Synthesis of Geometry Proof Problems

Chris Alvin, Louisiana State University

In Collaboration with

Supratik Mukhopadhyay, LSU
Sumit Gulwani, Microsoft Research, Redmond
Rupak Majumdar, Max Planck Institute for Software Systems

AAAI '14
Thursday, July 31, 2014

Motivation for Automatic Problem Generation

Difficulties students face with their mathematics education:

- Limited textbook problems,
- Overcoming absences and reteaching,
- Variance of time to mastery (slow or fast), and
- Acquiring problems using personally-designed criteria.

Difficulties teachers face educating students:

- Efficiently develop supplementary materials,
- Write multiple versions of exams, and
- Differentiate instruction effectively.

Why Geometry Domain?

- Problem synthesis techniques have been restricted to mostly algebraic domains. [Wolfram-Alpha, Gulwani et al. AAAI '12, etc.]

- Reasoning about diagrams is non-trivial.

- Automatic theorem proving is well-studied, but basic geometry solution and problem synthesis have yet to be explored.

Question Answered

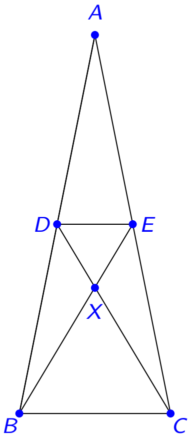
Can we synthesize (other) geometry proof problems (and their solutions) from a figure together with a set of properties true of that figure?

Our Contributions

- We formalize the notion of a geometry proof problem.
- We present a technique for generating proof problems over a geometric figure in a system we call *GeoTutor*.
- Our semi-automated approach takes a figure, analyzes, and generates problems within a few seconds.
- Supports queryable problem properties.

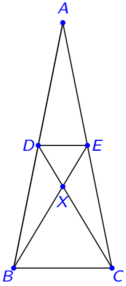
Textbook Problem

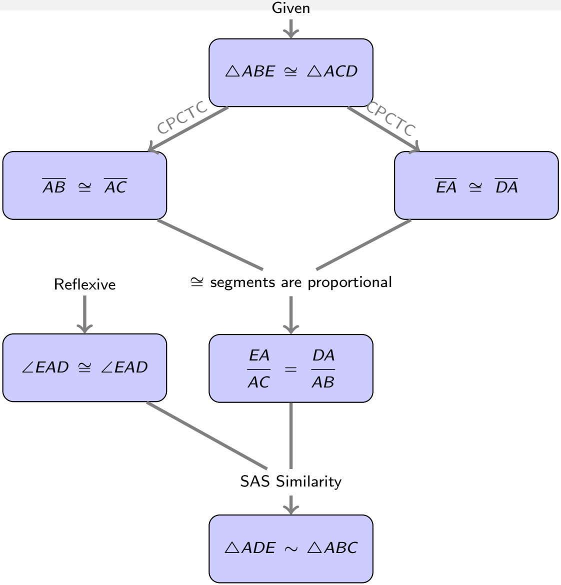
If $\Delta ABE \cong \Delta ACD$*, show that* $\Delta ADE \sim \Delta ABC$*.*



Demonstration

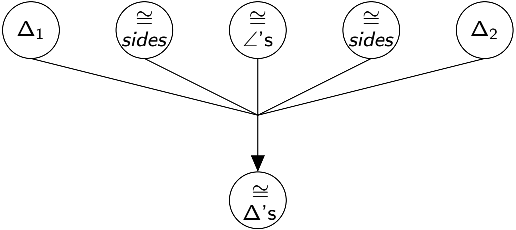
Solution


A
D
E
X
B
C


Given
$\triangle ABE \cong \triangle ACD$
CPCTC
CPCTC
$\overline{AB} \cong \overline{AC}$
$\overline{EA} \cong \overline{DA}$
Reflexive
$\cong$ segments are proportional
$\angle EAD \cong \angle EAD$
$\frac{EA}{AC} = \frac{DA}{AB}$
SAS Similarity
$\triangle ADE \sim \triangle ABC$

Internal Representation: Hypergraph

Geometric deductions can be written as logical propositions such as $P_1, P_2, \ldots, P_k \vdash P_n$ as evidenced below with the SAS congruence axiom.



We will use a **directed hypergraph** with edges being *many-to-one*.

Definition: Geometry Proof Problem

Definition

Let Fig be a figure and let Axioms be a set of geometry axioms. A **geometry proof problem** over (Fig, Axioms) is a pair (*assumptions*, *goals*), where the *assumptions* and *goals* are sets of explicit facts about Fig such that

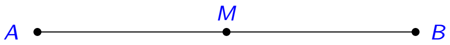
- an assumption is not a goal,
- the implicit facts of Fig, *assumptions*, and Axioms imply each goal in the set of *goals* using first-order reasoning.

Observe that a problem (and solution) is then a path in the hypergraph.

Problem Synthesis Example

Consider the following statement; we call it the *Midpoint Theorem*.

If segment AB has midpoint M, then $2AM = AB$ and $2MB = AB$.



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

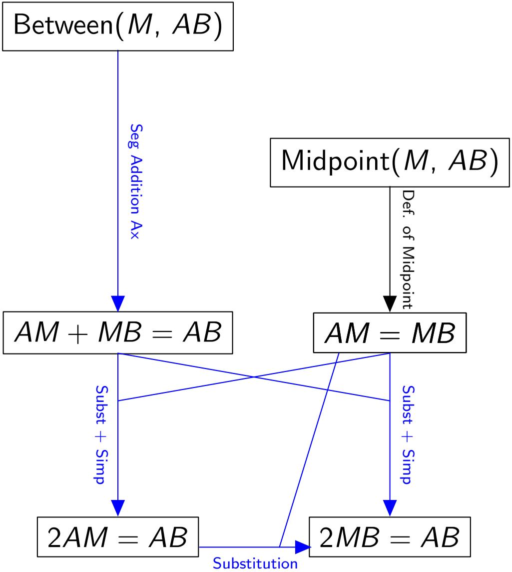
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

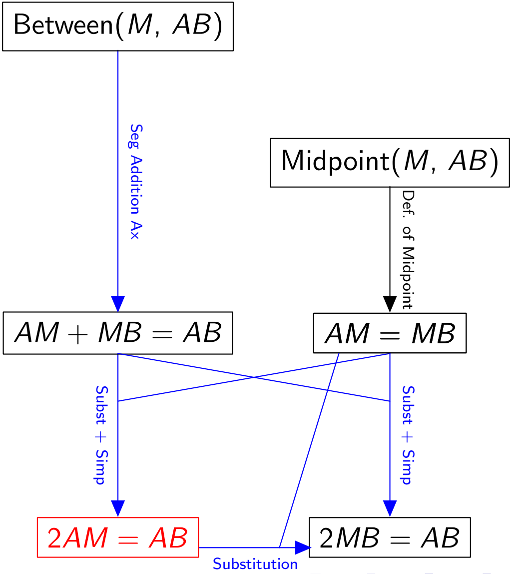
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

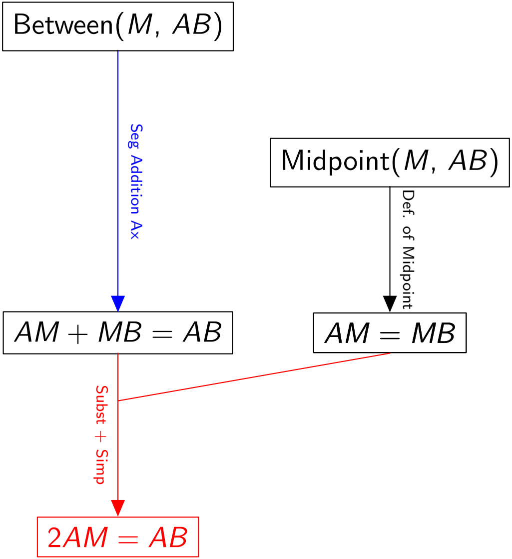
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

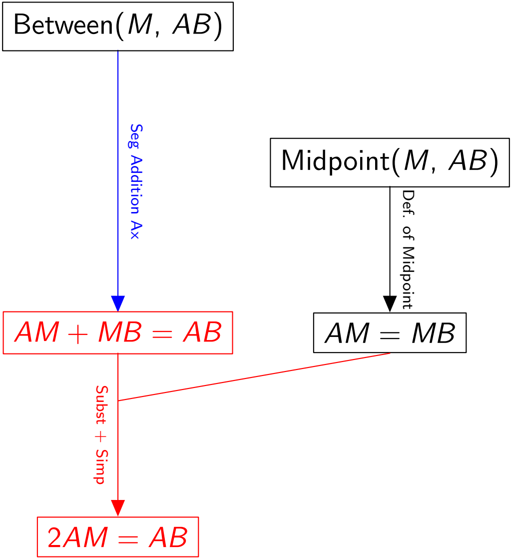
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

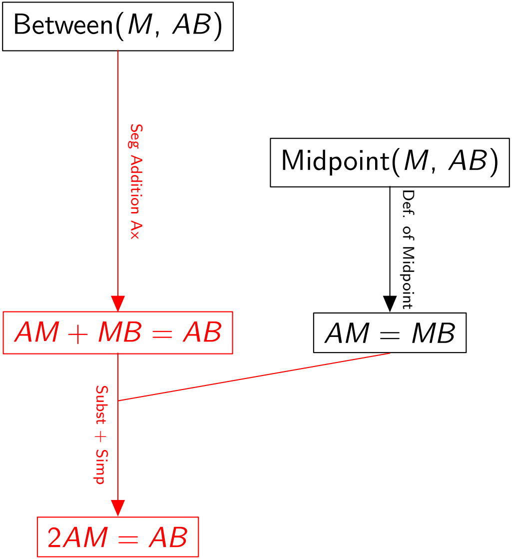
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

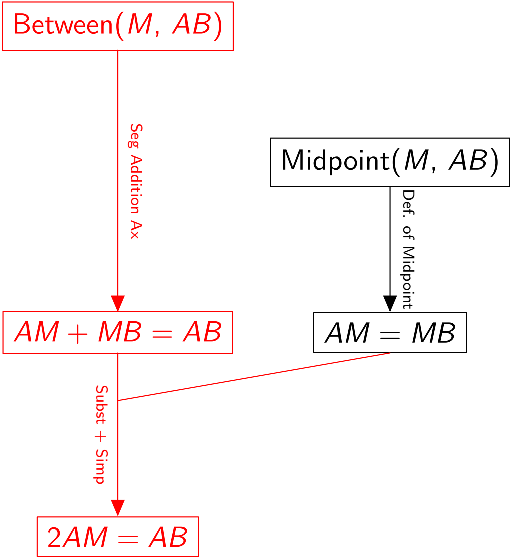
Midpoint(M, AB):

$AM + MB = AB$:

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

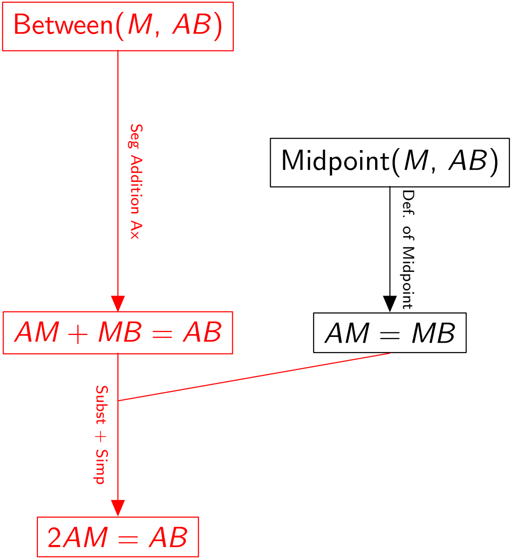
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

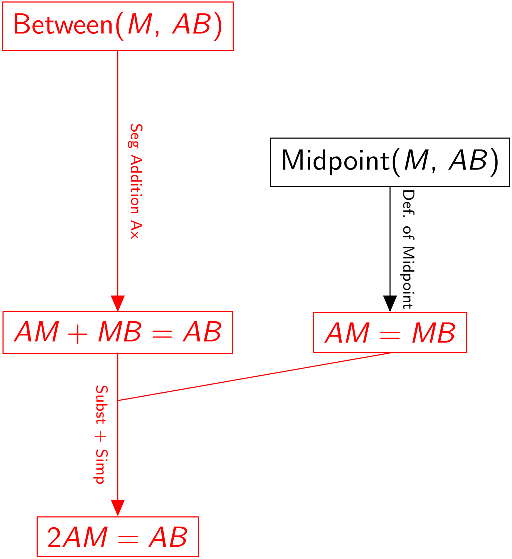
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

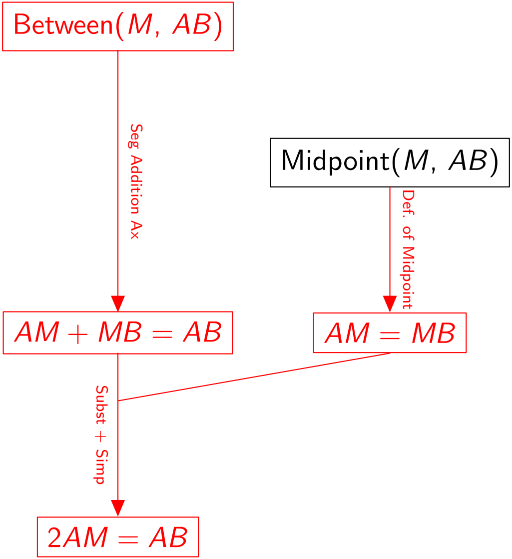
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

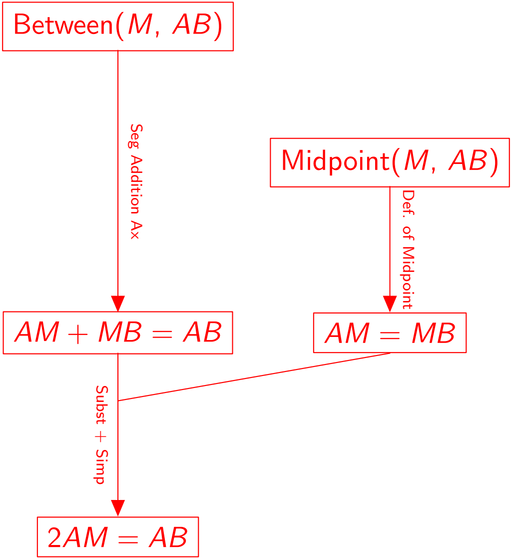
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

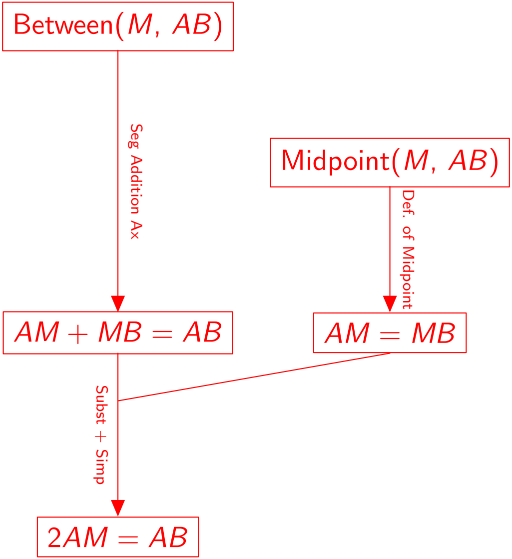
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:
(S): Midpoint(M, AB)
(G) $AM = MB$

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

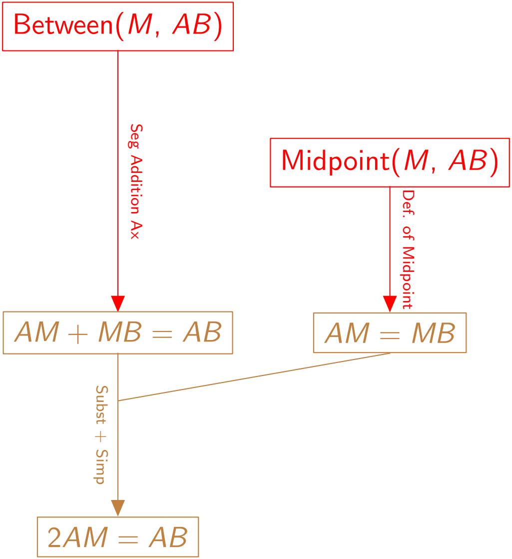
Midpoint(M, AB):

$AM + MB = AB$:
 (S): Between(M, AB)
 (G) $AM + MB = AB$

$AM = MB$:
 (S): Midpoint(M, AB)
 (G) $AM = MB$

$2AM = AB$:

$2MB = AB$:



Problem Synthesis: Midpoint Theorem

Generated Problems (Hashed by Goal)

Between(M, AB):

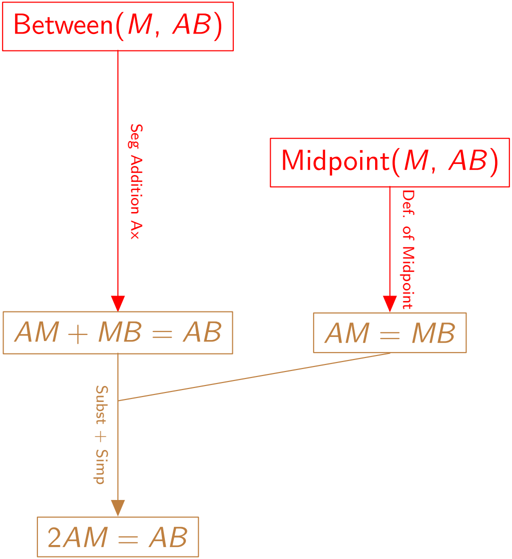
Midpoint(M, AB):

$AM + MB = AB$:
(S): Between(M, AB)
(G) $AM + MB = AB$

$AM = MB$:
(S): Midpoint(M, AB)
(G) $AM = MB$

$2AM = AB$:
(S): Between(M, AB), Midpoint(M, AB)
(P): $AM + MB = AB$
$AM = MB$
(G) $2AM = AB$

$2MB = AB$:



Interesting Geometry Proof Problem

Definition

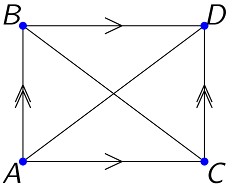
A geometry problem $(\textit{assumptions}, \textit{goals})$ over $(\text{Fig}, \text{Axioms})$ is **interesting** if the set of *assumptions* is minimal.

Interesting vs. Uninteresting

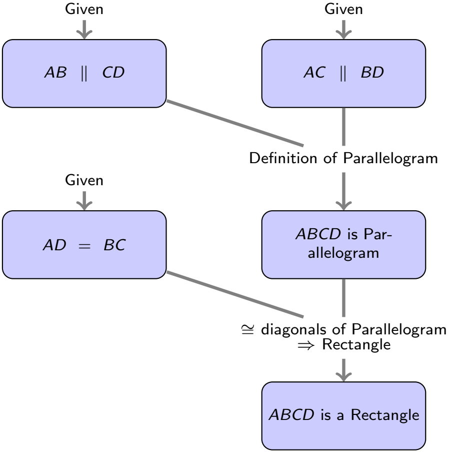
Assume:

- $AD = BC$,
- $AB \parallel CD$, and
- $AC \parallel BD$.

Goal: Prove *ABCD* is a rectangle.



Solution



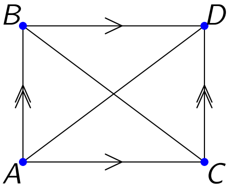
The assumptions are minimal to prove the goal; this problem is *interesting*.

Interesting vs. Uninteresting

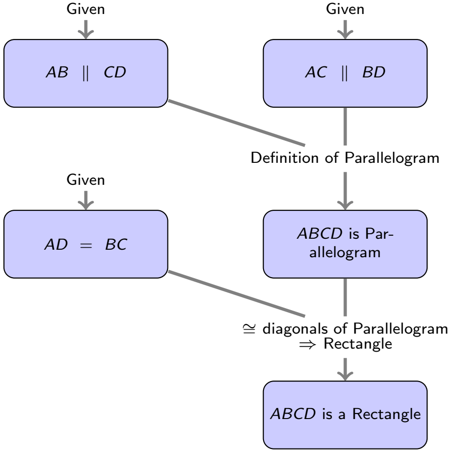
Assume:

- $AD = BC$,
- $AB \parallel CD$,
- $AC \parallel BD$, and
- $m\angle ACD = 90^o$

Goal: Prove $ABCD$ is a rectangle.



Solution



Adding assumption $m\angle ACD = 90^o$ results in an *uninteresting* problem.

Strict Geometry Proof Problem

Definition

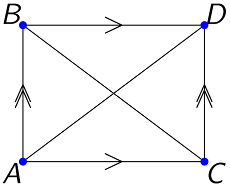
An interesting problem is **strict** if the set of *goals* is minimal.

Strict vs. Non-Strict

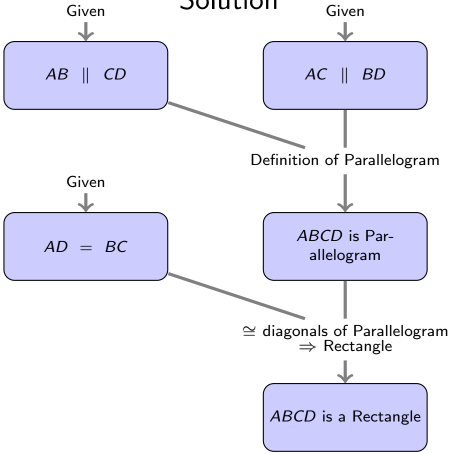
Assume:

- $AD = BC$,
- $AB \parallel CD$, and
- $AC \parallel BD$.

Goal: Prove $ABCD$ is a rectangle.



Solution



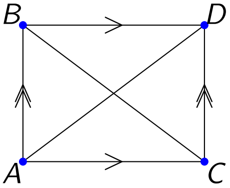
Since there is a single goal, the problem is vacuously *strict*.

Strict vs. Non-Strict

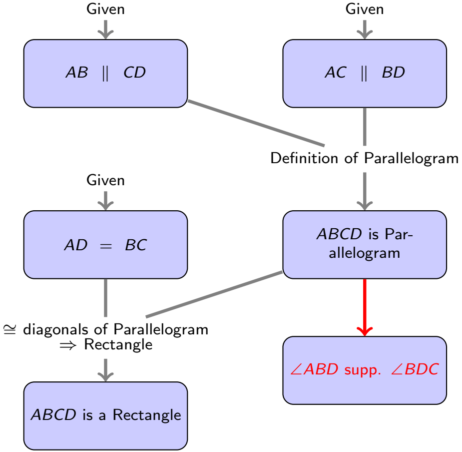
Assume:

- $AD = BC$,
- $AB \parallel CD$, and
- $AC \parallel BD$.

Goals: Prove $\angle ABD$ is supplementary to $\angle BDC$ and $ABCD$ is a rectangle.



Solution



Adding a goal not in the solution path results in a *non-strict* problem.

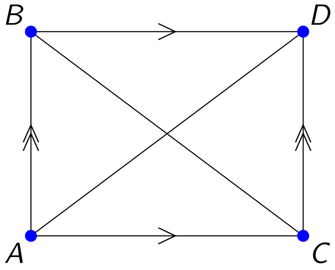
Complete Geometry Proof Problem

Definition

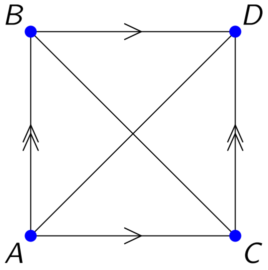
An interesting geometry problem (*assumptions*, *goals*) over (Fig, Axioms) is **complete** if the implicit facts of Fig, *assumptions*, and Axioms defines all explicit facts of the figure.

Complete vs. Interesting

Assume $AD = BC$, $AB \parallel CD$, and $AC \parallel BD$.
Goal: Prove quadrilateral $ABCD$ is a rectangle.



Complete



Interesting, Not Complete

The quadrilateral on the right is a square, alas, we cannot strengthen beyond a rectangle since no information is provided about congruent sides.

Evaluation Methodology

The corpus contained 110 figures and 155 **textbook problems** from textbooks in India and the United States.

Each textbook problem is defined as a triple: $T =< F_T, A_T, G_T >$ where:

- F_T denotes the set of intrinsic properties of the figure,
- A_T denotes the assumptions as stated in the textbook, and
- G_T the set of goals as stated in the textbook.

Our synthesis is **sound** if the respective set of generated interesting (or complete) problems contains the original problems stated in the textbook.

Results

Figures	110
Strictly Complete Textbook Problems	45
Strictly Interesting Textbook Problems	65
Ave. Generated 1-Goal Problems	37
Ave. Generated 2-Goal Problems	443
Time (secs / figure)	4.7

Table: Cumulative Results of Synthesis

Thank you for your attention. Any questions?