involve angles.

On the other hand, angle is one of most powerful notion used in theorem proving in the traditional proof method. The use of angle is one of the main reasons that traditional geometric proofs are often very elegant, skillful and interesting. But the traditional angle is a concept involving the the order relation in geometry, and is thus very difficult to fit into our mechanic proof system. For a detailed discussion about the order relation, see Section 5.2 in Part I of the paper.

Example 2.1 Two circles O and Q meet in two points A and B. A line passing through A meets circles O and Q in C and E. A line passing through B meets circles O and Q in D and F. Show that $CD \parallel EF$.

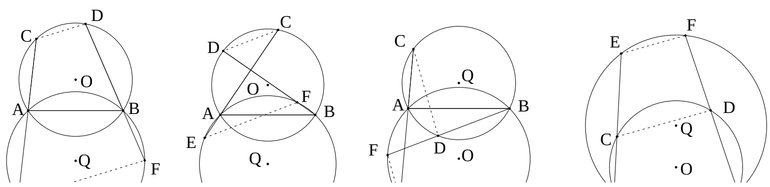


Figure 1 shows four possible diagrams for this example. If we use the traditional angle to prove this theorem, in each case we have to give different proofs. This makes it very difficult to develop a mechanical method based on the traditional angle.

To overcome this difficulty, we will use the concept of full-angle as a new notion to prove geometry statements. The concept of full-angle was explicitly used by Wu to express the predicate of angle congruence as an algebraic equation [10].

2.2 Basic Rules

In this paper, points are represented by capital English letters, and lines are represented by lowercase English letters or two distinct points on it.

Intuitively, a full-angle $\angle[u, v]$ is the angle from line u to line v. Note that u and v are not rays as in the definition for the ordinary angles. Two full-angles $\angle[l, m]$ and $\angle[u, v]$ are equal if there exists a rotation K such that $K(l) \parallel u$ and $K(m) \parallel v$. For the geometric meaning of the addition of full-angles, let l, m, u, and v be four lines and K be a rotation such that $K(l) \parallel v$. Then $\angle[u, v] + \angle[l, m] = \angle[u, K(m)]$.

Formally, full-angles can also be defined using the signed area and the Pythagoras difference. For details, see Section 4. But here, full-angle is defined as an ordered pair of lines which satisfies the following rules.

R1 For all parallel lines $AB \parallel PQ$, $\angle[0] = \angle[AB, PQ]$ is a constant.

R2 For all perpendicular lines $AB \perp PQ$, $\angle[1] = \angle[AB, PQ]$ is a constant.

R3 There is an operation "addition" between two full-angles which is commutative and associative.

R4 $\angle[1] + \angle[1] = \angle[0]$.

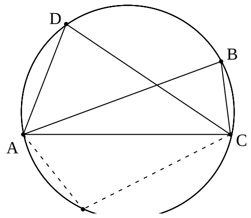
R5 $\angle[u, v] + \angle[0] = \angle[u, v]$.

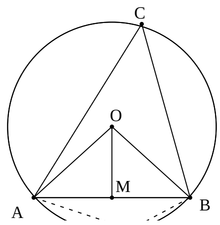
R6 If X is on line PQ, then $\angle[AB, PX] = \angle[AB, PQ]$.

R7 If PX is parallel to UV, then $\angle[AB, PX] = \angle[AB, UV]$.

R8 If PX is perpendicular to UV, then $\angle[AB, PX] = \angle[1] + \angle[AB, UV]$.

R9 If $XA = XB$ then $\angle[AX, AB] = \angle[AB, XB]$.





R10 (The Inscribed Angle Theorem) If A, B, C, and D are cyclic then $\angle[AD, CD] = \angle[AB, CB]$ (Figure 2).

R11 If O is the circumcenter of triangle ABC and M is the midpoint of AB then $\angle[AO, OM] = \angle[AC, BC]$ (Figure 3).

R12 If $MA = MB$ and A, B, P, M are cyclic then $\angle[PA, PM] = \angle[PM, PB]$.

R13 $\angle[AB, CD] = -\angle[CD, AB]$.

R14 For any line UV, $\angle[AB, CD] = \angle[AB, UV] + \angle[UV, CD]$.

Note that rule R6 is to *eliminate a point* (X) from a full-angle. Rules R7-R12 are to *eliminate lines* from the full-angles. Rule R13 ensures that the two lines in a full-angle will both be dealt with. R14 is a very special rule: if used like $\angle[AB, UV] + \angle[UV, CD] = \angle[AB, CD]$ then it eliminates line UV from the full-angles; if used in the other direction it adds a line to the full-angle $\angle[AB, CD]$. When no new results can be obtained by applying rules R1–R13 to a full-angle α, we use R14 to split α to two new full-angles by adding a line and the elimination process might continue with these two new full-angles.

Remark 2.2 The above properties already show the advantage of using full-angles. In R10, if using the traditional angle, we need two conditions (Figure 2): $\angle ADC = \angle ABC$ or $\angle ADC + \angle AD_1C = 180°$ and to distinguish these two cases, we need inequalities. This is the reason why the method based on full-angles can generate diagram independent proofs. Similarly, for rule R11 there are also two cases: $\angle AOM = \angle ACB$ or $\angle AOM + \angle AC_1B = 180°$ if the traditional angle is used.

The three control strategies introduced in Part I of this series will still be used here. To use Control strategy 1, we need to define a rank or order among the full-angles. Suppose that we have an order among the points, denoted by $<$. Then a full-angle can be represented canonically. A full-angle $\angle[AB, CD]$ is said to be in the *canonical form* if $A \geq B$, $C \geq D$, and ($A > C$ or ($A = C$ and $B > D$)). In what follows, all full-angles are assumed to be in the canonical form.

Definition 2.3 A full-angle $\angle[AB, CD]$ has lower rank than $\angle[PQ, RS]$, denoted by $\angle[AB, CD] < \angle[PQ, RS]$, if (1) $A < P$, or (2) $A = P$ and $B < Q$, or (3) $A = P, B = Q$, and $C < R$, or (4) $A = P, B = Q$, $C = R$, and $D < S$.

We use a new control strategy.

Control Strategy 4. Rule R14 can be used to add a line only if the two new full-angles can be further reduced to angles of lower rank.

2.3 Combined Rules

For Strategy 4, we need to remember the elimination results for the next two steps to see whether rule R14 is usable. In our prover, we actually builtin many frequently-used cases to simplify the search process. These builtin cases are one of the key factors for the search efficiency of our prover.

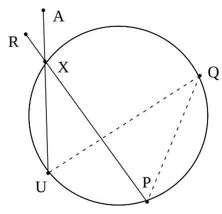
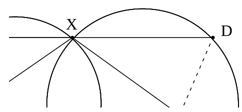
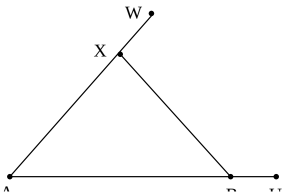


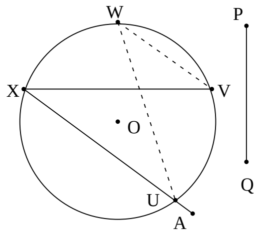
Figure 4



Rule R15. If coll(A, X, U), coll(P, X, R), and cyclic(X, P, Q, U) (Figure 4), then by R14, R6, and R10 $\angle[AX, BC] = \angle[AX, XP] + \angle[XP, BC] = \angle[QU, QP] + \angle[PR, BC]$.

Rule R16. If coll(X, C, D), coll(X, U, A), coll(X, V, B), cyclic(X, U, C, E), and cyclic(X, V, D, F) (Figure 5), then by R14, R6, and R10 $\angle[AX, BX] = \angle[AX, XC] + \angle[XD, XB] = \angle[EU, EC] + \angle[FD, FV]$. This rule includes an interesting special case: if $C = D$ then XC becomes the common chord of the two circles.





Rule R17. If $XA = XB$ and $coll(X, A, W)$ (Figure 6) then by R14 and R9 $\angle[AX, BX] = \angle[AX, AB] + \angle[AB, XB] = 2\angle[AX, AB] = 2\angle[AW, AB]$. Here W could be the same as X.
Rule R18. If $XA = XB$, $coll(X, A, W)$, and $coll(A, B, U)$ (Figure 6) then by R14, R6 and R9 $\angle[BX, CD] = \angle[BX, AB] + \angle[AB, CD] = \angle[AU, AW] + \angle[AU, CD]$.
Rule R19. If circumcenter(O, A, B, C) and midpoint(M, A, B) (Figure 3) then by R14, R6 and R11 $\angle[AO, AB] = \angle[AO, OM] + \angle[OM, AB] = \angle[AC, CB] + \angle[1]$. The above rule can be used to eliminate line OA or line AC according to the rank of the full-angles.

Rules R15-R19 are actually combinations of rule R6 and rules R9, R10, R11 using rule R14, i.e., we need to find a new line such that when we split a full-angle into two using rule R14, at least one of the new full-angles can be reduced to lower rank by rules R9, R10, or R11. Similarly, we can combine each pair of the rules R6–R12 to obtain new elimination rules. As an example, the following is a combination of R8 and R10.

Rule R20. If cyclic(X, W, U, V), coll(X, U, A), and perp(X, V, P, Q) (Figure 7) then $\angle[AX, BC] = \angle[AX, VX] + \angle[VX, BC] = \angle[WU, WV] + \angle[PQ, BC] + \angle[1]$.

Rule R21. If incenter(I, A, B, C) then we have $\angle[AI, AB] + \angle[BI, BC] + \angle[CI, CA] = \angle[1]$. There are twelve angles in this configuration among which there exist two independent ones and other full-angles can be expressed by the two angles. We can choose the two independent full-angles according to the order of the points. For instance, if $A < B < C < I$, a set of independent angle are $\angle[AI, AB]$ and $\angle[BI, BC]$.

2.4 The Prover

The prover based on full-angles is just an extension of the prover reported in [3] (part I) by adding the rules and control strategies given in this section. Thus the prover uses a backward chaining search method.

In the case of full-angles, the algebraic computation used in the proving process is much simpler, because we do not multiply two full-angles. Some special property of full-angles should be taken into account. First, $2\angle[1] = \angle[0]$ meaning that the addition of two right angles is a flat angle. Second, no division is allowed, e.g., from $2\alpha_1 = 2\alpha_2$ we have two possible conclusions $\alpha_1 = \alpha_2$ or $\alpha_1 = \alpha_2 + \angle[1]$.

Example 2.4 The following is a machine proof for Example 2.1 produced by our prover *automatically.* Note that this proof is valid for all the cases in Figure 1.

The Machine Proof (Point order: $A, B, C, D, E, F.$)$\angle[FE, DC]$

$(\angle[FE, DC] = \angle[FE, FB] + \angle[BD, DC] = \angle[EA, BA] - \angle[DC, DB]$,
because collinear(B,D,F), cyclic(F,E,B,A). (R15))

$= -\angle[EA, BA] + \angle[DC, DB]$

$(\angle[EA, BA] = \angle[DC, DB]$, because collinear($A$,$C$,$E$), cyclic($A$,$C$,$B$,$D$). (R10))

$= \angle[0]$

For each full-angle in the above proof, say $\angle[FE, DC]$, we need to assume that the two lines defining the full-angle are proper, i.e. $E \neq F$ and $C \neq D$. In all the machine proofs based on full-angles, we will implicitly make this assumption.

3 Geometry Information Bases and Forward Chaining

As we mentioned in Section 2.4, we use a backward chaining search strategy in the prover. But in the case of full-angles, we also use a forward chaining in a restricted fashion.

3.1 The Geometry Information Base and Forward Chaining

To use each of the rules presented in the preceding section, some conditions, like several points are cyclic or collinear, should be satisfied. This kind of information usually comes from the hypotheses of the geometry statement. However, there are some facts that are not in the hypotheses explicitly, but can be deduced from the hypotheses directly. For example, if $coll(A, B, C)$, $coll(A, B, D)$ and $A \neq B$ then $coll(B, C, D)$. To organize this kind of information properly affects not only the time efficiency but also the power of the prover.

In our prover, before proving a theorem we first collect all the obvious geometric information into a *geometry information basis* (abbr. GIB). This GIB will be used to provide information for the elimination rules. First, the hypotheses of the proposition should be put into the GIB. Then the prover will keep applying the following rules to all the information in GIB to get new information and to put the new information into GIB until nothing new can be obtained.

F1 If $l_1 \parallel l_2$ and $l_1 \parallel l_3$ then $l_2 \parallel l_3$.

F2 If $l_1 \perp l_2$ then $l_1 \perp l_3$ iff $l_2 \parallel l_3$.

F3 AB is the mediator of XY if and only if $AX = AY$ and $BX = BY$.

F4 If midpoint(O, C, A) then $AB \perp BC$ iff circumcenter(O, A, B, C).

F5 If $PA \perp PB$ then $QA \perp QB$ iff *cyclic*(A, B, P, Q).

F6 If M and N are the midpoints of AB and AC then $MN \parallel BC$.

F7 If $\angle[OU, OP] = \angle[OP, OV] \neq \angle[0]$, $OU \perp PU$, and $OV \perp PV$ then $OU = OV$ and $PU = PV$.

F8 If $AB \parallel AC$ then coll(A, B, C).

Note that some of the rules can be used in two directions. For example, F5 can be used in two ways: if $PA \perp PB$ and $QA \perp QB$ then *cyclic*(A, B, P, Q); if $PA \perp PB$ and *cyclic*(A, B, P, Q) then $QA \perp QB$.

The reader might have noticed that the GIB is built by doing a special kind of *forward chaining*. The purpose of this forward chaining is not to prove geometry theorems (though for some simple theorems the forward chaining can provide a proof) but to provide information for further backward chaining. We thus only use those simple rules and leave the difficult rules to the *backward chaining*, i.e., the elimination procedure we have discussed in the preceding section. In the forward chaining, we need to find and store all the possible geometry results. Therefore, the forward chaining is usually much time consuming than the backward chaining. The aim of our prover is to achieve higher efficiency and greater power by adopting the combination of the forward chaining and backward chaining with emphasis on the backward chaining.

A. Nevins was the first to recognize the importance of the forward chaining in automated geometry theorem proving [8]. Similar ideas were also used in [5].

Unlike backward chaining which can be automatically done by prolog, we have to write our forward chaining program. Since only some simple rules are used, the forward chaining is generally efficient.

Example 3.1 (Simson's Theorem)
Let D be a point on the circumcircle of triangle ABC. From D three perpendiculars are drawn to the three sides BC, AC, and AB of triangle ABC. Let E, F, and G be the three feet respectively. Show that E, F and G are collinear.

Besides the hypotheses, the GIB contains the following three circles.

cyclic(E, B, D, G), because $EB \perp ED$, $GB \perp GD$; (F5)
cyclic(E, C, D, F), because $EC \perp ED$, $FC \perp FD$; (F5)
cyclic(F, A, D, G), because $FA \perp FD$, $GA \perp GD$. (F5)

The Machine Proof (Point order: O, A, B, C, D, E, F, G.)
$\angle[GF, GE]$
($\angle[GF, GE] = \angle[GF, GD] + \angle[GD, GE] = \angle[FA, DA] - \angle[EB, DB]$,
because cyclic(F,G,D,A), cyclic(E,G,D,B). (R16))
$= \angle[FA, DA] - \angle[EB, DB]$
($\angle[FA, DA] = -\angle[DA, CA]$, because collinear($A$,$C$,$F$). (R6))
$= -\angle[EB, DB] - \angle[DA, CA]$
($\angle[EB, DB] = -\angle[DA, CA]$, because collinear($B$,$C$,$E$), cyclic($B$,$C$,$D$,$A$). (R10))
$= \angle[0]$

3.2 Forward Chaining Using Full-Angles

For some geometry statements, the GIB generated as in the preceding subsection does not contain enough information for the prover to produce a proof. In that case, we will obtain a larger GIB by doing more forward chaining. The main purpose of doing this is finding new circles in the statements. The following new rules are used to obtain equal angles from other geometry relations. For some rules, we need to add auxiliary conditions which are called *the non-degenerate (ndg) conditions* for these rules.

K1 $l_1 \parallel l_2$ iff there is another line l_3 such that $\angle[l_1, l_3] = \angle[l_2, l_3]$.

K2 If $l_1 \perp l_2$ and $l_3 \perp l_4$ then $\angle[l_1, l_3] = \angle[l_2, l_4]$.

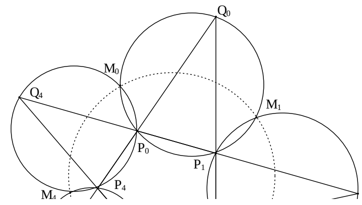
K3 A, B, C, and D are cyclic iff $\angle[AB, AC] = \angle[DB, DC]$. (The non-degenerate condition is $\neg$coll(A, B, C, D).)

K4 $OA = OB$ iff $\angle[OA, AB] = \angle[AB, OB]$. (The non-degenerate condition is $\neg$ coll(O, A, B).)

K5 If $MA = MB$ and A, B, P, M are cyclic then $\angle[PA, PM] = \angle[PM, PB]$.

There are two steps in the forward chaining involving equal full-angles. (1) Find all equal full-angles using rules K1–K5. (2) For any two equal full-angles, use the reverse of the rules K1–K5 to obtain new circles, new parallel lines, and new perpendicular lines.

We separate the forward chaining into two steps, because collecting equal angles is usually very time consuming and the proofs of most of the geometry theorems do not need this step. Also the geometric conditions obtained in this way may not obvious any more. Some of them may need detailed explanation as shown by Example 3.2. But as we mentioned before, this step is necessary to obtain proofs for some geometry theorems. Seven of the 110 geometry theorems in [4] need this step. Therefore, we use the following strategy: only if the prover fails to find a proof for a statement, this step will be used.



Example 3.2 (The Five Circle Theorem[1])
As in Figure 9, $P_0P_1P_2P_3P_4$ is a pentagon. $Q_i = P_{i-1}P_i \cap P_{i+1}P_{i+2}$, $M_i = circle(Q_{i-1}P_{i-1}P_i) \cap circle(Q_iP_iP_{i+1})$ (the subscripts are understood to be mod 5). Show that points M_0, M_1, M_2, M_3, M_4 are cyclic.

Besides the fives circles in the hypotheses, the prover finds five new circles: cyclic(M_3, M_0, P_4, Q_0, Q_2), cyclic(M_1, M_4, P_0, Q_1, Q_3), cyclic(M_1, M_3, P_2, Q_0, Q_3), cyclic(M_2, M_4, P_3, Q_1, Q_4), cyclic(M_2, M_0, P_1, Q_2, Q_4). The first of the five circles can be derived as follows.

(1). cyclic(M_3, M_0, P_4, Q_0, Q_2), because of (2) and (6).
(2). cyclic(M_3, P_4, Q_0, Q_2),because of (3).
(3). $\angle[P_4M_3, P_4Q_0] = \angle[Q_2M_3, Q_2Q_0]$, because of (4) and (5).
(4). $\angle[P_4M_3, P_4Q_0] = \angle[P_3M_3, P_3P_2]$, because of cyclic($P_3, P_4, Q_3, M_3$).
(5). $\angle[Q_2M_3, Q_2Q_0] = \angle[P_3M_3, P_3P_2]$, because of cyclic($P_2, P_3, Q_2, M_3$).
(6). cyclic(M_0, P_4, Q_0, Q_2), because of (7).
$\cdots$.

After building the GIB, the prover can easily give the following proof for cyclic(M_0, M_1, M_2, M_4).

The Machine Proof

$\angle[M_4M_1, M_4M_0] - \angle[M_2M_1, M_2M_0]$

($\angle[M_4M_1, M_4M_0] = \angle[M_4M_1, M_4P_0] + \angle[M_4P_0, M_4M_0] = \angle[M_1Q_1, Q_1P_0] - \angle[M_0P_4, P_4P_0]$, because cyclic($M_1$,$M_4$,$P_0$,$Q_1$), cyclic($M_0$,$M_4$,$P_0$,$P_4$). (R16))

$= -\angle[M_2M_1, M_2M_0] + \angle[M_1Q_1, Q_1P_0] - \angle[M_0P_4, P_4P_0]$

($\angle[M_2M_1, M_2M_0] = \angle[M_2M_1, M_2P_1] + \angle[M_2P_1, M_2M_0] = \angle[M_1P_2, P_2P_1] - \angle[M_0Q_4, Q_4P_1]$, because cyclic($M_1$,$M_2$,$P_1$,$P_2$), cyclic($M_0$,$M_2$,$P_1$,$Q_4$). (R16))

$= \angle[M_1Q_1, Q_1P_0] - \angle[M_1P_2, P_2P_1] + \angle[M_0Q_4, Q_4P_1] - \angle[M_0P_4, P_4P_0]$

($\angle[M_1Q_1, Q_1P_0] = \angle[M_1P_2, P_2P_1]$, because collinear($P_0$,$P_1$,$Q_1$), cyclic($Q_1$,$M_1$,$P_1$,$P_2$). (R10))

[1] We were informed of this problem by K. Abdali. It was proposed in the news group sci.math by Noam D. Elkies of Harvard University. The theorem was proved by Gerald A. Edgar of Ohio State University with Maple. However, for the general-purpose geometry theorem provers based on the algebraic methods, the proofs require exceedingly large amount of computer memory which are currently not available on most computer systems. Wen-Tsün Wu was later able to give a simple synthetic proof. Now an elegant proof for this difficult theorem can be produced totally automatically by our prover.

$= \angle[M_0Q_4, Q_4P_1] - \angle[M_0P_4, P_4P_0]$

$(\angle[M_0Q_4, Q_4P_1] = \angle[M_0P_4, P_4P_0]$, because collinear$(P_0,P_1,Q_4)$, cyclic$(Q_4,M_0,P_0,P_4)$. (R10))

$= \angle[0]$

4 A Complete Method for Full-angles

We will present a complete method for solving constructive problems involving full-angels using the area method. First, we need a formal definition of full-angles using the signed area and the Pythagoras difference[2]

Definition 4.1 An ordered pair of lines AB and CD determines a *full-angle*, denoted by $\angle[AB, CD]$, which satisfies the following properties.

1. $\angle[AB, CD] = \angle[PQ, UV]$ if and only if $S_{ACBD}P_{PUQV} = S_{PUQV}P_{ACBD}$ where S_{ACBD} and P_{ACBD} are the *signed area* and *Pythagoras difference* of the quadrilateral $ABCD$ [2] respectively. Thus the *tangent function* for the full-angle,

$$\tan(\angle[AB, CD]) = \frac{4S_{ACBD}}{P_{ADBC}}$$

   is a well defined geometry quantity.

2. For all parallel lines $AB \parallel PQ$, $\angle[0] = \angle[AB, PQ]$ is a constant.

3. For all perpendicular lines $AB \perp PQ$, $\angle[1] = \angle[AB, PQ]$ is a constant.

4. There exists an operation "+" for full-angles such that

   - $\angle[1] + \angle[1] = \angle[0]$;
   - the tangent function of the sum of two full-angles is defined as follows

$$\tan(\angle[AB, CD] + \angle[PQ, UV]) = \frac{\tan(\angle[AB, CD]) + \tan(\angle[PQ, UV])}{1 - \tan(\angle[AB, CD]) \tan(\angle[PQ, UV])}.$$

All the rules in Section 2 can be derived from the above definition. Actually some of them can be proved automatically with the following explanation.

With the concept of full-angles, the constructive geometry statements (see [2]) can be extended as follows. First the conclusion of a geometry statement could be an equation of full-angles. The second extension is more interesting: we can introduce a new kind of straight lines.

[2] In this section, we assume the reader is familiar with signed areas, Pythagoras differences, and the complete method for constructive geometry statements [2].

(ALINE P Q U W V) which is the line l passing through P such that $\angle[PQ, l] = \angle[UW, WV]$. We assume that $P \neq Q$, $U \neq W$, and $W \neq V$.

With this new type of lines, we can introduce new constructions such as taking the intersection of an ALINE with another line or a circle. To provide methods of eliminating points introduced by these new constructions from area and Pythagoras difference, we need only to reduce an ALINE to an ordinary line. To do that, we need a construction used in [2]:

(TRATIO R Q P r) introduce a point R such that $RQ \perp QP$ and $(RQ/QP =)\frac{4S_{RQP}}{P_{QPQ}} = r$ where r could be a number, a variable or an expression in geometry quantities.

Proposition 4.2 If UW is not perpendicular to WV, line l =(ALINE P Q U W V) is the same as line PR where R is introduced by construction (TRATIO R Q P $\frac{4S_{UWV}}{P_{UWV}}$).

Proof. Let the line passing through point Q and perpendicular to PQ meet line l in R. Then R is introduced by construction (TRATIO R Q P r), where

$$r = \frac{4S_{RQP}}{P_{QPQ}} = \frac{4S_{QPR}}{P_{QPR}} = \tan(\angle[RPQ]) = \tan(\angle[VWU]) = \frac{4S_{UWV}}{P_{UWV}}. \quad \blacksquare$$

Theorem 4.3 We have a complete method of proving theorems in the class of constructive geometry statements involving full-angles.

Proof. If the conclusion of a geometry statement is an expression $\sum_{i=1}^{k} n_i\alpha_i = 0$ where n_i are integers and α_i are full-angles, then we can prove the equivalent result $\tan(n_1\alpha_1) = \tan(\sum_{i=2}^{k} n_i\alpha_i)$ which can be represented by area and Pythagoras difference. Hence the area method in [2] can be used to eliminate points from it. If the constructive description of the statement needs ALINE, then by Proposition 4.2 we can always describe the statement constructively without using the notion of ALINE. Thus a constructive geometry statement involving full-angles actually belongs to the class of constructive statements defined in [2] and hence can always be proved or disproved with the area method. ∎

For a geometry statement involving full-angles, the proof produced with the area method as described in the proof of Theorem 4.3 is generally longer than the proof produced with the full-angle method. For Example 3.2, the area method even fails to produce a proof due to the computer memory limit. This is why we introduce the incomplete full-angle method as a complement to the area method.

5 Possible Future Refinements

5.1 The Search for More Geometric Notions

The geometric notions with related high level geometry lemmas play a decisive role in producing short and readable proofs of difficult geometry theorems. We have successfully used geometric quantities such as the signed area, the Pythagoras difference, the signed volume, the full-angle and their properties as basic notions. Other promising notions include trigonometric functions and similar triangles. Note that some of the commonly used geometry quantities such as the length of a segment the absolute area are not used by us. The reason is that it is difficult to produce diagram independent proofs based on this kind of geometry quantities. But it is still worth exploring how to use them to prove theorems automatically, since they are easier to understand.

At the present time, our method can only deal with geometry theorems not involving inequality. It is an interesting research topic to find proper notions that can generate human understandable proofs for geometry inequalities.

Another interesting direction is to extend the methods based on geometric invariants to non-Euclidean geometries, higher dimensional geometries, and differential geometries.

5.2 Database and Forward Chaining

At the present time, we only use some very simple rules in the forward chaining. To use complicated rules to do forward chaining is generally very time consuming. But it is still worth exploring the full power of forward chaining, since a successful forward chaining method is more powerful than the backward chaining and can be used to discover new theorems. This has been shown by Example 3.2. The difference in power between the backward chaining and the forward chaining used by us is caused by the using of rank among full-angles and control strategy 1 in part I of this series.

Our current GIB can be extended in the following three steps. (1) Find all the equal full-angles in the geometry statement. At the present time, the GIB only contains those equal full-angles obtained without using algebraic computation. (2) Find all the linear relations among the full-angles in the geometry statements. (3) Find all the geometry relations like collinear, parallel, and perpendicular among the points in the geometry statement using the area method. To build a larger GIB, the techniques developed in the field of deductive database [1] may be useful.

5.3 The Combination of the GIB and the Area Method

At the present time, the GIB is only used by the method based on full-angles. We believe a combination of the GIB and the area method could reduce the lengths of many proofs. For example, from the following result (which is easy to prove) we see that from information about full-angles, areas and Pythagoras differences could be simplified.

Proposition 5.1 (Co-angle Theorem) If $\angle[AB, BC] = \angle[XY, YZ]$, $\angle[AB, BC] \neq \angle[1]$, and $\angle[AB, BC] \neq \angle[0]$, then $\frac{S_{ABC}}{S_{XYZ}} = \frac{P_{ABC}}{P_{XYZ}} = \lambda$ where $\lambda^2 = \frac{AB^2 \cdot BC^2}{XY^2 \cdot ZY^2}$.

6 Experiment Results

The prover is implemented using the SB-Prolog on a SPARC-10 Workstation. The following table contains some statistics for Examples 2.1,3.1,3.2, and the examples in the Appendix.

We include six indexes for each theorem. Steps is the length or steps of the shortest proof obtained. Maxt is the number of full-angles occurring in the maximal algebraic formula in the shortest proof. Ftime is the time needed to obtain the first proof using the depth-first search. Mtime is the time needed to obtain the shortest proof using the depth-first search. Btime is the time needed to obtain the first proof using the depth-first iterative deepening search [9]. Allprs is the number of all the proofs generated using depth-first search.

Examples	steps	maxt	ftime (secs)	mtime (secs)	btime (secs)	allprs
2.1	4	2	0.17	0.28	0.51	4
3.1	5	2	0.32	1.26	1.5	2572
3.2	5	4	6.91	228.8	266.9	27268*
1	4	3	0.16	0.58	0.46	2
2	4	3	0.32	27.83	4.66	3772
3	6	3	2.11	49.05	57.2	27583*
4	6	4	0.71	9.32	1.72	14
5	4	3	1.15	2.56	4.34	3
6	6	2	3.06	1671.2	304.1	179*

The * in the table means that the program does not finish after running more than ten hours.

The following table gives the average performance of the prover for the 110 in [4].

110 theorems	steps	maxt	ftime (secs)
[4]	6.45	3.06	1.66

References

[1] F. Bancilhon & R. Ramakrishnan, An Amateur's Introduction to Recursive Query Processing Strategies, *Proc. of ACM SIGMOD Conference*, edited by C. Zanioilo., p.16-52, 1986.

[2] S. C. Chou, X. S. Gao, & J. Z. Zhang, *Machine Proofs in Geometry*, World Scientific, 1994.

[3] S. C. Chou, X. S. Gao, & J. Z. Zhang, Automated Generation of of Readable Proofs with Geometric Invariants, I. Multiple and Shortest Proof Generation, the same volume.

[4] S. C. Chou, X. S. Gao, & J. Z. Zhang, A Collection of 110 Geometry Theorems and Their Machine Proofs Based on Full-Angles, TR-94-4, CS Dept., WSU, November, 1994.

[5] H. Coelho & L. M. Pereira, Automated Reasoning in Geometry Theorem Proving with Prolog, *J. of Automated Reasoning*, vol. 2, p. 329-390, 1986.

[6] H. Gerlentner, J.R. Hanson, and D.W. Loveland, Empirical Explorations of the Geometry-theorem Proving Machine, *Proc. West. Joint Computer Conf.*, 143-147, 1960.

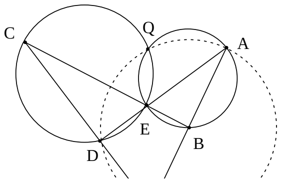
[7] P. C. Gilmore, An Examination of the Geometry Theorem Proving Machine, *Artificial Intelligence*, 1, p. 171–187.

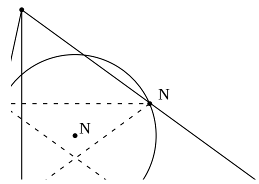
[8] A.J. Nevins, Plane Geometry Theorem Proving Using Forward Chaining, *Artificial Intelligence*, 6, 1-23.

[9] M. Stickel, A Prolog Technology Theorem Prover: Implementation by an Extended Prolog Compiler, *J. of Automated Reasoning*, 4(4), 353–380, 1985.

[10] Wu Wen-tsün, *Basic Principles of Mechanical Theorem Proving in Geometries*, Volume I: Part of Elementary Geometries, Science Press, Beijing (in Chinese), 1984.

Appendix. Examples and Their Machine Produced Proofs





Example 1 (Miquel Point Theorem) Four lines form four triangles. Show that the circumcircles of the four triangles passes through a common point.

Point order: A, B, C, D, E, Q, P.

Hypotheses: cyclic(A, B, E, Q), cyclic(C, D, E, Q), coll(A, B, P), coll(C, D, P), coll(B, C, E), coll(A, D, E).

Conclusion: cyclic(D, A, P, Q).

The Machine Proof

$\angle[PD, PA] - \angle[QD, QA]$

($\angle[PD, PA] = \angle[DC, BA]$, because collinear$(C,D,P)$, collinear$(A,B,P)$. (R6))

$= -\angle[QD, QA] + \angle[DC, BA]$

($\angle[QD, QA] = \angle[QD, QE] + \angle[QE, QA] = -\angle[EC, DC] + \angle[EB, BA]$, because cyclic$(D,Q,E,C)$, cyclic$(A,Q,E,B)$. (R16))

$= \angle[EC, DC] - \angle[EB, BA] + \angle[DC, BA]$

($\angle[EC, DC] - \angle[EB, BA] = -\angle[DC, BA]$, because collinear$(B,C,E)$. (R14))

$= \angle[0]$

Example 2 (Nine Point Circle Theorem) Let the midpoints of the sides AB, BC, and CA of $\triangle ABC$ be L, M, and N, and AD the altitude on BC. Show that L, M, N, and D are on the same circle.

Point order: A, B, C, D, L, M, N.

Hypotheses: midpoint(M, B, C), midpoint(N, A, C), midpoint(L, A, B), foot(D, A, B, C).

Conclusion: cyclic(L, D, M, N).

The Machine Proof

$-\angle[NL, ND] + \angle[ML, MD]$

($\angle[NL, ND] = -\angle[ND, CB]$, because $NL \parallel BC$. (R7))

$= \angle[ND, CB] + \angle[ML, MD]$

$(\angle[ND, CB] = \angle[ND, DC] = \angle[DA, CA] + \angle[1]$, because circumcenter($N$,$D$,$C$,$A$). (R19))

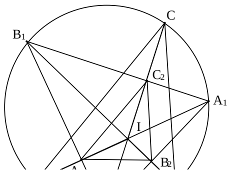
$= \angle[ML, MD] + \angle[DA, CA] + \angle[1]$

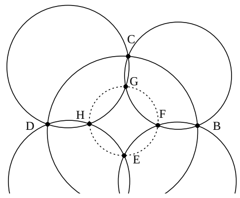
$(\angle[ML, MD] + \angle[DA, CA] = -\angle[MD, DA]$, because $ML \parallel CA$. (R14))

$= -\angle[MD, DA] + \angle[1]$

$(\angle[MD, DA] = \angle[1]$, because $MD \perp DA$. (R2))

$= \angle[0]$





Example 3 [3] Let ABC be inscribed in a circle and let A_1, B_1, and C_1 be the midpoints of the arc BC, CA, and AB respectively. Show that the pedal triangle of triangle $A_1B_1C_1$ is homothetic to triangle ABC.

Point order: $A, B, C, A_1, B_1, C_1, A_2, B_2, C_2$.

Hypotheses: incenter(I, A, B, C), cyclic(A, B, C, A_1, B_1, C_1), coll(A_1, A, I), coll(B_1, B, I), coll(C_1, C, I), foot(A_2, A_1, B_1, C_1), foot(B_2, B_1, A_1, C_1), foot(C_2, C_1, A_1, B_1).

Conclusion: para(A, B, A_2, B_2).

The Machine Proof

$-\angle[B_2A_2, BA]$

$(\angle[B_2A_2, BA] = \angle[B_2A_2, B_2A_1] + \angle[A_1C_1, BA] = \angle[A_2B_1, B_1A_1] + \angle[C_1A_1, BA]$,
because collinear(A_1,B_2,C_1), cyclic(B_2,A_2,A_1,B_1). (R15))

$= -\angle[A_2B_1, B_1A_1] - \angle[C_1A_1, BA]$

$(\angle[A_2B_1, B_1A_1] = \angle[C_1B, A_1B]$, because collinear($A_2$,$B_1$,$C_1$), cyclic($B_1$,$C_1$,$A_1$,$B$). (R10))

$= -\angle[C_1A_1, BA] - \angle[C_1B, A_1B]$

$(\angle[C_1B, A_1B] = \angle[C_1B, BA] + \angle[BA, A_1B] = -\angle[C_1A, BA] - \angle[A_1B, BA]$,
because $C_1B = C_1A$. (R17))

$= -\angle[C_1A_1, BA] + \angle[C_1A, BA] + \angle[A_1B, BA]$

$(\angle[C_1A_1, BA] = \angle[C_1A_1, C_1A] + \angle[AC_1, BA] = \angle[C_1A, BA] + \angle[A_1B, BA]$,
because cyclic(C_1,A_1,A,B). (R15))

$= \angle[0]$

[3] This example is a problem proposed in American Mathematics Monthly, 1993 (Problem 10317)

Example 4 (Miquel's Theorem) If four circles are arranged in sequence, each two successive circles intersecting, and a circle pass through one point of each pair of intersection, then the remaining intersections lie on another circle.

Point order: A, B, C, D, E, G, F, H.
Hypotheses: cyclic(A, B, C, D), cyclic(A, B, E, F),
cyclic(B, C, F, G), cyclic(C, D, G, H), cyclic(D, A, E, H).
Conclusion: cyclic(E, G, F, H).

The Machine Proof

$\angle[HG, HE] - \angle[FG, FE]$

($\angle[HG, HE] = \angle[HG, HD] + \angle[HD, HE] = \angle[GC, DC] - \angle[EA, DA]$,
because cyclic(G,H,D,C), cyclic(E,H,D,A). (R16))

$= -\angle[FG, FE] + \angle[GC, DC] - \angle[EA, DA]$

($\angle[FG, FE] = \angle[FG, FB] + \angle[FB, FE] = \angle[GC, CB] - \angle[EA, BA]$,
because cyclic(G,F,B,C), cyclic(E,F,B,A). (R16))

$= \angle[GC, DC] - \angle[GC, CB] - \angle[EA, DA] + \angle[EA, BA]$

($\angle[GC, DC] - \angle[GC, CB] = -\angle[DC, CB]$. (R14))

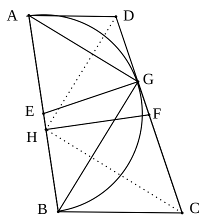
$= -\angle[EA, DA] + \angle[EA, BA] - \angle[DC, CB]$

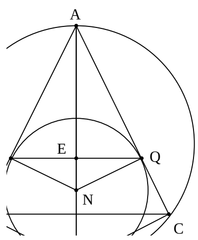
($-\angle[EA, DA] + \angle[EA, BA] = \angle[DA, BA]$. (R14))

$= -\angle[DC, CB] + \angle[DA, BA]$

($\angle[DC, CB] = \angle[DA, BA]$, because cyclic($C$,$D$,$B$,$A$). (R10))

$= \angle[0]$





Example 5 [4] In quadrilateral $ABCD$, $BC \parallel AD$ and the circle with AB as its diameter is tangent to CD. Show that the circle with CD as its diameter is tangent to AB.

Point order: A, B, C, D, E, G, F, H.
Hypotheses: para(A, D, B, C), midpoint(E, A, B), foot(G, E, C, D),
perp(A, G, B, G), midpoint(F, C, D), foot(H, F, A, B).
Conclusion: perp(C, H, H, D).

[4] This example is from the 1984 International Mathematical Olympiad

The Machine Proof
$-\angle[HD, HC] + \angle[1]$

$(\angle[HD, HC] = \angle[HD, HG] + \angle[HG, HC] = \angle[GB, CB] - \angle[GA, DA]$,
because cyclic(D,H,G,A), cyclic(C,H,G,B). (R16))

$= -\angle[GB, CB] + \angle[GA, DA] + \angle[1]$

$(\angle[GA, DA] = \angle[GA, CB]$, because $DA \parallel BC$. (R7))

$= -\angle[GB, CB] + \angle[GA, CB] + \angle[1]$

$(\angle[GB, CB] = \angle[GA, CB] + \angle[1]$, because $GB \perp AG$. (R8))

$= \angle[0]$

This example uses the forward chaining discussed in Subsection 3.2. Points A, D, G, H are cyclic, because $\angle[AD, CD] = \angle[AB, GH]$. This equation is further derived from the following two equations: $\angle[AD, CD] = \angle[EF, CD]$, because $EF \parallel AD$; $\angle[AB, GH] = \angle[EF, CD]$, because G, E, F, H are cyclic. G, E, F, H are cyclic, because $EG \perp CD$ and $HF \perp AB$. The fact that B, C, G, H are cyclic can be proved similarly.

Example 6 [5] In triangle ABC, $AB = AC$. A circle is tangent to the circumcircle of triangle ABC and is tangent to AB, AC at P and Q. Show that the midpoint of the PQ is the incenter of triangle ABC.

Point order: P, Q, A, N, D, B, C, E.
Hypotheses: cyclic(A, B, C, D), cong(A, P, A, Q), cong(A, B, A, C), cong(D, B, D, C), coll(N, A, D), foot(P, N, A, B), foot(Q, N, A, C), circumcenter(N, D, P, Q), coll(E, P, Q), coll(E, A, D).
Conclusion: eqangle(A, B, E, E, B, C).

The Machine Proof
$-\angle[EB, CB] - \angle[EB, BA]$

$(\angle[EB, CB] = \angle[BE, PE] = \angle[BD, DP]$, because $CB \parallel EP$.)

$= -\angle[EB, BA] - \angle[BD, DP]$

$(\angle[EB, BA] = \angle[ED, DP]$, because collinear$(A,B,P)$, cyclic$(B,E,P,D)$. (R10))

$= -\angle[ED, DP] - \angle[BD, DP]$

$(\angle[BD, DP] = -\angle[DP, NP]$, because $BD \parallel PN$. (R7))

$= -\angle[ED, DP] + \angle[DP, NP]$

$(\angle[ED, DP] = \angle[DP, NP]$, because $\angle[AD, DP] = \angle[DP, PN]$.)

$= \angle[0]$

This example uses the forward chaining discussed in Subsection 3.2.

[5] This problem is from the 1978 International Mathematical Olympiad