An Empty Hexagon in Every Set of 30 Points

Marijn J.H. Heule

joint work with Manfred Scheucher

30th International Conference on Tools and Algorithms for the Construction and Analysis of Systems

Luxembourg April 8, 2024
Points in General Position

A finite point set $S$ in the plane is in general position if no three points in $S$ are on a line.

Throughout this talk, every set is in general position.
**k-Holes**

A **k-hole** (in $S$) is a convex $k$-gon containing no other points of $S$.

- **5-hole**
- **not a 6-hole**

$h(k)$: the **smallest** number of points that contain a $k$-hole.

For $k$ fixed, does every **sufficiently large** point set in general position contain $k$-holes?
**k-Holes Overview**

For k fixed, does every *sufficiently large* point set in general position contain k-holes?

- 3 points ⇒ ∃ 3-hole (trivial)
- 5 points ⇒ ∃ 4-hole [Klein ’32]
- 10 points ⇒ ∃ 5-hole [Harborth ’78]
- Arbitrarily large point sets with no 7-hole [Horton ’83]

Main open question: what about 6-hole?

- Sufficiently large point sets contain a 6-hole [Gerken ’08 and Nicolás ’07, independently]
- **Conjecture**: h(6) = 30 (proved in TACAS’24 paper)
Lowerbound for 4-Hole: $h(4) > 4$

Clearly, any 3-point set in general position has a 3-hole.

Some sets with four points have no 4-hole, so $h(4) > 4$: 

![Diagram of a triangle with a point inside it to illustrate a 4-hole](attachment:triangle_with_point.png)
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]
Upperbound for 4-Hole: $h(4) = 5$ [Klein, 1930s]

Happy ending problem
Lowerbound for 5-Hole: $h(5) \geq 10$

All 5-gons in these 9 points have an inner point: $h(5) = 10$
Lowerbound for 6-Hole: \( h(6) \geq 30 \)

29 points, no 6-hole \cite{Overmars'02}

- Found using simulated annealing... is now easy using SAT
- This contains 7-gons. Each 9-gon contains a 6-hole

Empty Hexagon
No Lowerbound for 7-Hole: Horton’s Construction

$2^5$ points, no 7-hole
Orientation Variables

No explicit coordinates of points

Instead for every triple $a < b < c$, one orientation variable $O_{a,b,c}$ to denote whether point $c$ is above the line $ab$

Triple orientations are enough to express $k$-gons and $k$-holes

WLOG points are sorted from left to right

Not all assignments are realizable
  - Realizability is hard [Mnëv '88]
  - Additional clauses eliminate many unrealizable assignments
Inside Variables

We introduce inside variables $I_{x;a,b,c}$ which are true if and only if point $x$ is in the triangle $abc$ with $a < x < b$ or $b < x < c$.

Four possible cases:

![Diagram of four cases]

Empty Hexagon
Inside Variables

We introduce inside variables $I_{x;a,b,c}$ which are true if and only if point $x$ is in the triangle $abc$ with $a < x < b$ or $b < x < c$.

Four possible cases:

The left two cases with $a < x < b$:

$I_{x;abc} \leftrightarrow \left((O_{abc} \rightarrow (\overline{O_{axb}} \land O_{axc})) \land (\overline{O_{abc}} \rightarrow (O_{axb} \land \overline{O_{axc}}))\right)$

The right two cases with $b < x < c$:

$I_{x;abc} \leftrightarrow \left((O_{abc} \rightarrow (O_{axc} \land \overline{O_{bxc}})) \land (\overline{O_{abc}} \rightarrow (\overline{O_{axc}} \land O_{bxc}))\right)$
We introduce **hole variables** $H_{abc}$ which are true if and only if no points occur with the triangle $abc$ with $a < b < c$.

$$\bigwedge_{a<x<c} I_{x;abc} \rightarrow H_{abc}$$

Simple 6-hole encoding:

$$\bigvee_{a,b,c \in X} H_{abc} \quad \forall \ X \subset S \text{ with } |X| = 6$$
6-Hole Encoding: One Triangle-is-Empty Check Required

Trusted 6-hole encoding uses $O(n^6)$ clauses with 20 literals:

$$\bigvee_{a,b,c \in X} \overline{H_{abc}} \quad \forall \ X \subset S \text{ with } |X| = 6$$

Example

Consider an assignment with

- $O_{abd} = 0$ and $O_{bdf} = 0$
- $O_{ace} = 1$ and $O_{cef} = 1$
6-Hole Encoding: One Triangle-is-Empty Check Required

Trusted 6-hole encoding uses $O(n^6)$ clauses with 20 literals:

$$\bigvee_{a,b,c \in X} \overline{H_{abc}} \quad \forall X \subset S \text{ with } |X| = 6$$

Example

Consider an assignment with

- $O_{abd} = 0$ and $O_{bdf} = 0$
- $O_{ace} = 1$ and $O_{cef} = 1$
- $H_{ade} = 1$

This implies the existence of a 6-hole!
6-Hole Encoding: One Triangle-is-Empty Check Required

Trusted 6-hole encoding uses $O(n^6)$ clauses with 20 literals:

$$\bigvee_{a,b,c \in X} \overline{H_{abc}} \quad \forall \ X \subset S \text{ with } |X| = 6$$

Example

Consider an assignment with

- $O_{abd} = 0$ and $O_{bdf} = 0$
- $O_{ace} = 1$ and $O_{cef} = 1$
- $H_{ade} = 1$

This implies the existence of a 6-hole!

Clause to prevent this: $O_{abd} \lor O_{bdf} \lor \overline{O_{ace}} \lor \overline{O_{cef}} \lor \overline{H_{ade}}$
6-Hole Encoding: One Triangle-is-Empty Check Required

Trusted 6-hole encoding uses $O(n^6)$ clauses with 20 literals:

$$\bigvee_{a,b,c \in X} \overline{H_{abc}} \quad \forall X \subset S \text{ with } |X| = 6$$

**Example**

Consider an assignment with

- $O_{abd} = 0$ and $O_{bdf} = 0$
- $O_{ace} = 1$ and $O_{cef} = 1$
- $H_{ade} = 1$

This implies the *existence of a 6-hole!*

Clause to prevent this: $O_{abd} \lor O_{bdf} \lor \overline{O_{ace}} \lor \overline{O_{cef}} \lor \overline{H_{ade}}$

This encoding is 5 times larger, but much easier to solve
k-Hole Encoding Using $O(n^4)$ Clauses

Shorter clauses, thus more propagation, but still $O(n^6)$

Example

Introduce $O(n^3)$ auxiliary variables:

- $A_{acd}$: a 4-gon above the line $ad$
  \[ O_{abc} \land O_{bcd} \rightarrow A_{acd} \]
- $B_{ac'd}$: a 4-gon below the line $ad$
  \[ O_{ab'c'} \land O_{b'c'd} \rightarrow B_{ac'd} \]
- Combine them to block 6-holes
  \[ \overline{A_{acd}} \lor \overline{B_{ac'd}} \lor \overline{H_{acc'}} \]

This reduces the size of the encoding to $O(n^4)$ clauses
Symmetry Breaking: Sorted & Rotated Around Point 1

1. Place leftmost point at origin.
2. Stretch points to the right to be within $y = x$ and $y = -x$.
3. Rotate by 45 degrees.
4. Projective transformation:
   $$(x, y) \mapsto \left( \frac{y}{x + \epsilon}, \frac{1}{x + \epsilon} \right)$$
Realizability Constraints

Under the assumption that points are sorted from left to right

<table>
<thead>
<tr>
<th></th>
<th>$O_{abc}$</th>
<th>$O_{abd}$</th>
<th>$O_{acd}$</th>
<th>$O_{bcd}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Block multiple sign changes with $\Theta(n^4)$ (ternary) clauses [Felsner & Weil ’01]
Impact of the Encoding

Four different encodings of a random subproblem

- $T$: the trusted encoding
- $O_1$: the explicit encoding with a single empty triangle
- $O_2$: reduce the size of $O_1$ with auxiliary variables to $O(n^4)$
- $O_3$: $O_2$ without redundant clauses

<table>
<thead>
<tr>
<th>$\Gamma$</th>
<th>#var</th>
<th>#clause</th>
<th>#conflict</th>
<th>#propagation</th>
<th>time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T$</td>
<td>62930</td>
<td>1171942</td>
<td>1082569</td>
<td>1338662627</td>
<td>243.07</td>
</tr>
<tr>
<td>$O_1$</td>
<td>62930</td>
<td>5823078</td>
<td>228838</td>
<td>282774472</td>
<td>136.20</td>
</tr>
<tr>
<td>$O_2$</td>
<td>75110</td>
<td>667005</td>
<td>211272</td>
<td>343388591</td>
<td>45.49</td>
</tr>
<tr>
<td>$O_3$</td>
<td>75110</td>
<td>436047</td>
<td>234755</td>
<td>340387692</td>
<td>39.46</td>
</tr>
</tbody>
</table>
Problem Partitioning

Partitioning to split the problem into easier subproblems

▶ Original problem UNSAT iff all subproblems UNSAT
▶ Split on variables $O_{a,a+1,a+2}$ starting from the middle
▶ One parameter: the length $\ell$, roughly $1.83^\ell$ cubes
▶ Tested on: 24 points contain 6-hole or 7-gon

<table>
<thead>
<tr>
<th>$\ell$</th>
<th>#cubes</th>
<th>avg time (s)</th>
<th>max time (s)</th>
<th>total (h)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>312 418</td>
<td>6.99</td>
<td>66.86</td>
<td>606.55</td>
</tr>
<tr>
<td>19</td>
<td>89 384</td>
<td>13.61</td>
<td>123.70</td>
<td>337.96</td>
</tr>
<tr>
<td>17</td>
<td>25 663</td>
<td>34.29</td>
<td>293.10</td>
<td>244.50</td>
</tr>
<tr>
<td>15</td>
<td>7393</td>
<td>112.61</td>
<td>949.50</td>
<td>231.27</td>
</tr>
<tr>
<td>13</td>
<td>2149</td>
<td>431.26</td>
<td>3 347.59</td>
<td>257.44</td>
</tr>
<tr>
<td>11</td>
<td>629</td>
<td>1 847.46</td>
<td>11 844.05</td>
<td>322.79</td>
</tr>
<tr>
<td>9</td>
<td>188</td>
<td>7 745.14</td>
<td>32 329.05</td>
<td>404.47</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>32 905.90</td>
<td>105 937.76</td>
<td>521.01</td>
</tr>
</tbody>
</table>

Empty Hexagon
Empty Hexagon Theorem Summary

Theorem: $h(6) = 30$

- Partitioned problem using 312,418 cubes ($\ell = 21$)
- Total runtime: 17,000 CPU hours on AWS
- Linear speedups using 1,000 machines
- Proof: 180 terabytes in unprocessed LRAT format
- Validated with formally-verified checker
The optimization steps are validated or part of the proof

**Concurrent** solving and proof checking for the first time

- The solver pipes the proof to a verified checker
- This avoids storing/writing/reading huge files
- Verified checker can easily catch up with the solver

CMU students have formalized and verified all parts in Lean

- Paper submitted to ITP ’24
Conclusions

Theorem
\[ h(6) = 30 \]

SAT appears to be the most effective technology to solve a range of problems in computational geometry

Many interesting open problems:

- Minimum number of 4-gons / 5-gons / 6-gons
- Determine whether \( g(7) = 33 \)
- Unbalanced configurations (points can be collinear)