# A Web–based Intelligent System for Geometric Discovery

Francisco Botana

Departamento de Matemática Aplicada
Universidad de Vigo
Campus A Xunqueira, 36005 Pontevedra, Spain
fbotana@uvigo.es

**Abstract.** An open web–based tool for automatic discovery in elementary Euclidean geometry, webDiscovery, is described. It is based in recent findings in automatic discovery in geometry. A user–defined geometric construction is uploaded to a Java Servlet server, where two computer algebra systems, CoCoA and Mathematica, return the discovered facts about the construction. webDiscovery can be efficiently used in mathematics education, linkage design and testing and computer aided geometric design. The system can be tested at `rosalia.uvigo.es/sdge/web/2D`.

## 1 Introduction

Since the birth of Artificial Intelligence research, considerable attention has been paid to develop computer programs for automatic geometry reasoning. First attempts [9, 20, 22] were based on the synthetic approach, while from the 80's onwards the algebraic approach (mainly Wu's method and the Groebner basis method) [25, 6, 13] and the geometric invariant approach [7] revived interest in the field.

Two new interactive environments for learning geometry [17, 12] appeared also in the 80's and defined a field of computer aided instruction referred to as dynamic geometry. Dynamic geometry software (see also [23, 8, 1]) refers to computer programs where accurate construction of geometric configurations can be done. The key characteristic of this software is that unconstrained parts of the construction can be moved and, as they do, all other elements automatically self–adjust, preserving all dependent relationships and constraints [16]. Dynamic geometry environments have been considered as instruments for theorem proving and discovery [18, 11, 19].

In this paper, an open web–based tool for automatic discovery in elementary Euclidean geometry is proposed. Section 2 explains what is understood by automatic discovery and gives an overall description of the mathematical bases of the approach. The implementation on the server's side is discussed in Section 3, and some examples and limitations of the proposal are given in Section 4. Finally, some ways for further development are proposed.

## 2 Automatic Discovery in Euclidean Geometry

While automatic proving deals with verifying geometric statements, and automatic derivation relates to finding geometric formulae holding among prescribed geometric magnitudes, automatic discovery refers to finding complementary hypotheses for arbitrary statements to become true [21], or, in other words, to finding the missing hypotheses so that a given conclusion follows from a given incomplete set of hypotheses [15].

Our program follows an algebraic approach. It performs automatic discovery in Euclidean geometry via algorithmic commutative algebra and algebraic geometry using Groebner bases [4]. An in–depth description of the mathematical framework used can be found in [14, 21, 3].

Roughly speaking, the procedure is as follows. A statement (a finite set of hypotheses and a thesis) of the subclass of geometric constructions expressed using the concepts of parallelism, incidence and congruence, is considered, where the conclusion does not follow from the hypotheses. Symbolic coordinates are assigned to the points of the construction (where every free point gives up two new free variables $u_i, u_{i+1}$, and every bounded point gives up two new dependent variables $x_j, x_{j+1}$), and the hypotheses $h_1, \ldots, h_n$ and the thesis $t$ are rewritten as polynomials in $\mathbf{Q}[\mathbf{u}, \mathbf{x}]$. Since the geometric information of the construction is contained in the polynomial system, any Groebner basis of the system will contain the same knowledge. Thus, eliminating the dependent variables in the ideal $(hypotheses, thesis)$, the vanishing of every element in the elimination ideal $(hypotheses, thesis) \cap \mathbf{Q}[\mathbf{u}]$ is a necessary condition for the statement to hold.

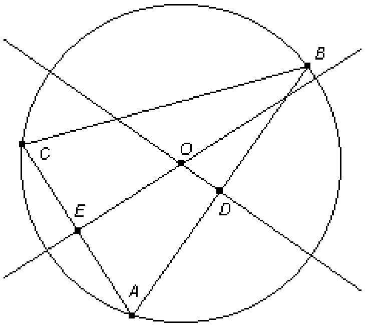


**Fig. 1.** When is $O$ on side $BC$?

In order to illustrate the approach, we use a simple example showing that a necessary condition for the alignment of the circumcenter of a triangle on one of its sides is the rightness of the triangle. The geometric construction of Fig. 1 consists of three free points $A(0, 0)$, $B(1, 0)$, $C(u_1, u_2)$, (there is no restriction in assuming $A$ and $B$ as origin and unit points), the perpendicular bisectors of sides $AB$ and $AC$ defined by means of their midpoints $D(x_1, x_2)$, $E(x_3, x_4)$, and the circumcenter $O(x_5, x_6)$ lying on both lines. Thus, the polynomials are $x_1 - 1/2, x_2, x_3 - u_1/2, x_4 - u_2/2, x_5 - x_1, u_2(x_6 - x_4) - u_1(x_5 - x_3)$. Adding as condition the collinearity of $O, B, C$, (that is, the polynomial $x_6(u_1 - x_5) - (u_2 - x_6)(x_5 - 1)$), the elimination of dependent variables returns $1/2u_1^3 + 1/2u_1u_2^2 - u_1^2 + 1/2u_1$, a polynomial whose zeros are those of $u_1$ (note that in this case $A, B, C$ are collinear, that is, the triangle is degenerated), or those of $u_1^2 + u_2^2 - 2u_1^2 + 1$, which can be understood as a perpendicularity condition between sides $AB$ and $AC$.

## 3 System Description

Since the main topic in this paper is geometric theorem discovery and web–availability, it is necessary to emphasize the way it works on the server's side. webDiscovery uses webMathematica, a Java servlet technology allowing remote access to the symbolic capabilities of Mathematica. Furthermore, CoCoA [5], an efficient system specialized in operations over commutative rings of polynomials, is used to perform computationally expensive tasks such as factoring and Groebner bases computations. A sketch of the main tasks in webDiscovery is shown in Fig. 2.

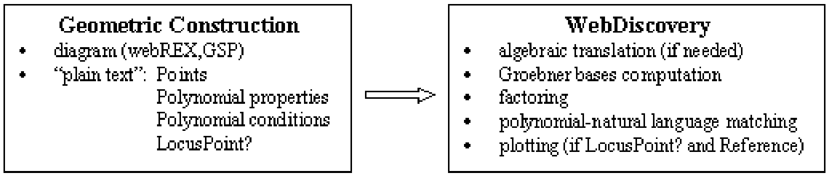
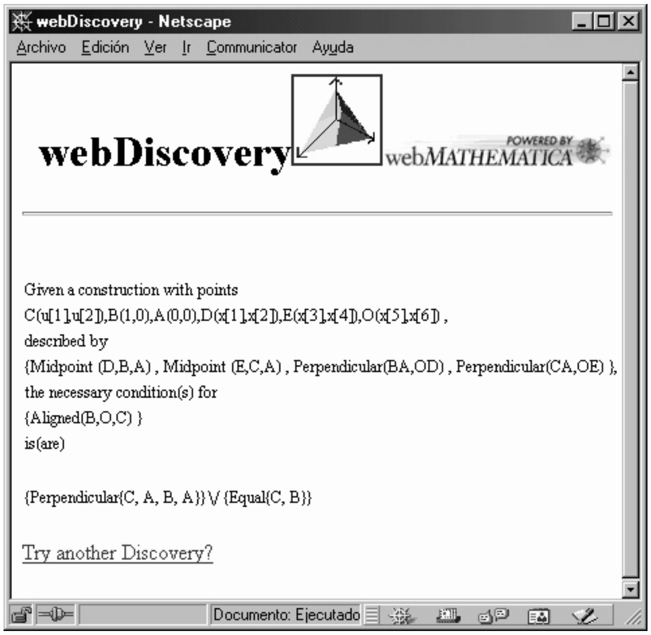

Geometric Construction
• diagram (webREX,GSP)
• "plain text": Points
Polynomial properties
Polynomial conditions
LocusPoint?
WebDiscovery
• algebraic translation (if needed)
• Groebner bases computation
• factoring
• polynomial-natural language matching
• plotting (if LocusPoint? and Reference)

**Fig. 2.** The architecture and main tasks of webDiscovery

The first step in discovery consists in uploading the geometric construction into the server. This is done via a plain text file written by the user or generated by a dynamic geometry environment. The text file contains the polynomial and linguistic knowledge about the construction, and its structure is illustrated as follows for the case of the circumcenter (where an * denotes an optional line):

```
Points
C(u[1],u[2])
B(1,0)
A(0,0)
D(x[1],x[2])
E(x[3],x[4])
O(x[5],x[6])
(blank line)
Properties
```

```
(x[1]-(1+0)/2)
(x[2]-(0+0)/2)
(x[3]-(u[1]+0)/2)
(x[4]-(u[2]+0)/2)
((0-u[2]))*(x[6]-x[4])+((0-u[1]))*(x[5]-x[3])
((0-0))*(x[6]-x[2])+((0-1))*(x[5]-x[1])
(blank line)
Conditions
(x[6]-0)*(u[1]-x[5])-(u[2]-x[6])*(x[5]-1)
(blank line)
LingProperties
Midpoint(D,B,A)*
Midpoint(E,C,A)*
Perpendicular(BA,OD)*
Perpendicular(CA,OE)*
(blank line)
LingConditions
Aligned(B,O,C)*
(blank line)
LocusPoint
```



Given a construction with points
C(u[1],u[2]),B(1,0),A(0,0),D(x[1],x[2]),E(x[3],x[4]),O(x[5],x[6]) ,
described by
{Midpoint (D,B,A) , Midpoint (E,C,A) , Perpendicular(BA,OD) , Perpendicular(CA,OE) },
the necessary condition(s) for
{Aligned(B,O,C) }
is(are)

{Perpendicular{C, A, B, A}} \/ {Equal{C, B}}

Try another Discovery?

**Fig. 3.** The necessary conditions for the alignment of the circumcenter

Although it is possible to write by hand this file or similar ones, there is a dynamic geometry program, webREX [24], which can export geometric constructions in this format. The use of a well–known dynamic geometry environment, *The Geometer's Sketchpad* [12], for performing automatic discovery through the web has also been investigated, and preliminary results can be found at `http://193.146.36.49/Discovery`.

Once the file has been uploaded into the server, a Mathematica Server Page is launched. It reads the file, initializing variables and deciding which class of task (general or locus discovery) is needed. An initialization file for CoCoA, containing the ideal generated by the properties and conditions polynomials, is also written out, and CoCoA, launched inside Mathematica, computes a Groebner basis of this ideal. For the case of the circumcenter, the basis ideal, as returned by CoCoA, is `Ideal(1/2u[1]^3 + 1/2u[1]u[2]^2 - u[1]^2 + 1/2u[1])`. Each generator is factored (a task also done by CoCoA), and a process of logical expansion is performed on the conjunction of the generators in order to remove repeated factors. Each element in the final conjunctive normal form is matched against a database of linguistic predicates, returning a compound proposition of elementary geometric facts and/or equations, if the pattern matching has not succeeded for any of the factors (Fig. 3).

If we were trying to discover a locus, the returned factors would be classified as points, lines, conics or general curves. Additionally, the Mathematica graphic abilities are used to plot the locus.

## 4 Examples and Limitations

webDiscovery can be used to discover (or rediscover) a vast class of Euclidean properties. The following examples illustrate its abilities.

### 4.1 General Discovery

Example 1 (Simson Steiner theorems). *Given a triangle $ABC$ and a point $X$ on its plane, which are the necessary conditions for the collinearity of the perpendicular projections $M, N, P$ of $X$ on the triangle sides?* (Fig. 4)

Selecting $A$ as origin and $B$ as unit, and imposing the alignment of $M, N, P$, the system finds $\mathrm{Aligned}(A, B, C) \vee \mathrm{OnCircle}(A, B, C, X)$ as necessary conditions for the statement to become true. Note that the triangle being not degenerated, the well–known thesis of Simson Wallace theorem has been found.

Using again the construction, webDiscovery easily allows a generalization. If we change the imposed condition to $\mathrm{area}(M, N, P) = 1$, the system returns as necessary condition the equation $2\,u(5)^2-4\,u(5)^3+2\,u(5)^4+2\,u(6)^2-4\,u(5)\,u(6)^2+4\,u(5)^2\,u(6)^2+2\,u(6)^4-u(6)^3\,u(7)+u(6)^3\,u(7)^2+u(5)\,u(6)^2\,u(8)-u(5)^2\,u(6)^2\,u(8)-u(6)^4\,u(8)+u(6)^3\,u(8)^2=0$, that is, $X(u_7, u_8)$ must lie on a circle. As a simple calculation shows, this circle is concentric with the circumcircle, thus rediscovering Steiner theorem. webDiscovery does not return any linguistic statement

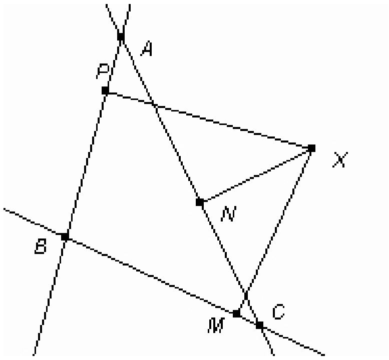
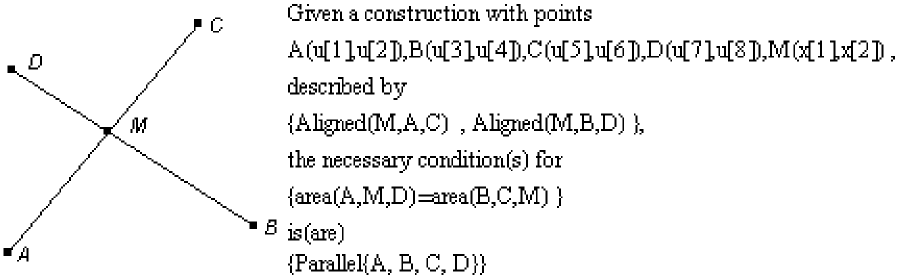


**Fig. 4.** Construction for discovering Simson theorem

since there are none in the database involving the free points $A, B, C, X$ that matches the equation.

Example 2. *Given a quadrilateral $ABCD$ and the intersection points of its diagonals, $M$, a necessary condition for the equality of the areas of triangles $AMD$ and $BCM$ is the parallelism of sides $AB$ and $CD$.* (Fig. 5)



```
Given a construction with points
A(u[1],u[2]),B(u[3],u[4]),C(u[5],u[6]),D(u[7],u[8]),M(x[1],x[2]) ,
described by
{Aligned(M,A,C) , Aligned(M,B,D) },
the necessary condition(s) for
{area(A,M,D)=area(B,C,M) }
is(are)
{Parallel(A, B, C, D)}
```

**Fig. 5.** Construction and discovery result for Example 2

### 4.2 Loci Search

Except for the most simple cases, such as lines, circles and perhaps the conics, visualizing loci is a hard task due to common difficulties when considering various objects with different movements. Although there exist dynamic geometry programs that can plot loci, they are restricted to a narrow class of loci and they can not in general find the analytic expression of loci [2].

Example 3. Recalling Steiner theorem, where only an equation was returned as the result of discovery, webDiscovery is able to return more specific knowledge. Using $A$ and $B$ to define a rectangular coordinate system, the strategy used substitutes all free points symbolic coordinates, except those of the locus point (which must be described as such), by their numeric values. Thus, the equation just involves the coordinates of the locus point. This equation is then linguistically interpreted as a circle, and the Mathematica abilities to plot implicit equations are used, as shown in Fig. 6. It must be noted that when using the locus discovery option, no general discovery is performed. For this theorem, the point $C$ is not a general one, but a fixed point in the plane.

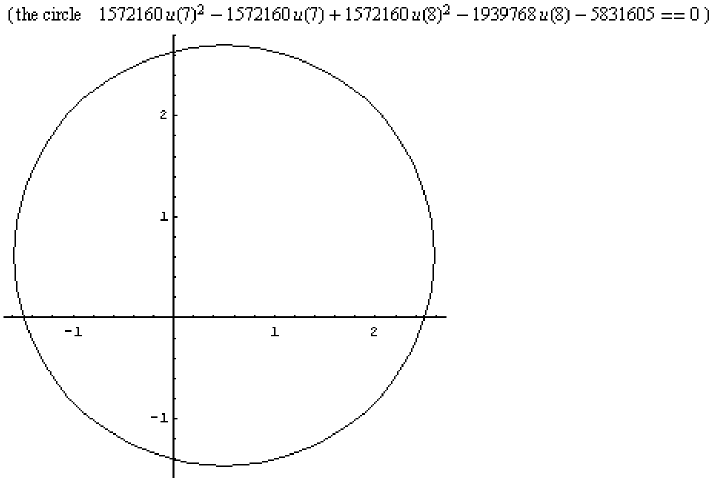
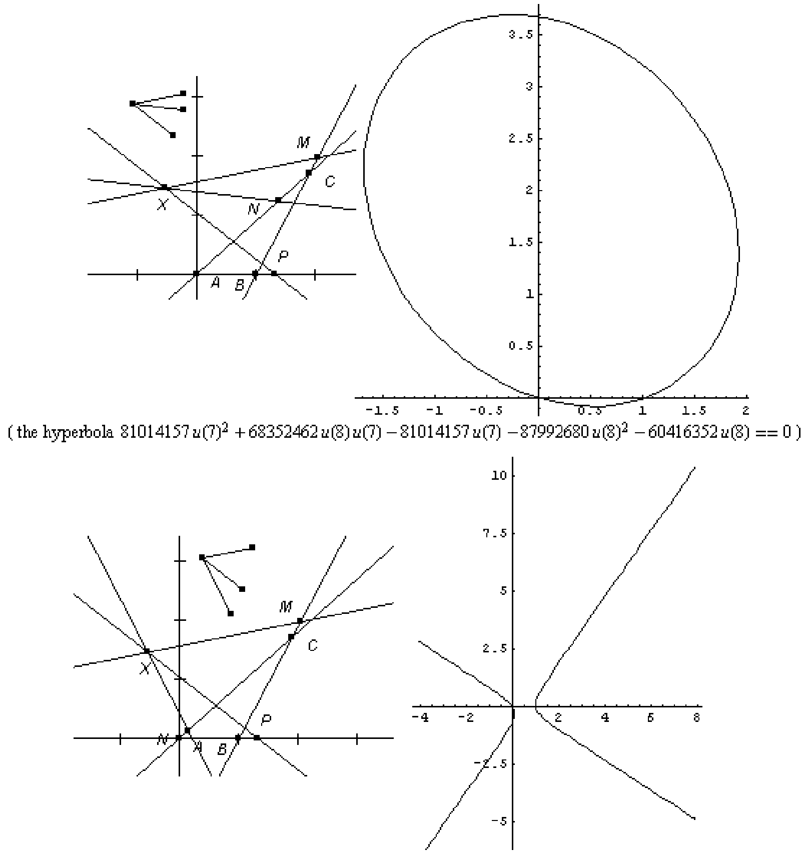


**Fig. 6.** The locus of $X$ such that the oriented area of triangle $MPN$ is 1

Redefining points $M, N, P$ to be the projections on the triangle sides along three directions, not all three equal, nor parallel to the respective sides, the locus of points $X$, such that the oriented area of triangle $MNP$ remains constant, is a general conic, a result recently proved in [10]. Fig. 7 shows the constructions and the obtained loci, an ellipse and a hyperbola.

Example 4 (A four bar linkage). Consider a mechanism with four linked bars: the foundation link $AC$ (taken as unit segment), the input or driver link $AB$, the output or driven link $CD$, and the connecting link $BD$ (Fig. 8, left). A common way to test the behavior of this linkage consists of obtaining the locus of the midpoint $X$ of the connecting link. webDiscovery plots the locus of $X$ as shown in Fig. 8, right, and it finds that this curve is the sextic

$$\frac{4563}{128000} - \frac{15483\,u(11)}{32000} + \frac{246529\,u(11)^2}{96000} - \frac{508\,u(11)^3}{75} + \frac{373\,u(11)^4}{40} - \frac{19\,u(11)^5}{3} + \frac{5\,u(11)^6}{3} + \frac{3087\,u(12)^2}{6400} - \frac{308\,u(11)\,u(12)^2}{75} + \frac{679\,u(11)^2\,u(12)^2}{60} - \frac{38\,u(11)^3\,u(12)^2}{3} + 5\,u(11)^4\,u(12)^2 + \frac{239\,u(12)^4}{120} - \frac{19\,u(11)\,u(12)^4}{3} + 5\,u(11)^2\,u(12)^4 + \frac{5\,u(12)^6}{3} = 0$$



( the ellipse $244701\,u(7)^2+102726\,u(8)\,u(7)-244701\,u(7)+224060\,u(8)^2-823023\,u(8)==0$ )

( the hyperbola $81014157\,u(7)^2+68352462\,u(8)\,u(7)-81014157\,u(7)-87992680\,u(8)^2-60416352\,u(8)==0$ )

**Fig. 7.** Two cases in a generalization of Simson Steiner theorems

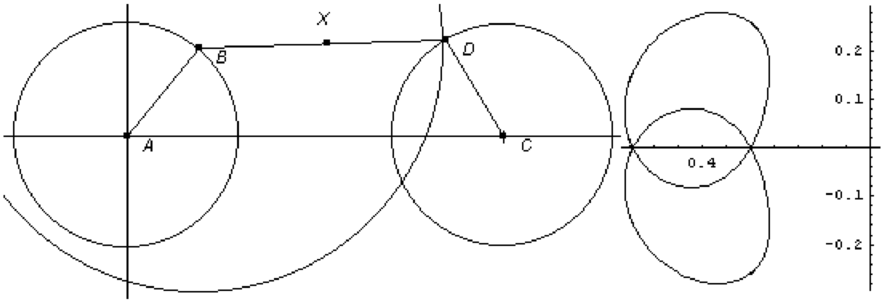


**Fig. 8.** A four bar linkage and the locus of $BD$ midpoint

### 4.3 Limitations

There are several known limitations in webDiscovery. On a side, just a limited number of natural language statements (mainly dealing with equality, collinearity and cocircularity of points, and parallelism/perpendicularity of lines) are managed in the matching process of equations, when not doing loci discovery. Furthermore, currently unavoidable simplification processes in the symbolic routines sometimes impede the expression of the discovered conditions in natural language. The diagonals of a quadrilateral meet at their midpoint if it is a parallelogram, but only one condition of parallelism is returned using the `Parallel` predicate, the other remaining in equational form.

On the other side, the proposed method is complete only in an algebraically closed field containing $\mathbf{Q}$, $\mathbf{C}$ for instance [14]. Nevertheless, the method works for a surprisingly vast amount of Euclidean constructions. Additionally, some tricks have been implemented. Consider for example a discovery involving the zeros of $x^2 + y^2$. Although $x^2 + y^2 = 0 \to x = 0, y = 0$ is not a theorem in $\mathbf{C}$, webDiscovery will return *the point* $(0, 0)$. Another source of imprecision in the system is due to the internal substitution of expressions involving distances by their square values, in order to keep the analytic knowledge in polynomial form, a prerequisite for using CoCoA. This decision will sometimes produce incorrect results.

## 5 Conclusion and Further Work

A web–based system for carrying out automatic discovery in the domain of plane Euclidean geometry has been described. webDiscovery is particularly well suited for obtaining loci of points that describe algebraic curves, a difficult task in other systems for dynamic geometry or of parametric CAD. Ongoing research is being conducted to extend its domain for dealing with non polynomial equations and inequalities, thus allowing a consideration of new relations between geometric elements.

## References

1. Botana, F., Valcarce, J.L.: A dynamic-symbolic interface for geometric theorem discovery. *Computers and Education*, 38(1-3), 21–35 (2002)
2. Botana, F.: Interactive versus symbolic approaches to plane loci generation in dynamic geometry environments. *Lecture Notes in Computer Science*, 2330, 211-218 (2002)
3. Botana, F., Valcarce, J.L.: A software tool for the investigation of plane loci. (submitted to *Mathematics and Computers in Simulation*)
4. Buchberger, B.: Groebner bases: an algorithmic method in polynomial ideal theory. In N.K. Bose, *Multidimensional systems theory*, Reidel, Dordrecht, 184-232 (1985)
5. Capani, A., Niesi, G., Robbiano, L.: CoCoA, a system for doing Computations in Commutative Algebra. Available via anonymous ftp from: `cocoa.dima.unige.it`
6. Chou, S.C.: *Mechanical Geometry Theorem Proving*. Reidel, Dordrecht (1988)

7. Chou, S.C., Gao, X.S., Zhang, J.Z.: *Machine Proofs in Geometry.* World Scientific, Singapore (1988)
8. Gao, X.S., Zhang, J.Z., Chou, S.C.: *Geometry Expert.* Nine Chapters, Taiwan (1998)
9. Gerlentner, H., Hansen, J.R., Loveland, D.W.: Empirical explorations of the geometry theorem proving machine. *Proc. West. Joint Computer Conf.*, 143–147 (1960)
10. Guzmán, M.: An extension of the Wallace–Simson theorem: projecting in arbitrary directions. *American Mathematical Monthly*, 106(6), 574–580 (1999)
11. Hanna, G.: Proof, explanation and exploration: an overview. *Educational Studies in Mathematics*, 44(1–2), 5–23 (2002)
12. Jackiw, N.: *The Geometer's Sketchpad v 4.0.* Key Curriculum Press, Berkeley (2002)
13. Kapur, D.: Using Groebner bases to reason about geometry problems. *Journal of Symbolic Computation*, 2, 399–408 (1986)
14. Kapur, D.: A refutational approach to geometry theorem proving. *Artificial Intelligence*, 37, 61–93 (1988)
15. Kapur, D., Mundy, J.L.: Wu's method and its application to perspective viewing. *Artificial Intelligence*, 37, 15–36 (1988)
16. King, J., Schattschneider, D.:*Geometry Turned On.* MAA, Washington (1997)
17. Laborde, J. M., Bellemain, F.: *Cabri Geometry II.* Texas Instruments, Dallas (1998)
18. Laborde, J.M., Straesser, R.: Cabri Géomètre, a microworld of geometry for guided discovery learning. *Zentralblatt für Didaktik der Mathematik*, 22(5), 171–177 (1990)
19. Laborde, C.: Dynamic geometry environments as a source of rich learning contexts for the complex activity of proving. *Educational Studies in Mathematics*, 44(1–2), 151–161 (2002)
20. Nevins, A.J.: Plane geometry theorem proving using forward chaining. *Artificial Intelligence*, 6, 1–23 (1975)
21. Recio, T., Vélez, M. P.: Automatic discovery of theorems in elementary geometry. *Journal of Automated Reasoning*, 23, 63–82 (1999)
22. Reiter, R.: A semantically guided deductive system for automatic theorem proving. *IEEE Transactions on Computers*, C–25(4), 328–334 (1976)
23. Richter–Gebert, J., Kortenkamp, U.: *The Interactive Geometry Software Cinderella.* Springer, Berlin (1999)
24. Valcarce, J.L., Botana, F.: webREX. Available from: `http://rosalia.uvigo.es/sdge/web/2D/webREXDemo.zip`
25. Wu, W. T.: *Mechanical Theorem Proving in Geometries.* Springer, Vienna (1994)