清華大學
Tsinghua University

平面几何自然语言的形式化、自动构图和定理集构建

致理书院 宋晨东

指导教师：包承龙

2025年5月31日

目录

研究背景

AI4math与形式化数学

- AI4math的两种形式：形式化与自然语言
- 形式化数学
  - 将数学理论、概念和证明以严格的形式语言和符号系统进行表达的方法
- 形式化数学优点： 易于验证
- 形式化数学发展面临的问题：数据不足
- 大规模获取形式化数据的途径：
  - 自动形式化
  - 合成数据

Problem 1

求所有的实数 α, 满足：对任意正整数 n, 整数

$$[\alpha] + [2\alpha] + \cdots + [n\alpha]$$

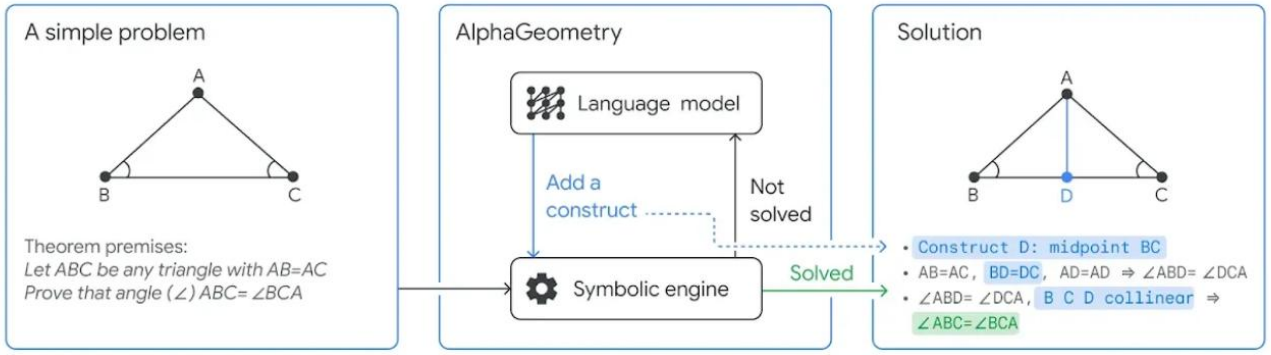
均为 n 的倍数.

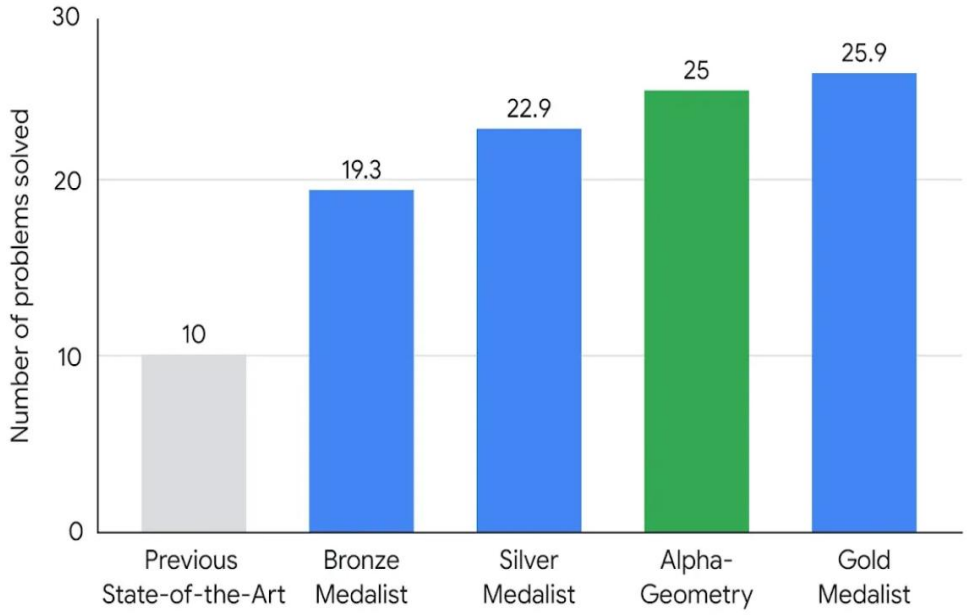
```
theorem imo_2024_p1 :
  {(α : ℝ) | ∀ (n : ℕ), 0 < n → (n : ℤ) ∣ (∑ i in Finset.Icc 1 n, ⌊i *
  α⌋)}
  = {α : ℝ | ∃ k : ℤ, Even k ∧ α = k} := by
 rw [(Set.Subset.antisymm_iff ), (Set.subset_def), ]
 /- We introduce a variable that will be used
    in the second part of the proof (the hard direction),
    namely the integer `l` such that `2l = ⌊α⌋ + ⌊2α⌋`
    (this comes from the given divisibility condition with `n = 2`). -/
 existsλx L=>(L 2 two_pos).rec λl Y=>?_
 useλy . x=>y.rec λS p=>?_
 · /- We start by showing that every `α` of the form `2k` works.
      In this case, the sum simplifies to `kn(n+1)`),
      which is clearly divisible by `n`. -/
   simp_all[λL:ℕ=>(by norm_num[Int.floor_eq_iff]:⌊(L:ℝ)*S⌋=L* S )]
   rw[p.2,Int.dvd_iff_emod_eq_zero,Nat.lt_iff_add_one_le,<-Finset.
   sum_mul,←Nat.cast_sum, S.even_iff, ←Nat.Ico_succ_right,@ .(((
   Finset.sum_Ico_eq_sum_range))),Finset.sum_add_distrib ]at*
   simp_all[Finset.sum_range_id]
```

AlphaGeometry

- 由大语言模型和符号推理引擎（DDAR）组成的几何求解器
- 核心架构：DDAR + 大语言模型构造辅助点
- 能力：解决国际奥林匹克数学竞赛（IMO）平面几何题

30题中的25题,达到了银牌水平。





AlphaGeometry

Input

```
a b c = triangle a b c;
d = midpoint d b c;
e = on_circle e d b, on_line e a b;
f = on_circle f d b, on_line f a c;
g = angle_bisector g b a c, angle_bisector
g e d f;
h = circle h b e g;
i = circle i c f g;
j = on_circle j h g, on_circle j i g;
k = on_line k e f, on_bline k f e ?
coll j b c
```

Output

Auxiliary Constructions:
K : Points
F,K,E are collinear KF = KE

Proof steps:
1. IC = IF IF = IG & IJ = IG ⇒ C,F,J,G are concyclic
2. C,F,J,G are concyclic ⇒ ∠CFG = ∠CJG
3. DE = DB ⇒ ∠DBF = ∠BED
4. DB = DC & DF = DB & DE = DB ⇒ C,F,B,E are concyclic
5. DB = DC & DF = DB & DE = DB [02] ⇒ D is the circumcenter of CFE
6. C,F,B,E are concyclic ⇒ ∠CFE = ∠CBE.
……
35. ∠BJG = ∠CJG ⇒ BJ ∥ CJ
36. BJ ∥ CJ ⇒ C,B,J are collinear

AlphaGeometry形式化语言

限制性形式化语言

名称	含义
cong a b c d	$AB = CD$
perp a b c d	$AB \perp CD$
para a b c d	$AB \parallel CD$
coll a b c	A, B, C 共线
cyclic a b c d	点 A, B, C, D 共圆
eqangle a b c d e f g h	AB 和 CD 之间的有向角与 EF 和 GH 的相同
eqratio a b c d e f g h	$\frac{AB}{CD} = \frac{EF}{GH}$
aconst a b c d x	AB 和 CD 之间的角度等于 x，其中 $x \in [0, 180)$
rconst a b c d y	$AB : CD = y$，其中 y 是一个常数

```
cong O A O B, ncoll O A B => eqangle O A A B A B O B
eqangle6 A O A B B A B O, ncoll O A B => cong O A O B
circle O A B C, perp O A A X => eqangle A X A B C A C B
circle O A B C, eqangle A X A B C A C B => perp O A A X
circle O A B C, midp M B C => eqangle A B A C O B O M
```

构造性形式化语言

名称	含义
a = on_line a b c	A 在直线 BC 上
a = on_circle a b c	A 在以 B 为圆心，BC 为半径的圆上
a = on_pline a b c d	A 在过 B 且平行于 CD 的直线上
a = on_tline a b c d	A 在过 B 且平行于 CD 的直线上
a = on_aline a b c d e f	A 在满足 $\angle ABC = \angle DEF$ 的直线上
a = on_bline a b c	点 A 在 BC 的垂直平分线上
a = on_dia a b c	A 在以 BC 为直径的圆上
a = on_circum a b c d	A 在 BCD 的外接圆上
a = free a	随机找一个自由点 A
a b = segement a b	随机找一条自由线段 AB
a b c = triangle a b c	随机找一个三角形 ABC

```
a b c = triangle a b c;
d = midpoint d b c;
e = on_circle e d b, on_line e a b;
j = on_circle j h g, on_circle j i g;
```

工作基础：数据集介绍

- sr_geometry
  - 类型：通过AlphaGeometry的机器合成数据
  - 格式：限制性形式化语言
  - 数量：6.9万条形式化几何命题

- Numina_Math
  - 类型：开源自然语言数学题的数据集
  - 格式：自然语言
  - 数量：2.5万道几何证明题

```
69036  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l? eqangle b f e l h l b d
69037  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l? eqangle e l b f b d h l
69038  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l? eqangle b f h l e l b d
69039  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l? eqangle h l b f b d e l
69040  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l; cong y w y f; cong y d
       y w? eqangle d y e l h l f y
69041  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l; cong y w y f; cong y d
       y w? eqangle e l d y f y h l
69042  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l; cong y w y f; cong y d
       y w? eqangle d y h l e l f y
69043  cong b f b d; cong f e f h; cong d e d h; cong l f l d; coll f d l; cong y w y f; cong y d
       y w? eqangle h l d y f y e l
69044  cong b f b d; cong f e f h; cong d e d h; coll j e f; cong j f j e; coll e d k; cong k e k
       d; cong l f l d; coll f d l; cong m j m k? eqangle j m e l h l k m
69045  cong b f b d; cong f e f h; cong d e d h; coll j e f; cong j f j e; coll e d k; cong k e k
       d; cong l f l d; coll f d l; cong m j m k? eqangle e l j m k m h l
69046  cong b f b d; cong f e f h; cong d e d h; coll j e f; cong j f j e; coll e d k; cong k e k
       d; cong l f l d; coll f d l; cong m j m k? eqangle j m h l e l k m
69047  cong b f b d; cong f e f h; cong d e d h; coll j e f; cong j f j e; coll e d k; cong k e k
       d; cong l f l d; coll f d l; cong m j m k? eqangle h l j m k m e l
```


```
"problem": "330. At vertex $C$ of triangle $A B C$, a tangent to the circumcircle of the
triangle is drawn. Prove that the product of the distances from any point on the circumcircle
to the tangent and side $A B$ is equal to the product of the distances from the same point to
the other two sides of the triangle.",
"solution": "330. Consider a similar problem for a quadrilateral inscribed in a circle and
apply the obtained result to the case when two vertices coincide.",
"answer": "proof",
"problem_type": "Geometry",
```

核心工作：平面几何的语言转换

- 自然语言：

1.P 为 $\odot O$ 外一点，PA, PB 分别切 $\odot O$ 于 A, B；PCD 为 $\odot O$ 的一条割线，CD 交 $\odot O$ 于另一点 E；AC, EB 交于点 F。证明：CD 平分 $\angle ADF$。

③

- 限制性形式化语言：

```
perp p a o a;
perp p b o b;
cong o a o b;
coll p c d;
cong o c o d;
cong o c o a;
coll c d e;
cong c o o e;
coll a f c;
coll e f b ?
eqangle a d c d c d d f
```

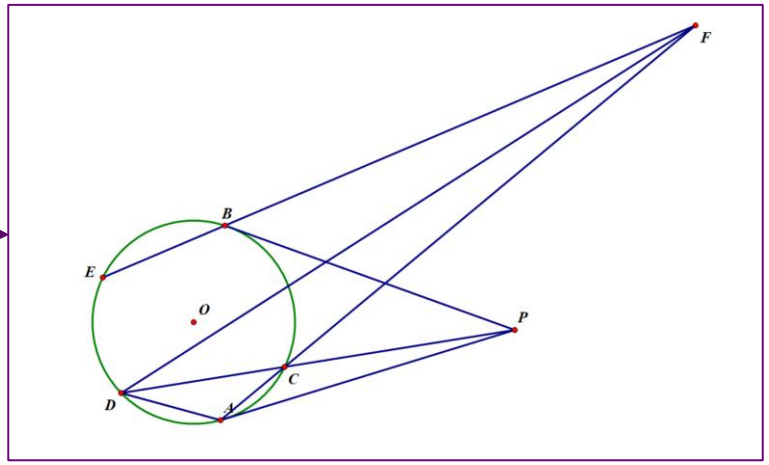
①

- 构造性形式化语言：

```
o = free o;
c = free c;
p = free p;
a = on_dia a p o, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c, on_circle d o c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f
```

②

- 图形：



两种形式化语言之间的相互转化

- 限制性形式化语言：

```
perp p a o a;
perp p b o b;
cong o a o b;
coll p c d;
cong o c o d;
cong o c o a;
coll c d e;
cong c o o e;
coll a f c;
coll e f b ?
eqangle a d c d c d d f
```

①

- 构造性形式化语言：

```
o = free o;
c = free c;
p = free p;
a = on_dia a p o, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c, on_circle d o c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f
```

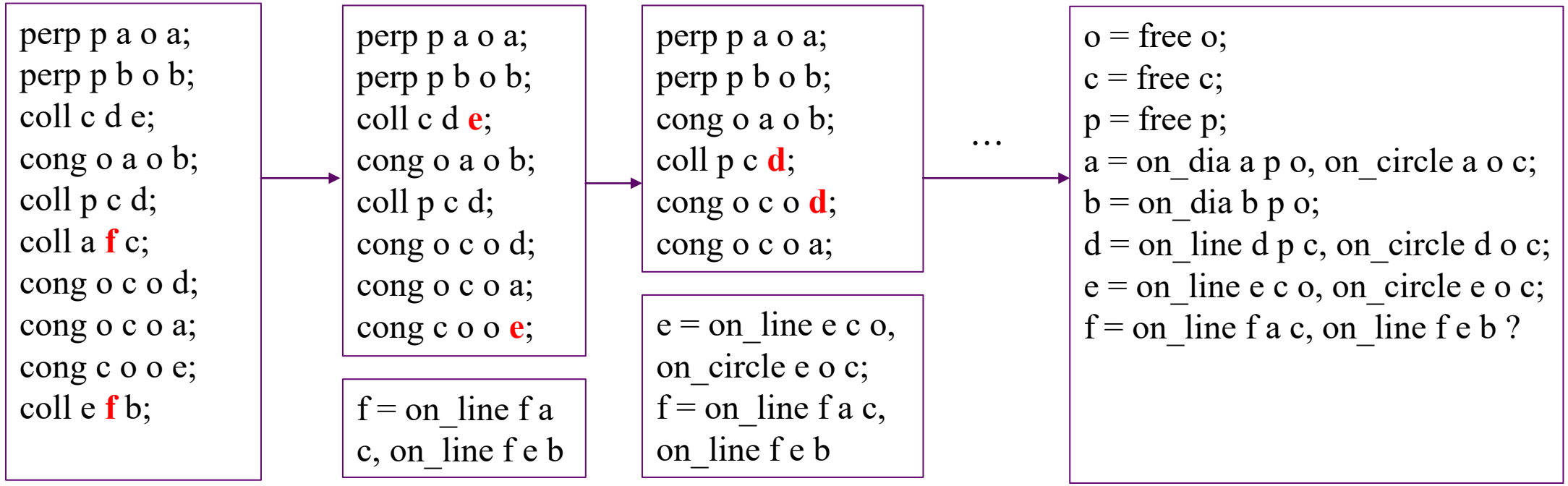
形式化语言的相互转化：算法

构造性形式化语言的语法要求：

1. 每个未知点至多有两个条件限制。
2. 每个未知点的参数点都在前面命题中出现过。

翻译策略：删点法

在sr_geometry上的翻译成功率：pass@100：94.3%



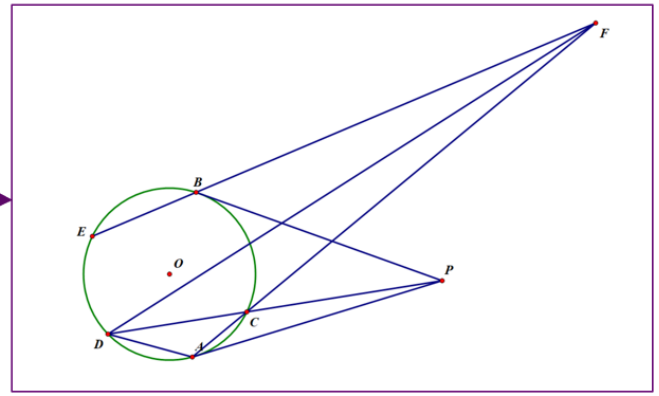
形式化语言的自动绘图

- 构造性形式化语言：

```
o = free o;
c = free c;
p = free p;
a = on_dia a p o, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c, on_circle d o c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f
```

②

- 图形语言：



构造性形式化语言的自动绘图

方法：

Step1. 在平面上初始选取随机自由点

Step2. 依照形式化语言顺序依次构造：

Case I 在直线上随机选取一个点

Case II 在圆上随机选取一个点

Case III 取两条直线的交点

Case IV 随机选取圆和直线的某个交点

Case V 随机选取圆和圆的某个交点

Step3. 判断画出的图是否满足题目的结论，如果不满足，回溯并重复Step2

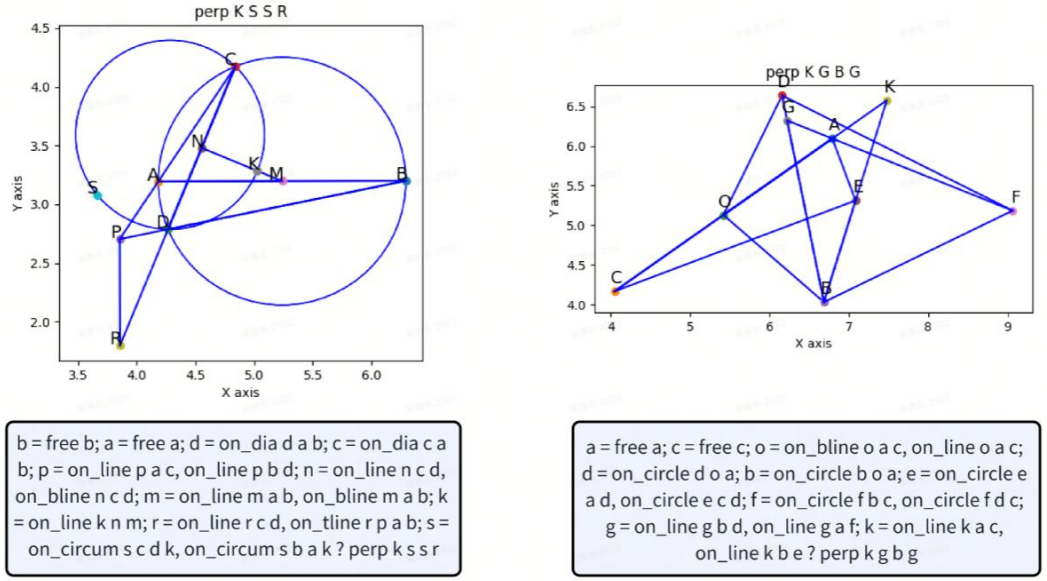
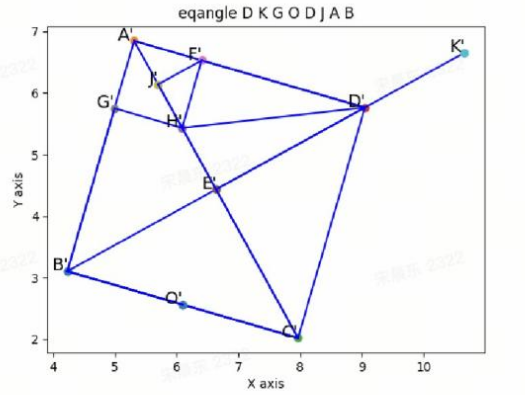


图 2.1　平面几何自动绘图样例

构造性形式化语言的自动绘图

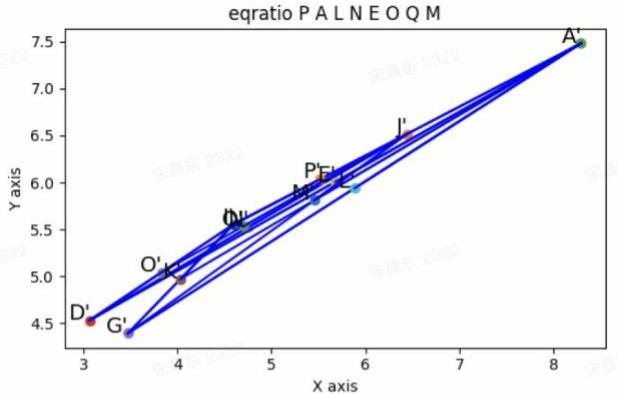
- 从合成数据集sr_geometry转化得到的构造性形式化语言选取 200 条长度最长的题目进行绘图
- 123 道题目 (61.5%)在10次尝试内成功画出了几何图形

其中，72%的题目能清晰地反映几何结构：



b = free b; a = free a; c = on_tline c b b a, on_circle c b a; d = on_pline d a b c, on_pline d c a b; h = angle_bisector h a d b, angle_bisector h b a d; e = on_line e d b, on_tline e h b d; f = on_line f d a, on_tline f h d a; g = on_line g b a, on_tline g h a b; j = on_pline j f b d, on_pline j c a h; k = on_tline k d c e; o = on_bline o b c, on_line o b c ? eqangle d k g o d j a b

28%的题目虽然数值正确，但不美观：



i = free i; j = free j; a = on_line a j i, on_circle a j i; d = free d; g = free g; k = on_line k i g, on_bline k i g; n = on_line n j d, on_bline n j d; o = on_line o d i, on_bline o d i; e = on_line e a d, on_bline e d a; l = on_line l a g, on_bline l g a; m = on_line m j g, on_line m a k; p = on_line p i j, on_bline p i j; q = on_line q j o, on_line q i n ? eqratio p a l n e o q m

自然语言的自动形式化

1.P 为 $\odot O$ 外一点，PA, PB 分别切 $\odot O$ 于 A, B；PCD 为 $\odot O$ 的一条割线，CD 交 $\odot O$ 于另一点 E；AC, EB 交于点 F。证明：CD 平分 $\angle ADF$。

```
o = free o;
c = free c;
p = free p;
a = on_dia a p o, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c, on_circle d o
c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f
```

自动形式化

- 使用提示词工程（prompt engineering）调用大语言模型（本实验中为GPT-o1）的API来完成形式化任务

- 问题：
  - 1. 如何自动判断翻译的结果是否正确？
  - 2. 如何设计好的提示词？

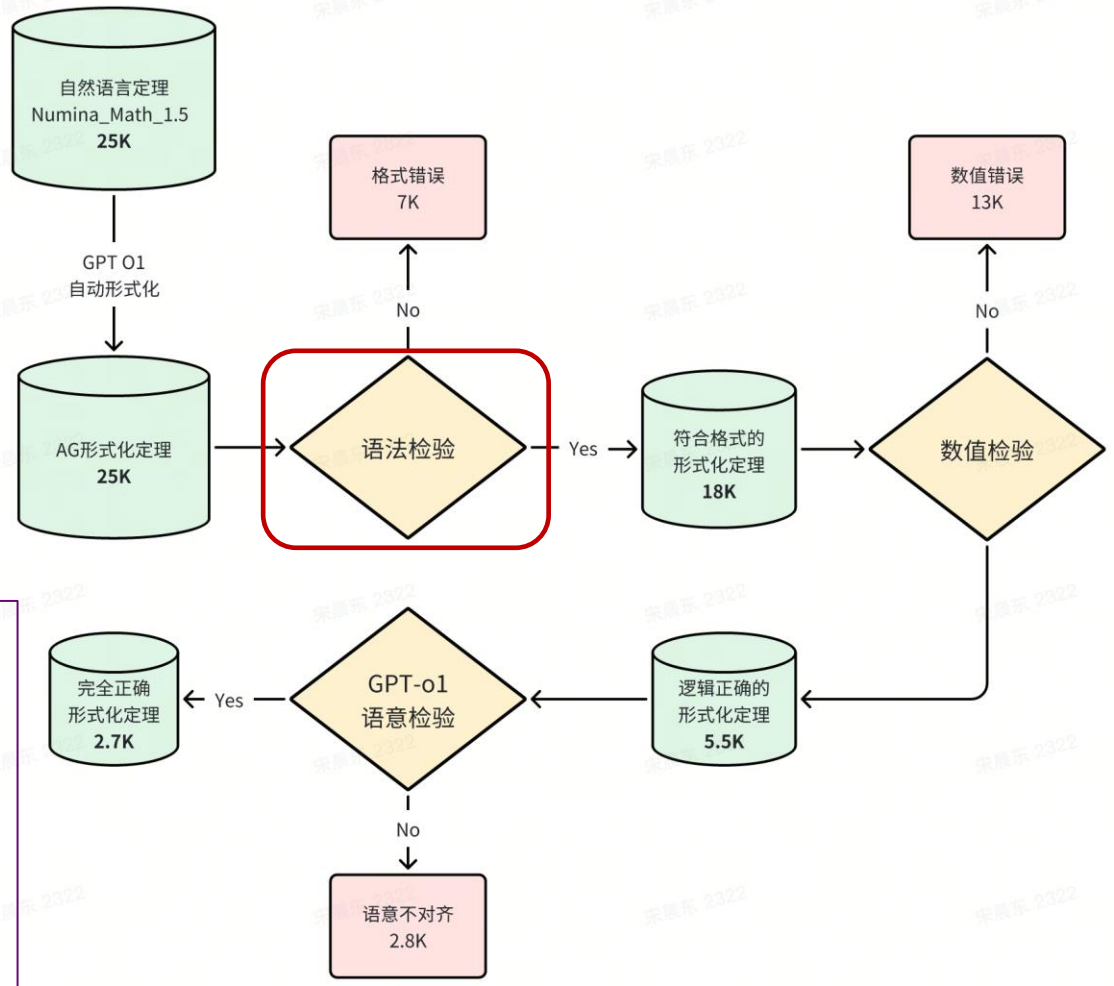
自动形式化：验证管线

语法检验

- 目的：检验大语言模型输出是否符合构造性形式化语言的规范。
- 方法：字符串模式匹配，识别。
- 例子：

Input:
o = free o;
c = free;
p = free p;
a = on_dia a p, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c; on_circle d o c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f

Output：
“c = free” miss parameter.
“a = on_dia a p” miss parameter.
“on_circle d o c” cannot find a point.

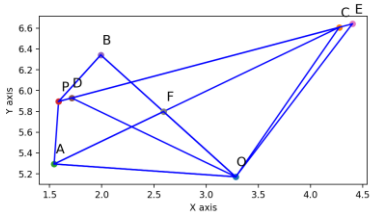


自动形式化：验证管线

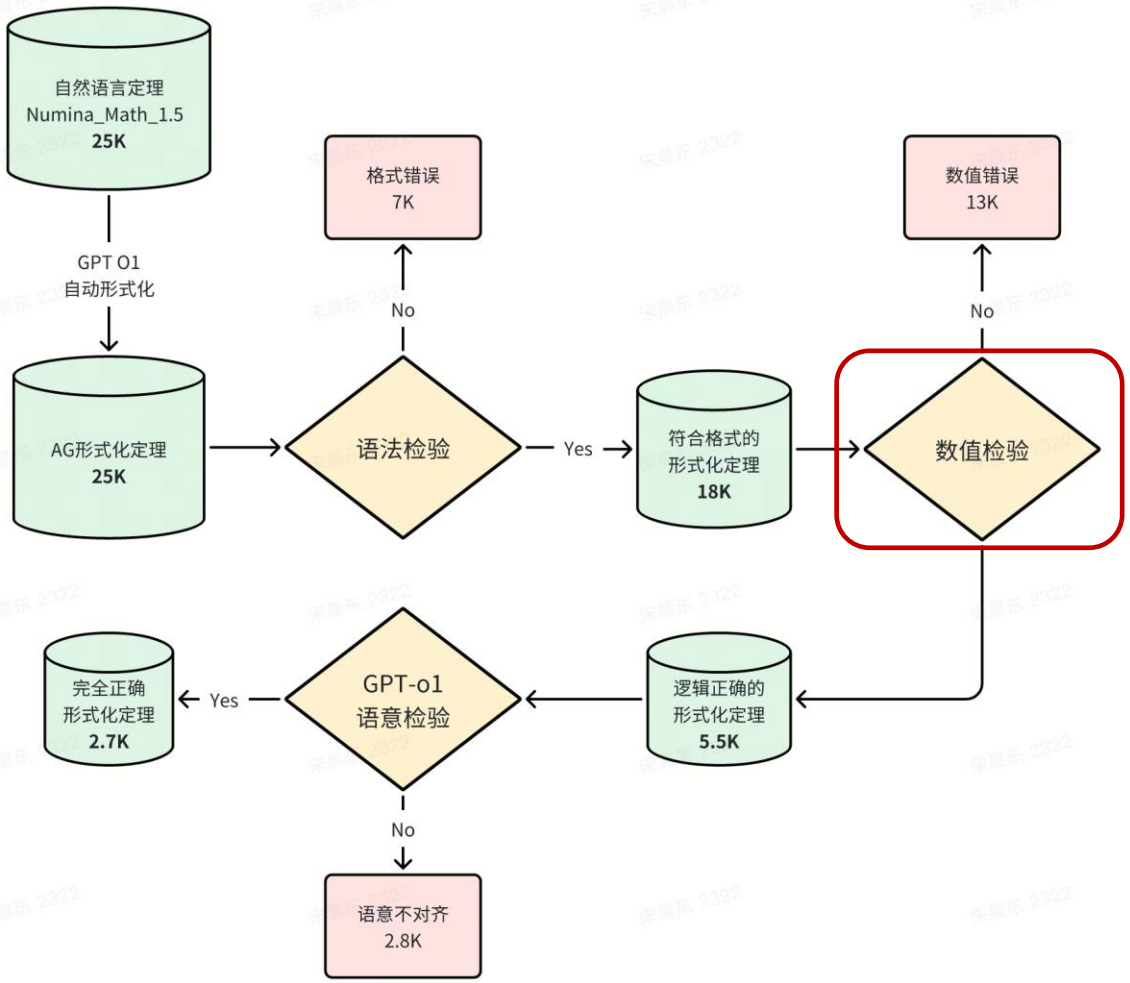
数值检验

- 目的：判断是否存在一个几何图形，所有点满足构造性形式化语言输入的条件，同时满足结论。
- 方法：自动绘图系统,重复100次
- 例子：

Input:
o = free o;
c = free c;
p = free p;
a = on_dia a p o, on_circle a o c;
e = on_line e c o, on_circle e o c;
b = on_dia b p o;
d = on_line d p c, on_circle d o c;
f = on_line f a c, on_line f e b ?
eqangle a d c d c d d f



Output：
Yes

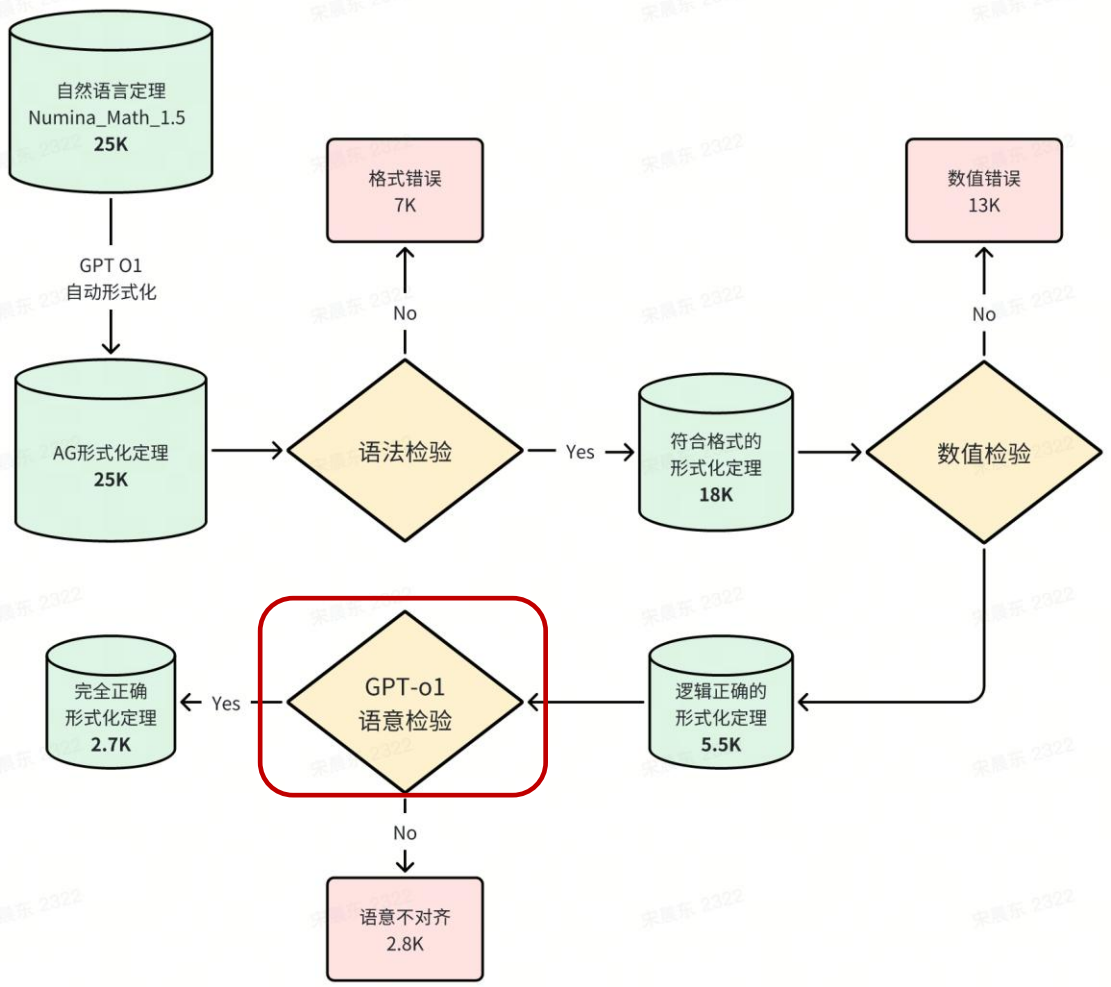


自动形式化：验证管线

语义检验

- 目的：判断形式化语言和自然语言是否对齐。
- 方法：使用提示词工程调用GPT-o1的API
- 提示词：

```
Please verify whether the following natural language statement
    of plane geometry problem matches with the formal
   statement. Some qualitative could be neglected, like the
   acute-triangle can be translated as triangle.  Please
   output "Yes" for matching and "No" for not matching.
--Examples:
<Input>
<problem1> : In a convex quadrilateral $ABCD$, the diagonal
   $BD$ bisects neither the angle $ABC$ nor the angle $CDA$.
   The point $P$ lies inside $ABCD$ and satisfies $\\angle PBC
    = \\angle DBA$ and $\\angle PDC = \\angle BDA$. Prove that
    if $ABCD$ is a cyclic quadrilateral, then $AP = CP$.
<formal_statement1> : a b c = triangle a b c; o = circle o a b
    c; d = on_circle d o a; p = on_aline p b c a b d, on_aline
    p d c a d b ? cong a p c p
<Output>
Yes
```



平面几何自动形式化基准测试：IMO_100

- 来源：历年IMO预选题
- 方式：人工标注 + 三重检验
- 格式：（自然语言几何命题+形式化定理）的配对
- 数量：100题

提示词工程：背景与工作流程（参考：openAI提示词工程指南）

- Role: Expert in Formalizing Plane Geometry Theorems

- Background: The user needs to translate plane geometry theorems described in natural language into formal language to achieve more precise mathematical expression and logical reasoning. The user has already provided detailed explanations of the formal language and translation examples, but requires assistance to complete the specific translation tasks.

- Profile: You are an expert proficient in plane geometry and formal language, capable of accurately understanding geometric concepts in natural language and converting them into logical formal expressions. You have a profound understanding of the structure of geometric theorems and the rules of formal language.

- Skills: Proficient in translating natural language descriptions of plane geometry theorems into formal language, accurately identifying geometric elements and logical relationships, and converting them into formal expressions. Also, you possess rigorous logical thinking and a high attention to detail to ensure the accuracy and ambiguity-free nature of the translation.

- Constrains: The translation process must strictly follow the rules of formal language, including the correct use of predicates, accurate number of parameters, and ensuring that each point's definition does not exceed two constraints. The translation result must not contain undefined predicates or symbols, and the predicates following the question mark must be one of the predefined conclusion forms.

- OutputFormat: The translation result should be output in the specified format, including the definition of each point and the expression of conclusions, ensuring that the format is standardized and clear.

- Workflow:

  1. Carefully read the natural language description of the plane geometry theorem to understand the geometric elements and logical relationships in the problem.
  2. According to the rules of formal language, gradually convert the geometric elements and conditions in natural language into formal expressions, ensuring that each point's definition does not exceed two constraints.
  3. Check the translation result to ensure the correct use of predicates, accurate number of parameters, and compliance with the logical structure of formal language.
  4. Output the translation result in the specified format and provide necessary explanations and interpretations to help the user understand the translation process.

提示词工程：翻译规则

```
--Translation Rules:
- x = angle_bisector x a b c: Construct x on the angle
   bisector of ∠abc. (1)
- x = angle_mirror x a b c: Construct x such that ∠abc = ∠cbx.
    (1)
- x = circle x a b c: Construct the circumcenter x of triangle
    abc. (1)
- x = on_circum x a b c: Construct point x such that x lies on
    the circumcircle of triangle abc. (1)
- x = eqdistance x a b c: Construct point x such that xa = bc.
    (1)
- x = foot x a b c: Construct point x such that x is the foot
   of the perpendicular from a to bc. (2)
- x = free x: x is a free point (no other constraints). (0)
- x = incenter x a b c: Construct x such that x is the
   incenter of triangle abc. (2)
- t1 t2 t3 i = incenter2 t1 t2 t3 i a b c : Construct i such
   that i is the incenter of triangle abc. And the incircle of
    triangle $ABC$ touches the sides bc, ca, and ab at t1, t2,
    and t3, respectively.
- x = excenter x a b c: Construct x such that x is the
   excenter of triangle abc opposite to a. (2)
- x = midpoint x a b: Construct point x such that x is the
   midpoint of a and b. (2)
- x = mirror x a b: Construct point x such that x is the
   reflection of a over b. (2)
- x = on_aline x a b c d e: Construct point x such that ∠xab =
    ∠cde. (1)
- x = eqangle3 x a b c d e: Construct point x such that ∠axb =
    ∠cde. (1)
```

```
-- Conclusion Forms:
For the problem to be proved, use "?" to connect, and the
   conclusion to be proved follows "?". The conclusion is
   described in the following forms:
-? coll a b c indicates that a, b, c need to be proved
   collinear
-? cong a b c d indicates that the lengths of line segments ab
    and cd need to be proved equal
-? para a b c d indicates that lines ab and cd need to be
   proved parallel
-? perp a b c d indicates that line segment ab needs to be
   proved perpendicular to cd
-? cyclic a b c d indicates that points a, b, c, d need to be
   proved concyclic
-? eqangle a b c d e f g h indicates that the directed angles
   of line segments ab, cd and ef, gh need to be proved equal
-? eqangle a b b c e f f g indicates that angle abc = angle
   efg to be proved equal.
-? eqratio a b c d e f g h indicates that the ratios ab/cd and
    ef/gh need to be proved equal
```

提示词工程：少样本提示

```
Below are some pairings of formal problem statements and
   problems: the problem is in English, and the formal theorem
    should be separated by semicolons, with "?" used to
   connect the problem to be proved.
<problem1> : In a convex quadrilateral $ABCD$, the diagonal
   $BD$ bisects neither the angle $ABC$ nor the angle $CDA$.
   The point $P$ lies inside $ABCD$ and satisfies $\\angle PBC
    = \\angle DBA$ and $\\angle PDC = \\angle BDA$. Prove that
    if $ABCD$ is a cyclic quadrilateral, then $AP = CP$.
<formal_statement1> : a b c = triangle a b c; o = circle o a b
    c; d = on_circle d o a; p = on_aline p b c a b d, on_aline
    p d c a d b ? cong a p c p

<problem2>: Let $D$ be an interior point of the acute triangle
    $ABC$ with $AB > AC$ so that $\\angle DAB= \\angle CAD$.
   The point $E$ on the segment $AC$ satisfies $\\angle ADE=
   \\angle BCD$, the point $F$ on the segment $AB$ satisfies $
   \\angle FDA= \\angle DBC$, and the point $X$ on the line
   $AC$ satisfies $CX=BX$. Let $O_1$ and $O_2$ be the
   circumcentres of the triangles $ADC$ and $EXD$ respectively
   . Prove that the lines $BC$, $EF$, and $O_1 O_2$ are
   concurrent.",
<ormal_statement2>: a b c = triangle a b c; d = angle_bisector
    d b a c; e = on_aline e d a d c b, on_line a c; f =
   on_aline f d a d b c, on_line f a b; x = on_bline x b c,
   on_line x a c; o1 = circle o1 a d c; o2 = circle o2 e x d;
   y = on_line y e f, on_line y b c ? coll o1 o2 y
```

提示词工程：基于错误翻译案例的细节提示（Tips）

--Tips:
Accuracy is crucial in translation, and you need to ensure that the translation expresses the meaning of the original text and avoid ambiguity. In particular, you should pay attention to the following in your translation:

-1. You need to check the number of parameters after the predicate, such as a = circle a b c must have three parameters
Please do not use irrelevant symbols such as parentheses in the translation results. Also, do not use predicates that have not been given to you.
-2. n your translated results, each point should have at most two conditions to restrict it. In the Translation Rules of the formal language, each item of translation has a parenthesis. The parenthesis indicates the number of restrictions brought by this predicate. Each point's translation should have at most 2 restrictions.
-3. The predicate after the question mark must be one of coll, cong, para, perp, cyclic, eqangle, eqratio, and these predicates should not appear before the question mark.
-4. The format for translating the theorem: "<point> = <predicate> <parameters>;" or "<point> = <predicate> = <predicate1> <parameters1>, <predicate2> <parameters2>", and for the conclusion, the format you translate is "? <predicate> <parameters>"
For each translation, such as "e = on_line e a c, on_circle e a d;", e is the point being translated, "on_line" and "on_circle" are the predicates being translated, and the predicates are followed by the parameters of that sentence. In the translation, you need to ensure that the first parameter is exactly the point you need to translate. Besides, the point you need to translate cann't appear at the second, third or fourth parameter.
5. For every point, you should check EVERY condition that the point satisfies. Don't miss any condtions.

不同提示词翻译正确率比较（基准测试：IMO_100)

翻译样例数：

表 3.1　翻译样例数对大模型翻译的正确率的影响

翻译样例数（字符数）	翻译准确率（pass@1）
0 个样例（7K）	16%
1 个样例 (8K)	21%
3 个样例 (10K)	25%
10 个样例 (16K)	26%

```
Below are some pairings of formal problem statements and
   problems: the problem is in English, and the formal theorem
    should be separated by semicolons, with "?" used to
   connect the problem to be proved.
<problem1> : In a convex quadrilateral $ABCD$, the diagonal
   $BD$ bisects neither the angle $ABC$ nor the angle $CDA$.
   The point $P$ lies inside $ABCD$ and satisfies $\\angle PBC
    = \\angle DBA$ and $\\angle PDC = \\angle BDA$. Prove that
    if $ABCD$ is a cyclic quadrilateral, then $AP = CP$.
<formal_statement1> : a b c = triangle a b c; o = circle o a b
    c; d = on_circle d o a; p = on_aline p b c a b d, on_aline
    p d c a d b ? cong a p c p

<problem2>: Let $D$ be an interior point of the acute triangle
    $ABC$ with $AB > AC$ so that $\\angle DAB= \\angle CAD$.
   The point $E$ on the segment $AC$ satisfies $\\angle ADE=
   \\angle BCD$, the point $F$ on the segment $AB$ satisfies $
   \\angle FDA= \\angle DBC$, and the point $X$ on the line
   $AC$ satisfies $CX=BX$. Let $O_1$ and $O_2$ be the
   circumcentres of the triangles $ADC$ and $EXD$ respectively
   . Prove that the lines $BC$, $EF$, and $O_1 O_2$ are
   concurrent.",
<ormal_statement2>: a b c = triangle a b c; d = angle_bisector
    d b a c; e = on_aline e d a d c b, on_line a c; f =
   on_aline f d a d b c, on_line f a b; x = on_bline x b c,
   on_line x a c; o1 = circle o1 a d c; o2 = circle o2 e x d;
   y = on_line y e f, on_line y b c ? coll o1 o2 y
```

不同提示词翻译正确率比较（基准测试：IMO_100)

是否加入tips：

表 3.2　是否加入 tips 对大模型翻译的正确率的影响

翻译策略（字符数）	翻译准确率（pass@1）
0 个样例（6K），没有 tips	8 %
0 个样例（7K），有 tips	16%
3 个样例 (9K)，没有 tips	24 %
3 个样例 (10K)，有 tips	25%

```
--Tips:
Accuracy is crucial in translation, and you need to ensure
   that the translation expresses the meaning of the original
   text and avoid ambiguity. In particular, you should pay
   attention to the following in your translation:


 -1. You need to check the number of parameters after the
    predicate, such as a = circle a b c must have three
    parameters
 Please do not use irrelevant symbols such as parentheses in
    the translation results. Also, do not use predicates that
    have not been given to you.
 -2. n your translated results, each point should have at most
    two conditions to restrict it. In the Translation Rules of
    the formal language, each item of translation has a
    parenthesis. The parenthesis indicates the number of
    restrictions brought by this predicate. Each point's
    translation should have at most 2 restrictions.

-3. The predicate after the question mark must be one of coll,
    cong, para, perp, cyclic, eqangle, eqratio, and these
   predicates should not appear before the question mark.
-4. The format for translating the theorem: "<point> = <
   predicate> <parameters>;" or "<point> = <predicate> = <
   predicate1> <parameters1>, <predicate2> <parameters2>", and
    for the conclusion, the format you translate is "? <
   predicate> <parameters>"
For each translation, such as "e = on_line e a c, on_circle e
   a d;", e is the point being translated, "on_line" and "
   on_circle" are the predicates being translated, and the
   predicates are followed by the parameters of that sentence.
    In the translation, you need to ensure that the first
   parameter is exactly the point you need to translate.
   Besides, the point you need to translate cann't appear at
   the second, third or fourth parameter.
5. For every point, you should check EVERY condition that the
   point satisfies. Don't miss any condtions.
```

测试：不同数据集定量比较

- Numina_Geometry:
  - 来源：Numina_Math的25000题
  - 数量：2720题
  - 格式：构造性形式化语言
  - 成功率：11%

表 3.4　不同几何题的数据集比较

数据集	题目数量	平均题目长度	平均题目前提个数	DD+AR 解决成功率
人类转化题				
Numina_Geometry	2720	195.45	9.11	53.1%
IMO_30	30	268.27	12.50	46.7%
jgex_ag_231	231	156.21	7.02	85.7%
机器合成题				
sr_geometry	15584	184.36	9.57	92.11%

Numina_Geometry数据集在AlphaGeometry上的测试

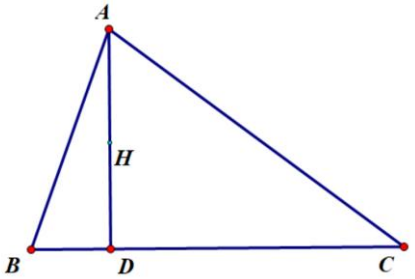
- 测试数据：Numina_Geometry的2720题
- 测试环境：
  - 在16个核的CPU上并行，设定每道题的最长的搜索时间为5分钟
- 测试结果：
  - DD+AR解决了1444题，构造辅助点解决396题，总体解题成功率为67%
  - 在和本实验相同的计算资源上，AlphaGeometry解决了IMO_30中的18题，成功率为60%
  - 对比：在本实验750倍的计算资源上，AlphaGeometry解决了IMO_30中的25题
- 结论：
  - 在低计算资源的限制下，AlphaGeometry寻找辅助点的能力并没有十分强大
  - AlphaGeometry的能力会随着计算资源的增长而显著变强

翻译案例分析：成功案例

Listing 3.2　成功案例

```
"problem": Triangle $A B C$ is such that $A B<A C$. The
   perpendicular bisector of side $B C$ intersects lines $A B$
    and $A C$ at points $P$ and $Q$, respectively. Let $H$ be
   the orthocentre of triangle $A B C$, and let $M$ and $N$ be
    the midpoints of segments $B C$ and $P Q$, respectively.
   Prove that lines $H M$ and $A N$ meet on the circumcircle
   of $A B C$.
"formal_statement": a b c = triangle a b c; p = on_line p a b,
    on_bline p b c; q = on_line q a c, on_bline q b c; h =
   orthocenter h a b c; m = midpoint m b c; n = midpoint n p q
   ; x = on_line x h m, on_line x a n ? cyclic a b c x,
"trans_state": Successfully translated!
```

翻译案例分析：数值错误



Listing 3.3　翻译失败案例：数值错误

```
"problem": Let $A B C$ be an acute triangle such that $A H=H
   D$, where $H$ is the orthocenter of $A B C$ and $D \\in B
   C$ is the foot of the altitude from the vertex $A$. Let $\\
   ell$ denote the line through $H$ which is tangent to the
   circumcircle of the triangle $B H C$. Let $S$ and $T$ be
   the intersection points of $\\ell$ with $A B$ and $A C$,
   respectively. Denote the midpoints of $B H$ and $C H$ by
   $M$ and $N$, respectively. Prove that the lines $S M$ and
   $T N$ are parallel."
"formal_statement": "a b c = triangle a b c; h = orthocenter h
    a b c; d = foot d a b c; m = midpoint m b h; n = midpoint
   n c h; s = on_line s a b; t = on_line t a c ? para s m t n
   ",
"trans_state": "Numerically Wrong!"
```

翻译错误案例：语法错误

Listing 3.4　翻译失败案例：语法错误

```
"problem": Let $A B C$ be an acute-angled triangle with
   circumscribed circle $k$ and centre of the circumscribed
   circle $O$. A line through $O$ intersects the sides $A B$
   and $A C$ at $D$ and $E$. Denote by $B^{\\prime}$ and $C
   ^{\\prime}$ the reflections of $B$ and $C$ over $O$,
   respectively. Prove that the circumscribed circles of $O D
   C^{\\prime}$ and $O E B^{\\prime}$ concur on $k$.
"formal_statement": "a b c = triangle a b c; o = circle o a b
   c; d = on_line d a b, on_line d o e; e = on_line e a c,
   on_line e o d; b' = mirror b' b o; c' = mirror c' c o; x =
   on_circum x o d c', on_circum x o e b' ? cyclic a b c x",
"trans_state": "Syntactically wrong at 'e'"
```

参考文献

[1] Yang K, Poesia G, He J, et al. Formal mathematical reasoning: A new frontier in ai[A]. 2024.

[2] Wu Y, Jiang A Q, Li W, et al. Autoformalization with large language models[J]. Advances in Neural Information Processing Systems, 2022, 35: 32353-32368.

[3] Murphy L, Yang K, Sun J, et al. Autoformalizing euclidean geometry[A]. 2024.

[4] Avigad J, Dean E, Mumma J. A formal system for euclid’s elements[J]. The Review of Symbolic Logic, 2009, 2(4): 700-768.

[5] Chou S C, Gao X S, Zhang J Z. A deductive database approach to automated geometry theorem proving and discovering[J]. Journal of Automated Reasoning, 2000, 25(3): 219-246.

[6] Kutzler B, Stifter S. On the application of buchberger’s algorithm to automated geometry theorem proving[J]. Journal of Symbolic Computation, 1986, 2(4): 389-397.

[7] Wenjun W. Basic principles of mechanical theorem proving in elementary geometries[J]. Selected Works Of Wen-Tsun Wu, 2008: 195.

参考文献

[8] Trinh T H, Wu Y, Le Q V, et al. Solving olympiad geometry without human demonstrations [J]. Nature, 2024, 625(7995): 476-482.

[9] Zhang C, Song J, Li S, et al. Proposing and solving olympiad geometry with guided tree search [A]. 2024.

[10] Sicca V, Xia T, Fédérico M, et al. Newclid: A user-friendly replacement for alphageometry [A]. 2024.

[11] Fu D, Chen Z, Xia R, et al. Trustgeogen: Scalable and formal-verified data engine for trustworthy multi-modal geometric problem solving[A]. 2025.

[12] Chervonyi Y, Trinh T H, Olšák M, et al. Gold-medalist performance in solving olympiad geometry with alphageometry2[A]. 2025.

[13] Krueger R, Han J M, Selsam D. Automatically building diagrams for olympiad geometry problems.[C]//CADE. 2021: 577-588.

[14] Li J, Beeching E, Tunstall L, et al. Numinamath: The largest public dataset in ai4maths with 860k pairs of competition math problems and solutions[J]. Hugging Face repository, 2024, 13: 9.

清華大學
Tsinghua University

谢谢大家

宋晨东

2025年6月5日