MM Research Preprints, 84–97
No. 18, Dec. 1999. Beijing

Automated Ordering for Automated Theorem Proving in Elementary Geometry – Degree of Freedom Analysis Method [1)]

Hongbo Li [2)] Minteh Cheng [3)]

Abstract. Wu's method on mechanical theorem proving in geometries is a well-known method. Ordering of geometric constructions is one of the most important and difficult problems when applying this method. In this paper we proposed a systematic mechanical ordering method based on geometric considerations. This method has been tested by about one hundred theorems in plane geometry.

Key words Wu's method, mechanical geometry theorem proving, automated ordering.

1. Introduction

Wu's method [Wu 78, 82] for mechanical geometry theorem proving is nowadays a well-known one. In this method, the first step is to introduce a coordinate system so that the geometric entities and relations are turned into algebraic ones, the coordinatization and algebralization procedure. The second step is to decide in an algorithmic manner whether the algebraic relations corresponding to the conclusion of a geometric theorem is formal consequence of the relations corresponding to the hypotheses of the theorem under some non degenerate conditions. When applying Wu's method, one difficulty is the problem of reducibility. It has been noticed that particular geometric considerations and deliberate choice of coordinates can help avoiding or at least lessening reducibility difficulties. Wu *et al.* pointed out, that there is a type of geometric theorems, called linear constructive type, which is a generalization of Hilbert's intersection point theorems, for which the hypotheses can be stated successively according to certain constructive statements by introducing new geometric entities, in each step the algebraic relations corresponding to the geometric ones being linear equations in the newly introduced geometric dependents as variables; that many geometric theorems seemingly not of this type can also be restated as this type by changing the manner of geometric statements. Much work has been done in this direction. In [Wu 82], [Chou 88], [Wang 89], Hilbert's type of geometric constructions is extended greatly by considering linear geometric constructions relating to circles, angles and metric properties. A typical example is that, given an intersection of a line with a circle, the construction of the other intersection is a linear geometric construction. In [Gao 90] there pointed out that the tangent functions of half angles can be used to resolve ambiguities in algebralization arising from indeterminancy of directions. In [Li 97] there proposed some new linear constructions, by which Thebault's Theorem (*cf.* [Chou 88]) is restated as a linear constructive type of theorem and is proved by hand very easily.

[1)] This paper was first written in 1993, last revised in 1995.

When the hypotheses of a geometric theorem need reformulation, thre first step of Wu's method actually takes five substeps:

1. Reformulate the hypotheses as a sequence of geometric statements.

2. Choose a coordinate system.

3. Separate the dependent variables from the independent ones.

4. Order all variables.

5. Turn the geometric statements into algebraic relations.

In usual application of Wu's method, the first four steps are artificial. In [Wang 90] there is an efficient heuristic ordering program making substeps 3 to 4 automatic. In this paper we propose a method to reformulate the hypotheses automatically, where by ordering we mean reformulating. The basic idea is to take the hypotheses of a theorem as a geometric diagram, thus disassemble the diagram into spare parts, analyse the spare parts and then assemble them in a different way to recover the original diagram. Since there is always some geometric objects which are key joints and which are difficult to be constructed in a linear manner, by assigning them full degree of freedom, they will contribute to simplifying the reconstruction of the diagram. We name the method *degree of freedom analysis method*, because at each step of assembly the constructions are controlled by various degrees of freedom. The basic idea of this method can be applied to other geometries, although we only realise it in plane geometry in this paper. Given a geometric theorem, our program can produce automatedly a kind of linear construction sequences or nonlinear ones leading to irreducible characteristic set, the linearity and irreducibility being known from the construction sequence directly. Nearly one hundred theorems in plane geometry are tested by our program. Experiments show that the method seems both efficient and intelligent.

In Section 2 there are some definitions. In Section 3 the method is described by introducing the implementation of our program. In Section 4 two working examples are given. Section 5 are some discussions.

2. Typical linear constructions

In elementary geometry, a geometric diagram can be represented by a finite number of *geometric objects*: points, directions and parameters. Directions are represented by unit vectors. The *configuration space* of a geometric problem is the set of geometric diagrams satisfying the hypotheses of the problem. Let a generic diagram be represented by m points, n parameters and directions. Then the configuration space can be represented as a subset of $\mathbf{E}^{2m+n}$. This subset is composed of a finite number of algebraic manifolds of distinct dimension. The largest dimension is irrelevant to the choice of geometric objects and is called the *total degree of freedom*.

The hypotheses of a geometric theorem can be taken as a construction sequence of a generic diagram, or equivalently, a construction sequence of the representing geometric objects. A construction of a geometric object is called irreducible if the characteristic set of the algebraic relations corresponding to the geometric construction is irreducible with respect to the newly introduced geometric dependents as variables. Especially, if the characteristic set is linear, the construction is called linear. If a construction is not irreducible, it is called reducible; if it is not linear, it is called nonlinear.

By an *order* of a geometric problem we mean a sequence of correct constructions of the geometric objects representing a generic diagram. A *linear order* is a sequence of linear constructions; an *irreducible order* is a sequence of irreducible constructions.

The study of linear constructions is fundamental for ordering. The linearity of a construction is often easily judged without computing the characteristic set, because of this, linear constructions are indispensable for judging whether or not an order is irreducible without computing the characteristic set. The topic of linear constructions has been well studied since the beginning of Wu's method. In the known linear constructions, some are basic, the others can be decomposed into a sequence of basic ones. Below we list some most often used linear constructions, called *typical linear constructions*. Those with asterisks are not basic ones.

Construction 1. Construct a free point, parameter or direction.

Construction 2. Given a polynomial equation linear with respect to variable x, assume that all the other variables have been constructed, construct x.

Construction 3. Given two points A, B, construct the point at the infinity of the line decided by them. The point at infinity can be represented by vector $A - B$ up to a nonzero scalar factor.

Construction 4. Given a point at infinity represented by vector A, construct another point at infinity B, assuming that the tangent function of the angle θ from vector A to vector B is given. (When $\theta \equiv \pi/2 \; mod \; \pi$, $B = A^{\sim}$, which is the image of the positive rotation of A by $\pi/2$; otherwise $B = A + \tan\theta A^{\sim}$.)

Construction 5. Given two points A, B, or a point A and a point at infinity C, construct a free point D on the line decided by them[4)]. ($D = A + t(B - A)$, or $A + tC$, where t is a scalar.)

Construction 6. Given two points A, B, or a point A and a point at infinity C, construct a point D on the line decided by them, such that the ratio of vector $D - A$ to $B - A$, or C, equals a given scalar.

Construction 7. Given two intersecting lines AB, CD, construct the intersection.

Construction 8*. Given triangle ABC, construct the center of the circumscribed circle of the triangle. (The center is the intersection of the two perpendicular bisectors of line segments AB, AC.)

Construction 9. (rational representation of free cyclic point) Given a circle decided by center O and a cyclic point A, or by three cyclic points A, B, C, construct a free cyclic point D. ($D = A + \frac{2}{t^2+1}((O - A) + t(O - A)^{\sim})$.)

[4)] When $C = P - Q$, the line passes through A and is parallel to line PQ; when $C = (P - Q)^{\sim}$, the line passes through A and is perpendicular to line PQ.

Construction 10. Given two points A, B, construct the image C of the positive rotation of B with respect to A, assuming that the tangent function of half the rotation angle θ is given. ($C = A + (B - A)\cos\theta + (B - A)^{\sim}\sin\theta$.)

Construction 11*. Given any two vertices of a positive regular n-polygon $A_1 A_2 \cdots A_n$, together with $w = \tan\frac{\pi}{n}$, construct the other vertices.
(When A_1, A_k are given, let O be the center of the circumscribed circle of the regular polygon, then $\frac{\angle A_1 O A_k}{2} = k\frac{\pi}{n}$, O can be constructed linearly. Now A_j can be constructed linearly from O, A_1 by $\frac{\angle A_1 O A_j}{2} = j\frac{\pi}{n}$.)

Construction 12*. Given line AB, construct the image of the reflection of a given point D with respect to line AB.

Construction 13*. Let B be an intersection of a line with circle O. Given B, O and the line, construct the other intersection C of the line with the circle. (C is the reflection of B with respect to the perpendicular drawn from O to the line.)

Construction 14*. Let I be the inner center of triangle ABC. Given I, B, C, construct A. (A is the intersection of the two reflections of line BC with respect to lines IB, IC.)

Construction 15. Let A be a point on circle O, B be a point not on circle O. Given A, O, B, construct points O_1, O_2, such that $O_1 - B$ is in the same (or opposite) direction with $O - A$, $O_2 - B$ is in the same (or opposite) direction with $O - A$, the circle with center O_1 and cyclic point B is inner tangent with circle O, the circle with center O_2 and cyclic point B is outer tangent with circle O. Also construct the radii of circles O_1, O_2.
(O_1 is decided by the equations

$$2(O_1 - O)\cdot(B - A) = (B - A)^2,\ \ (O_1 - B)\times(O - A) = 0. \tag{1}$$

$$(\text{or}\ \ (O_1 - O)\cdot(\frac{A+B}{2} - O) = (\frac{A+B}{2} - O)^2,\ \ (O_1 - B)\times(O - A) = 0.) \tag{2}$$

O_2 is decided by (2) (or (1)) with O_1 replaced by O_2. The radius of circle O_1 is $x = \frac{(O_1 - B)\cdot(O - A)}{(O - A)^2}$ (or $-x$). The radius of circle O_2 is x (or $-x$) with O_1 replaced by O_2.)

Construction 16. Let ABC be a triangle, $s = \tan\frac{\angle ABC}{2}$, $t = \tan\frac{\angle ACB}{2}$, a be the length of line segment BC. Let R, r be the radius of the circumscribed circle and the inscribed circle of the triangle respectively. Given a, s, t, construct R, r. ($R = \frac{a(1+s^2)(1+t^2)}{4(s+t)(1-st)}$, $r = \frac{ast}{s+t}$.)

Construction 17*. Let ABC be a triangle, the symbols a, s, t be the same as in Construction 16. Let O, I be the circumcenter and the inner center of triangle ABC respectively, I_A, I_B, I_C be the escenter of the triangle inside $\angle BAC$, $\angle ABC$, $\angle ACB$ respectively. Let r_A, r_B, r_C be the radius of circles I_A, I_B, I_C respectively. Given B, C, a, s, t, construct $O, I, I_A, I_B, I_C, r_A, r_B, r_C$.

3. Degree of freedom analysis method

By a *nice order* we mean either of the following two kinds of orders, the first one is a sequence of typical linear constructions; the second one is a parameter satisfying a nonlinear irreducible polynomial equation of one variable followed by a sequence of typical linear constructions. Our method aims to produce nice orders automatically.

First there are some definitions. The degree of freedom of a geometric object when setting it completely free is called the *complete degree of freedom*, abbrv. *c-deg*. At an instance of ordering, for an object, its degree of freedom is abbreviated *deg*; its number of typical linear constraints by a geometric construction is abbreviated *l-con*; its *typical linear degree of freedom*, abbr. *l-deg*, is the difference of its *c-deg* from its sum of *c-num* when it is not ordered yet, or its *l-deg* at the instance just before its ordering if it has been ordered. For a point A with $deg = 1$, if it is on one constructed circle, let $l\text{-}deg = 1$ (rational representation of free cyclic point). The *remaining total degree of freedom* at an instance of ordering, abbr. *r-tdeg*, is the difference of the total degree of freedom from the sum of the *l-deg*'s of the objects that have been ordered.

The procedure of automated ordering is:

Input. Hypotheses in geometric vocabulary: Assume that the hypotheses are correct and independent, i.e., inside the hypotheses there is neither contradiction nor implication. The hypotheses are expressed by a sequence of words in *geometric vocabulary*, with the order of introduction of the geometric objects identical with that in the original statement. There are about one hundred words in geometric vocabulary, which are similar to the usual vocabulary in geometry textbooks. For example, an isosceles triangle ABC with $AB = AC$ is expressed by $isosceles\text{-}triangle(A; B, C)$.

Step 1. Translating into basic vocabulary: First, from the hypotheses the *original order*, i.e., the order of the introduction of geometric objects, can be obtained.
Then translation is done, which is the first step of disassembly. There are about thirty words in basic vocabulary, which are simpler geometric relations.

In the following, "$\cdots$" represents a sequence of words or no word. "[ ]" means optional. The number at the end of each word is the constraint number which will be used in Step 3.

1. $point(A)$: *point* is the default head operator. It is always omitted. (0)
2. $ft(A)$: A is a foot. (0)
3. $cen(A)$: A is center of some circle. (0)
4. $dm(A, B)$: AB is a diameter of some circle. (0)
5. $dir(A)$: A is a direction. (0)
6. $par(A)$: A is a parameter. (0)
7. $seg(A, B)$: directed line segment from A to B. (0)
8. $ln(A, B, [\cdots])$: the straight line passing through $A, B, etc.$ ($n - 2$, n being the number of collinear points)
9. $cir(cen(O), A, [\cdots], [dm, \cdots])$: circle with center O, cyclic point A, *etc.* ($n - 2$, n being the number of elements in cir)

10. $cir(A, B, C, [\cdots], [dm, \cdots])$: ($n-3$, n being the number of elements in cir)
11. $sym(A, B; ln)$: A, B are symmetric with respect to ln. (2)
12. $sym([A]; ln_1; ln_2, ln_3$: ln_2, ln_3 are symmetric with respect to ln_1, A being the concurrent point of the three lines. When there are two distinct points on ln_2, ln_3 symmetric with respect to ln_3, this word is disused. (1)
13. $mid(A; B, C)$: A is midpoint of BC. (0)
14. $ch(A, B; ln, cir)$: appears in Step 3. AB is on ln and is a chord of cir.
15. $cir(cen(O), A, B, seg(A, B), [\cdots])$: appears in Step 3.
16. $rat(par; seg, seg$ or $dir)$: par is the ratio of two directed line segments or a directed line segment with a direction. In the hypotheses, generally any length should be represented by rat. (1)
17. $rad(par_1; A, B, C; par_2, par_3; par_4)$: par_1 is the radius of the circumscribed circle of triangle ABC, par_2, par_3 being $\tan\frac{\angle ABC}{2}$, $\tan\frac{\angle ACB}{2}$ respectively, par_3 being the signed length of directed line segment BC. (1)
18. $area(par; A, B, C)$: par is the directed area of positive triangle ABC. (1)
19. $pl(ln$ or $dir, [\cdots])$: abbreviates *parallel*. ($n-1$, n being the number of parallel lines or directions)
20. $pp(pl, pl)$: abbreviates *perpendicular*. (1)
21. $gcen(G; A, B, C)$: G is the gravity center of triangle ABC. (2)
22. $ocen(H; A, B, C; D, E, F)$: H is the orthocenter of triangle ABC, with D, E, F the foot drawn from A, B, C to the opposite sides respectively. (2)
23. $icen(I; A, B, C)$: I is the inner center of triangle ABC. (2)
24. $itan(cir(cen(O), A, [\cdots]), A; cir(cen(O_1), B, [\cdots]), B)$: Circle O is inner tangent with circle O_1, $O-A$ is in the same direction with O_1-B. It can also also represent that circle O is outer tangent with circle O_1, $O-A$ is in the opposite direction of O_1-B. (1)
25. $otan(cir(cen(O), A, [\cdots]), A; cir(cen(O_1), B, [\cdots]), B)$: Circle O is outer tangent with circle O_1, $O-A$ is in the same direction with O_1-B. It can also also represent that circle O is inner tangent with circle O_1, $O-A$ is in the opposite direction of O_1-B. The abbreviation is $outan$. (1)
26. $rot(dir, dir; par)$: par is the tangent function of half the angle of rotation of direction. (1)
27. $rot(seg, seg; par)$: par is the tangent function of half the angle of rotation of directed line segment. (1)
28. $rot(ln, ln; par)$: par is the tangent function of the angle of rotation of point at infinity. (1)
29. $gon([cen(O)], A_1, A_2, \cdots, A_n)$: appears in Step 3. $A_1A_2\cdots A_n$ is a positive regular n-polygon with center O. ($2(n-2)$ when there is no center, $2(n-1)$ otherwise)
30. $eqn([par_1, \cdots]; [par_2, \cdots]$; an equation): The equation is linear with respect to par_1, *etc.*, nonlinear with respect to par_2, *etc.* (1)

The translation can be illustrated by the following examples, where a single arrow means "be replaced by".

- $foot(D; A; line(B, C)) \longrightarrow$
  $ft(D), ln(D, B, C), pp(pl(ln(A, D)), pl(ln(D, B, C)))$.
- $midpoint(M, line\text{-}segment(A, B)) \longrightarrow$
  $ln(M, A, B),\ rat(\frac{1}{2}; seg(M, A), seg(B, A)),\ mid(M; A, B)$.
- $positive\text{-}equilateral\text{-}triangle(A, B, C) \longrightarrow$
  $par(w), eqn(; par(w); 3w^2 = 1), gon(A, B, C)$.

Step 2. Computing total degree of freedom: First, reasoning and grouping of geometric relations are done. The reasoning is a procedure of merging geometric relations. It changes neither the correctness nor the independency of the hypotheses.
Because of this, the total degree of freedom equals difference of the sum of all the *c-deg*'s from the number of constraints defined by different words in the result of reasoning and grouping.

Below we list some typical reasonings, where a double arrow means "imply", two words with the same name and footscript are identical.

1. $ln(A, B, [C, \cdots]), ln(A, B, [D, \cdots]) \longrightarrow ln(A, B, [C, D, \cdots])$.
2. $pl(ln(A, [B, \cdots]), ln(A, [C, \cdots]), [\cdots]) \Longrightarrow ln(A, [B, C, \cdots])$.
3. $pl(ln_1, ln_2, [\cdots]), pl(ln_1, ln_2, [\cdots]) \longrightarrow pl(ln_1, ln_2, ln_3, [\cdots])$.
4. $pp(pl_1, pl(ln_2, [\cdots])), pp(pl_1, pl(ln_3, [\cdots])) \longrightarrow$
   $pl(ln_2, ln_3, [\cdots]), pp(pl_1, pl(ln_2, ln_3, [\cdots])$.
5. $pp(pl(ln(A, B, [\cdots]), [\cdots]), pl(ln(A, C, [\cdots]), [\cdots])) \Longrightarrow ft(A)$.
6. $cir(cen(O), A, [B, \cdots]), cir(cen(O), A, [C, \cdots]) \longrightarrow cir(cen(O), A, [B, C, \cdots])$.
7. $cir(A, B, C, [D, \cdots]), cir(A, B, C, [E, \cdots]) \longrightarrow cir(A, B, C, [D, E, \cdots])$.

Step 3. Reasoning: This is the second step of disassembly, followed by the second step of reasoning and grouping of geometric relations.

1. $ocen(H; A, B, C; D, E, F) \Rightarrow \begin{cases} ln(H, A, D), & ln(H, B, E); \\ ln(H, A, D), & ln(H, C, F); \\ ln(H, B, E), & ln(H, C, F). \end{cases}$
2. $icen(I; A, B, C) \Rightarrow$
   $\begin{cases} sym(A; ln(I, A); ln(A, B), ln(A, C)), sym(B; ln(I, B); ln(A, B), ln(B, C)); \\ sym(A; ln(I, A); ln(A, B), ln(A, C)), sym(C; ln(I, C); ln(A, C), ln(B, C)); \\ sym(B; ln(I, B); ln(A, B), ln(B, C)), sym(C; ln(I, C); ln(A, C), ln(B, C)). \end{cases}$
3. $mid(M; A, B)$ is removed if there is no $cir(cen(O), A, B, [\cdots])$.
4. $mid(M; A, B), cir(cen(O_1), A, B, [\cdots]), \cdots, cir(cen(O_n), A, B, [\cdots]) \longrightarrow$
   $ft(M), ln(M, O_1, \cdots, O_n), pp(pl(ln(M, O_1, \cdots, O_n)), pl(ln(M, A, B)))$,
   $cir(cen(O_1), A, B, seg(A, B), [\cdots]), \cdots, cir(cen(O_n), A, B, seg(A, B), [\cdots])$.

5. If there is no $mid(M; A, B)$, $n > 1$, then:
$cir(cen(O_1), A, B, [\cdots]), \cdots, cir(cen(O_n), A, B, [\cdots]) \longrightarrow$
$ln_1(O_1, O_2, \cdots, O_n), sym(A, B; ln_1), cir(cen(O_1), A, B, seg(A, B), [\cdots]),$
$\cdots, cir(cen(O_n), A, B, seg(A, B), [\cdots])$.
6. If a circle is equipped without a center and only one diameter, then:
$cir(A, B, dm(A, B), C, [\cdots]) \longrightarrow ft(C), [ft, \cdots],$
$pp(pl(ln(A, C)), pl(ln(B, C))), [pp, \cdots]$.

For $ln_1(A, B, C, [\cdots])$, if there is only one circle cir_1 contains both A, B without $dm(A, B)$ and $seg(A, B)$, then a new word $ch(A, B, ln_1, cir_1)$ is produced.

Step 4. Testing the nicety of the original order: Prepend *null* to the original order. Call *nice-order-test*, with the original order as input. If the order is nice and no more nice order is required, exit.

Procedure *nice-order-test*

Order queue is the input order.

Order stack is the working stack, where except for the first position, pushing into an object is constructing it linearly. The initial value is empty. The *l-deg* of the first object in the stack is always set to be zero.

Construction sequence is the sequence of constructions for output.

begin *nice-order-test*

Substep 1. Push the first object of the order queue into the order stack. Put the independent typical linear constructions of the object into the construction sequence. Check whether there is any object in the order queue with negative *l-deg*, if there is any, go to Substep 3; else compute *r-tot*.

Substep 2. If *r-tot*> 0, go to Substep 1. If *r-tot* $= 0$ and the order queue is empty, go to Substep 4.

Substep 3. Output the order stack, return that the order is not nice.

Substep 4. Output the construction sequence, return that the order is nice.

end *nice-order-test*

Below we list the translations in basic vocabulary of the typical linear (partial) constructions of a geometric object A, each with the *c-num*. When an object B is constructed, set $B = 0$.

1. $ln(0, 0, A, [\cdots]) \Rightarrow (A : 1)$.
2. In *pl*, set $ln(0, 0, [\cdots]) = 0$. Then
$pl(0, ln(0, A, [\cdots]), [\cdots]) \Rightarrow (A : 1)$.
3. In *pp*, set $pl(0, [\cdots]) = 0$. Then $pp(0, pl(ln(0, A, [\cdots]), [\cdots])) \Rightarrow (A : 1)$.
4. $cir(cen(A), 0, 0, [B, \cdots]) \Rightarrow (A : 1)$; $cir(cen(A), 0, 0, 0, [\cdots]) \Rightarrow (A : 2)$.
5. $ch(0, A; 0, cir(center(0); 0, A, [\cdots])) \Rightarrow (A : 1)$.
6. $ch(0, A; 0, cir(0, 0, 0, [\cdots])) \Rightarrow (A : 1)$.

7. $gon(0,0,A,[\cdots]) \Rightarrow (A:2)$; $gon(cen(0),0,A,[\cdots]) \Rightarrow (A:2)$; $gon(cen(A),0,0,[\cdots]) \Rightarrow (A:2)$.
8. $gcen(0;0,0,A) \Rightarrow (A:2)$; $gcen(A;0,0,0) \Rightarrow (A:2)$.
9. In rat, rot, set $seg(0,0)=0$. Then $rat(0;0,seg(0,A)) \Rightarrow (A:1)$; $rat(0;0,dir(A)) \Rightarrow (A:1)$; $rat(par(A);0,0) \Rightarrow (A:1)$.
10. $rot(0,seg(0,A);0) \Rightarrow (A:1)$; $rot(0,direction(A);0) \Rightarrow (A:1)$; $rot(0,line(0,A,[\cdots]);0) \Rightarrow (A:1)$; $rot(0,0;parameter(A)) \Rightarrow (A:1)$.
11. In sym, rot, ch, set a constructed line to be 0. Then $sym([0];0;0,ln(A,[\cdots])) \Rightarrow (A:1)$.
12. $eqn(par(A),[0,\cdots];[0,\cdots];\text{an equation}) \Rightarrow (A:1)$.
13. $area(par(A);0,0,0) \Rightarrow (A:1)$; $area(0;0,0,A) \Rightarrow (A:1)$.
14. In $i(o)tan$, set a constructed circle to be 0. Then $i(o)tan(cir(cen(A);0,[\cdots]),0;0,0) \Rightarrow (A:1)$.
15. $rad(par(A);B,C,D;0,0,0) \Rightarrow (A:1)$; $rad(0;B,C,D;0,0,par(A)) \Rightarrow (A:1)$.
16. In a $dup(\text{group}_1;\text{group}_2)(\text{words; words; }[\cdots])$:
    If point $A \in \text{group}_1$, then if at least one "words" provide A with nonzero *c-num*, $(A:1)$; otherwise $(A:0)$.
    If point $A \in \text{group}_2$, then if at least one "words" provide A with at least two *c-num*'s, $(A:1)$; else if at least one "words" provides A with a *c-num*, $(A:1)$; else $(A:0)$.

Step 5. Setting up precedence table: To speed up ordering, a precedence table of geometric objects is needed. Such a table is mainly based on experience of ordering.

In our method, making use of the feature that the total degree of freedom of a plane geometric theorem is usually between 6 and 8, we propose three kinds of precedence tables: 1-table, 2-table and 3-table.
1-table is used to choose the first geometric with *c-num*=2, or the first two objects if the first chosen one has *c-num*=1. In our program, the 1-table is:

tangent functions $\succ$ *vertices of angles* $\succ$ *points on bisectors of angles* $\succ$ *directions* $\succ$ *ratios* $\succ$ *common cyclic points ordered by the number of passing through circles* $\succ$ *cyclic points which are also feet* $\succ$ *cyclic points ordered by the number of passing-through lines* $\succ$ *feet* $\succ$ *centers* $\succ$ *points ordered by the number of passing-through lines* $\succ$ *others.*

2-table is used to choose one object right after 1-table is not used. It is by moving ahead those objects that are closely related to the objects in the order stack, *e.g.*,
tangent functions $\succ$ *vertices of the angles* $\succ$ *points on bisectors of the angles* $\succ$ *points on sides of the angles;*
directions $\succ$ *relevant ratios;*
common peripharal point of two circles $\succ$ *the other common cyclic point;*
cyclic point $\succ$ *cyclic points ordered by the number of passing-through lines;*
other point $\succ$ *collinear points ordered by the number of passing-through lines.*

2-table is produced during ordering, i.e., in Step 6 below.

3-table is used to choose the rest objects. In our program, it is obtained by interchanging *feet* and *cyclic points ordered by the number of passing through lines* in 1-table.

Step 6. Ordering – degree of freedom analysis: In the following program we make use of the feature that a theorem in plane geometry always has at least 5 degrees of freedom.

Object has two domains when ordering: the name and the *l-deg*.

Order stack is the working stack.

Failure set is the depositary for non nice orders and outdated nice ones.

i is the ordinal number of the precedence table to be used. The initial value is 1.

len is the number of objects in the order stack at an instance.

sgn $\in \{0, 1\}$ is used in changing precedence table when ordering. The initial value is 0.

Procedure *record*

begin *record*

If $len = 1$, output that no nice order can be found, exit. Otherwise, pop up all the objects at the end of the stack with $l\text{-}deg = 0$, record the instance of the stack in the failure set; pop up the last object, if $len \in \{2 + sgn, 3 + sgn\}$, let $i = i - 1$; return the instance of the stack.

end *record*

begin *main program*

Substep 1. If there is a parameter defined by a nonlinear polynomial equation of one variable, push it into the order stack; otherwise let the first object of the stack be *null*.

Substep 2. According to i-table, choose an object outside the stack; if the instance of the stack after pushing into the object belongs to the failure set, choose another object outside the stack and repeat the test. If no object can be chosen, call *record*, go back to the beginning of Substep 2. Else push the chosen object into the stack; if $len = 2$ and the *l-deg* of the object just pushed is 1, let $sgn = 1$ and go back to the beginning of Substep 2.

Substep 3. If there is any object outside the stack with negative *l- deg*, call *record* and go back to Substep 2. Else push all the objects outside the stack with $l\text{-}deg = 0$ into the stack, compute *r-tot*:

- If $r\text{-}tot > 1$, then if $i < 3$, let $i = i + 1$. Go back to Substep 2.
- If $r\text{-}tot = 1$, then if there is no object outside the stack with $l\text{-}deg = 1$, call *record*, go back to Substep 2.
- If $r\text{-}tot = 0$, then if there is any object outside the stack, call *record*, go back to Substep 2, else go to Substep 5.

Substep 4. According to 3-table, choose an object outside the stack with $l\text{-}deg =$ 1. If the instance of the stack after pushing into the object belongs to the failure set, choose another object outside the stack with *l- deg* $= 1$ and repeat the test. If no object can be chosen, call *record*, go back to Substep 2. Else push the chosen object into the stack; push all the objects outside the stack with *l-deg* $= 0$ into the stack; if there is any object outside the stack, call *record*, go back to the beginning of Substep 4.

Substep 5. A nice order is output from the stack. Call *nice-order- test* to produce the construction sequence.

Substep 6. If no more ordering is required, exit; otherwise a positive integer n is to be asked by man-machine dialogue. n means that it is better that the next nice order be the same with the present order in the first n objects. The default value of n is $m - 1$, where m is *len* after poping up all the objects at the end of the stack with *r-deg*$= 0$. When n is given, then if n is greater than the default value, output that n is unsuitable, go back to the beginning of Substep 6; else check the *r-tot* at the instance of the stack after poping up all but the first n objects, if it equals 1, go back to Substep 4, otherwise go back to Substep 2.

end *main program*

This ends Step 6.

This is the end of the whole method.

4. Working examples

We have tested about one hundred theorems from plane geometry using a program based on degree of freedom analysis method. The program is written in *Mathematica 1.2* and run on a LEO 486/33. In the program, we avoid the use of rational representation of free cyclic point. In the experiments, the first nice order is produced in $4.39 \sim 17.36$ seconds; from the second order on, a nice order takes in average $1.09 \sim 12.05$ seconds. Timing of the first order is made from the beginning of translation, so it is generally longer. Below we give two working examples.

Example 1. (Simson's Theorem) Let A, B, C, D be points on a circle with center O. Let E, F, G be the foot drawn from D to lines AB, BC, AC respectively. Then E, F, G are collinear.

The hypotheses in geometric vocabulary are:

$$\begin{array}{l} circle(center(O), A, B, C, D), \\ foot(E; D, line(A, B)), \\ foot(F; D, line(B, C)), \\ foot(G; D, line(A, C)). \end{array}$$

The result from translation is:

$$\begin{array}{l} cen(O), ft(E), ft(F), ft(G), \\ cir(cen(O), A, B, C, D), \\ ln(A, B, E), ln(B, C, F), ln(A, C, G), \\ pp(pl(ln(D, E)), pl(ln(A, B, E))), \\ pp(pl(ln(D, F)), pl(ln(B, C, F))), \\ pp(pl(ln(D, G)), pl(ln(A, C, G))). \end{array}$$

The result from assortment is by adding to the above result the following:

$$\begin{array}{l} ch(A,B,ln(A,B,E),cir(cen(O),A,B,C,D)), \\ ch(B,C,ln(B,C,F),cir(cen(O),A,B,C,D)), \\ ch(A,C,ln(A,C,G),cir(cen(O),A,B,C,D)). \end{array}$$

The total degree of freedom is: $2 \times 8 - (3+3+4-1) = 7$. The original order is: O, A, B, C, D, E, F, G. It is not a nice order. The first object which is not linearly constructed is: B.

The precedence table is: $D \succ A, B, C \succ E, F, G \succ O$. 3-table is: $D \succ E, F, G \succ A, B, C \succ O$.

In the order stack, the first object is *null*, the second one is D. 2-table is: $A, B, C \succ E, F, G \succ O$. The third object is A.

The first nice order is: $null, D^2, A^2, F^2, B^1, E^0, O^0, C^0, G^0$. It takes 6.17 seconds. The constructions are:

$$\begin{array}{ll} B^1: & pp(pl(ln(D,F)),pl(ln(B,C,F))); \\ E^0: & ln(A,B,E),pp(pl(ln(D,F)),pl(ln(B,C,F))); \\ O^0: & cir(cen(O),A,B,C,D); \\ C^0: & ln(B,C,F),ch(B,C,ln(B,C,F),cir(cen(O),A,B,C,D)); \\ G^0: & ln(A,C,G),pp(pl(ln(D,G)),pl(ln(A,C,G))). \end{array}$$

We list some other nice orders produced later:

$$\begin{array}{ll} null, D^2, E^2, F^2, B^0, A^1, O^0, C^0, G^0 & 1.93 \text{ seconds;} \\ null, D^2, E^2, A^1, C^2, G^0, O^0, B^0, F^0 & 2.01 \text{ seconds;} \\ null, A^2, B^2, E^1, F^2, D^0, O^0, C^0, G^0 & 3.33 \text{ seconds.} \end{array}$$

Example 2. Let P, A, B, C be points on a circle with center O. Draw three circles with PA, PB, PC as the diameters respectively. Let P_1 be the other intersection of the circle of diameter PA with that of diameter PB, P_2 be the other intersection of the circle of diameter PB with that of diameter PC, P_3 be the other intersection of diameter PA with that of diameter PC. Then P_1, P_2, P_3 are collinear.

The hypotheses in geometric vocabulary are:

$$\begin{array}{l} circle(center(O),P,A,B,C), \\ diameter(P,A,circle(P,A,P_1)), \\ diameter(P,B,circle(P,B,P_2)), \\ diameter(P,C,circle(P,C,P_3)), \\ intersection(P_1,circle(P,A,P_1),circle(P,B,P_2)), \\ intersection(P_2,circle(P,B,P_2),circle(P,C,P_3)), \\ intersection(P_3,circle(P,A,P_1),circle(P,C,P_3)). \end{array}$$

The result from translation is:

$$\begin{array}{l} cen(O),cir(cen(O),P,A,B,C), \\ dm(P,A,cir(P,A,P_1)),dm(P,B,cir(P,B,P_2)),dm(P,C,cir(P,C,P_3)), \\ cir(P,A,P_1),cir(P,B,P_1,P_2),cir(P,B,P_2), \\ cir(P,C,P_2,P_3),cir(P,A,P_1,P_3),cir(P,C,P_3). \end{array}$$

The result from reasoning is:

$$\begin{aligned}
&cen(O), ft(P_1), ft(P_2), ft(P_3),\\
&cir(cen(O), P, A, B, C),\\
&ln(A, B, P_1), ln(B, C, P_2), ln(A, C, P_3),\\
&pp(pl(ln(P, P_1)), pl(ln(A, B, P_1))),\\
&pp(pl(ln(P, P_2)), pl(ln(B, C, P_2))),\\
&pp(pl(ln(P, P_3)), pl(ln(A, C, P_3))),\\
&ch(A, B, ln(A, B, P_1), cir(cen(O), P, A, B, C)),\\
&ch(B, C, ln(B, C, P_2), cir(cen(O), P, A, B, C)),\\
&ch(A, C, ln(A, C, P_3), cir(cen(O), P, A, B, C)).
\end{aligned}$$

The total degree of freedom is: 7. The original order is: $O, P, A, B, C, P_1, P_2, P_3$. It is not a nice order. The first object which is not linearly constructed is: A.

The precedence table is: $P \succ A, B, C \succ O \succ P_1, P_2, P_3$. 3-table is: $P \succ P_1, P_2, P_3 \succ A, B, C \succ O$.

In the order stack, the first object is $null$, the second one is P. 2-table is: $A, B, C \succ O \succ P_1, P_2, P_3$. The third object is A.

The first nice order is: $null, P^2, A^2, P_2^2, B^1, P_1^0, O^0, C^0, P_3^0$. It takes 6.03 seconds. The constructions are:

$$\begin{aligned}
B^1 :&\quad pp(pl(ln(P, P_2)), pl(ln(B, C, P_2)));\\
P_1^0 :&\quad ln(A, B, P_1), pp(pl(ln(P, P_1)), pl(ln(A, B, P_1)));\\
O^0 :&\quad cir(cen(O), P, A, B, C);\\
C^0 :&\quad ln(B, C, P_2), ch(B, C, ln(B, C, P_2), cir(cen(O), P, A, B, C));\\
P_3^0 :&\quad ln(A, C, P_3), pp(pl(ln(P, P_3)), pl(ln(A, C, P_3))).
\end{aligned}$$

Some other nice orders produced after the first one are:

$$\begin{aligned}
&null, P^2, P_1^2, P_2^2, B^0, A^1, O^0, C^0, P_3^0 &&\text{2.23 seconds;}\\
&null, P^2, P_1^2, A^1, C^2, O^0, P_3^0, B^0, P_2^0 &&\text{1.89 seconds;}\\
&null, P_1^2, P_2^2, A^2, B^1, P^0, O^0, C^0, P_3^0 &&\text{3.06 seconds.}
\end{aligned}$$

5. Discussions

There are three kinds of ordering techniques:

1. Use the same geometric objects as in the original order of the geometric problem, reconstruct them in a different order.

2. Use geometric objects different from those in the original order, reconstruct the configuration space.

3. Remove part of the hypotheses, take the conclusion as part of the hypotheses, prove that the original configuration space is included in the new configuration space up to some degenerate cases. Generally if we can prove that the removed part of the hypotheses still holds, this inclusion can be ensured by some uniqueness properties of geometric constructions.

References

[Wu 78] Wentsun Wu, On the decision problem and the mechanization of theorem in elementary geometry, Scientia Sinica 21, 1978.

[Wu 82] Wentsun Wu, Toward mechanization of geometry – some comments on Hilbert's *Grundlagen der Geometrie, Acta Math. Sci.* 2, 1982.

[Wand and Hu 87] Dongming Wang and Sen Hu, A mechanical theorem proving system for constructible theorems in elementary geometry, Chinese J. Sys. Sci. and Math. Sci. 7, 1987.

[Chou 88] S. C. Chou, *Mechanical Geometry Theorem Proving*, D. Reidel, Dordrecht, 1988.

[Wang 89] Dongming Wang, On Wu's Method for proving constructive geometric theorems, *MM Research Preprints* 3, 1989.

[Gao 90] Xiaoshan Gao, Transcendental functions and mechanical theorem proving in elementary geometries, *J. Automated Reasoning* 6, 1990.

[Li 97] Li Hongbo, Ordering in mechanical geometry theorem proving, *Science in China*, Series A, **40**(3): 225–233, 1997.