The Necessity Principle and Its Implementation in Mathematics Instruction

Guershon Harel

University of California, San Diego

harel@math.ucsd.edu

A "Theorem" in Mathematics Pedagogy

1. Solid knowledge of mathematics is necessary but not sufficient for quality teaching.

2. Effective teachers know mathematics and understand basic principles of mathematics learning and teaching.

N J/S	Courses	Average Grade
25	Differential Equations & Linear Algebra	3.05
	Elementary Linear Algebra	2.91

A=4, B=3, C=2, D=1

	C o r r e c t	I n c o r r e c t	N o A n s w e r
Linear Independence			
Can 3 vectors in R^2 be independent?	48	52	0
Suggest 3 independent vectors in R^3	52	36	12
Define "linear independence"	28	68	4
Span			
Define "span"	36	58	6
Subspace			
Give an example of a subspace	16	60	24
Fundamental Theorem of LA			
$r(A_{4x7}) = 3$. Find $Nullity(A)$	32	52	16
Matrix Transformation			
Are matrix transformations linear?	54	24	32
What relation exists between Columnspace(A) and Image(A)?	20	16	64

Concept	Responses
Independence	• Dependence is when reducing the arbitrary numbers are $c_1 = c_2 = \cdots = c_n = 0$ • Independence is when there aren't zero divisors.
Subspace	Vector Space: $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ Subspace: $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$
Fundamental Theorem	Column space is equal to the dimensionality of the space the image is projected in.
Matrix Transformation	A matrix transformation is linear if [it is] 1-1.

Linear Algebra Textbooks

"So far we have defined a mathematical system called a real vector space and noted some of its properties ....

[In what follows], we show that each vector space V studied here has a set composed of a finite number of vectors that completely describe V. It should be noted that, in general, there is more than one such set describing V. We now turn to a formulation of these ideas."

Following this, the text defines the concepts:

Linear independence

Span

Basis

and proves related theorems.

Is Students' Intellectual Need Considered?

- Can this "motivation" help students see a need for the pivotal concepts, *linear independence*, *span*, and *basis*?
- Can it constitute a need for the concept *finitely generated vector space* (alluded to in the "motivating paragraph)?
- Can it help students see how these three pivotal concepts contribute to the characterization of *finitely generated vector space*?

From a chapter on eigen theory

"In this section we consider the problem of factoring an $n \times n$ matrix A into a product of the form XDX^{-1}, where D is diagonal. We will give necessary and sufficient condition for the existence of such a factorization and look at a number of examples. We begin by showing that eigenvectors belonging to distinct eigenvalues are linearly independent."

The Necessity Principle

For students to learn what we intend to teach them, they must have a need for it, where ‘need’ refers to ***intellectual need***, not social or economic need.

To implement the necessity principle:

1. Recognize what constitutes an *intellectual need* for a particular population of students, relative to the concept to be learned.
2. Present the students with a sequence problems that correspond to their intellectual need, and from whose solution the concept **may** be elicited.
3. Help students elicit the concept from the problem solution.

Necessitating Diagonalization

Stage 1: The need to solve the system:

$$(1) \quad \begin{cases} AY(t) = Y'(t) \\ Y(0) = C \end{cases}$$

Stage 2: The search: the system is analogized to:

$$(2) \quad \begin{cases} ay(t) = y'(t) \\ y(0) = c \end{cases}$$

Using the known solution for system (2), $y(t) = ce^{at}$, students offer $Y(t) = e^{tA}C$ for system (1).

Stage 3: Necessitating the definition of e^{B}

Stage 4: Necessitating the concepts of eigenvalue and eigenvector

$$Y(t) = e^{tA}C = \left(\sum_{i=0}^{\infty}(1/i!)(tA)^i\right)C = \left(\sum_{i=0}^{\infty}(t^i/i!)A^i\right)C = \sum_{i=0}^{\infty}(t^i/i!)A^iC.$$

If so happens that $AC = \lambda C$ for some scalar λ, then the solution is *easily computable*.

$$Y(t) = \sum_{i=0}^{\infty}(t^i/i!)A^iC = \sum_{i=0}^{\infty}(t^i/i!)\lambda^iC = \sum_{i=0}^{\infty}((\lambda t)^i/i!)C = e^{\lambda t}C.$$

Stage 5: Necessitating the concept of diagonalization:

Is it always the case that the condition vector C is an eigenvector of the coefficient matrix A?

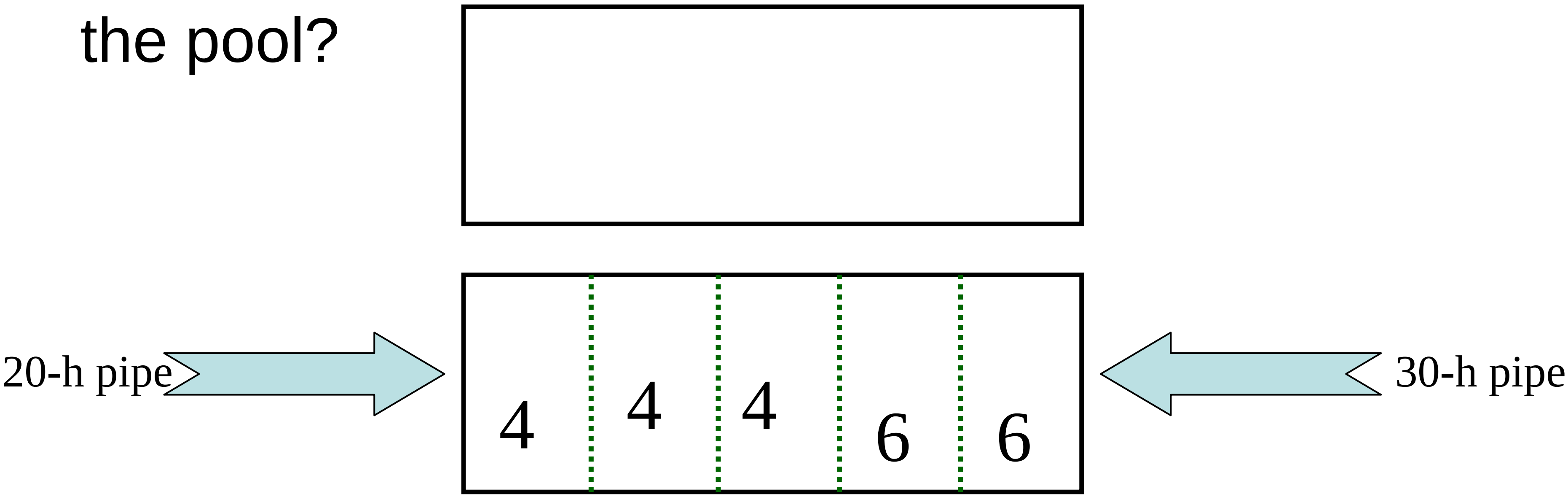
Implementing the necessity principle in elementary mathematics

How can algebraic reasoning be intellectually necessitated for students?

The case of elementary word problems.

Harriet’s solution to the *pool problem*:

A pool is connected to two pipes. One pipe can fill up the pool in 20 hours, and the other in 30 hours. How long will it take the two pipes together to fill up the pool?



Towns A and B are 300 miles apart. At 12:00 PM, a car leaves A toward B, and a truck leaves B toward A. The car drives at 80 m/h and the truck at 70 m/h. When and where will they meet?

Students' reasoning:

After 1 hour, the car drives 80 miles and truck 70 miles.

Together they drive 150 miles.

In 2 hours they will together drive 300 miles.

Therefore,

They will meet at 2:00 PM.

They will meet 160 miles from A.

Towns A and B are ~~300~~ miles apart. At 12:00 PM, a car leaves A toward B, and a truck leaves B toward A. The car drives at 80 m/h and the truck at 70 m/h. When and where will they meet?

Towns A and B are ~~300~~ 100 miles apart. At 12:00 PM, a car leaves A toward B, and a truck leaves B toward A. The car drives at 80 m/h and the truck at 70 m/h. When and where will they meet?

Students' reasoning:

It will take them less than one hour to meet.

It will take them more than 30 minutes to meet.

They will meet closer to B than to A.

Let's try some numbers:

$$\frac{50}{60}80 + \frac{50}{60}70 \stackrel{?}{=} 100$$

$$\frac{40}{60}80 + \frac{40}{60}70 \stackrel{?}{=} 100 \quad (Yes!)$$

Towns A and B are 118 miles apart. At 12:00 PM, a car leaves A toward B, and a truck leaves B toward A. The car drives at 80 m/h and the truck at 70 m/h. When and where will they meet?

Necessitating the concept of unknown

Students' reasoning:

It will take them less than one hour to meet;

It will take them more than 30 minutes to meet;

They will meet closer to B than to A.

Let's try some numbers:

$$\frac{50}{60}80 + \frac{50}{60}70 \stackrel{?}{=} 118$$

$$\frac{40}{60}80 + \frac{40}{60}70 \stackrel{?}{=} 118$$

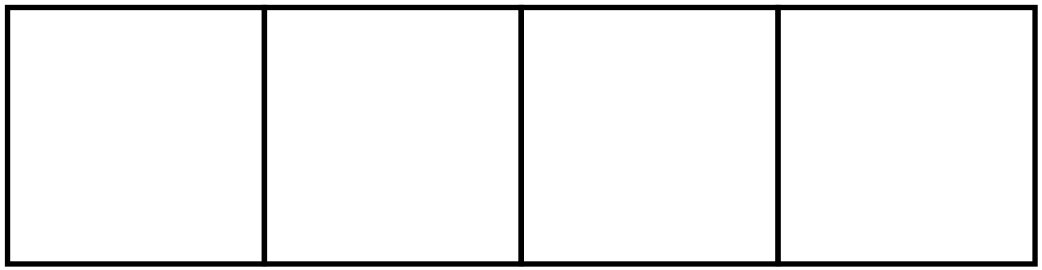
Repeated Reasoning

Tom and John are roommates. They decided to paint their room. Tom can paint the room in 4 hours and John, a perfectionist, in 8 hours. How long would it take them to paint their room if they work together?

Kate:	… $4x + 8x$ …
Teacher:	What is x?
Kate:	… x?... x is the house.
Teacher :	You want to find x, … so you want to find the house?

Tom and John are roommates; they decided to paint their room. Tom can paint the room in 4 hours and John, a perfectionist, in 8 hours. How long would it take them to paint their room if they work together?

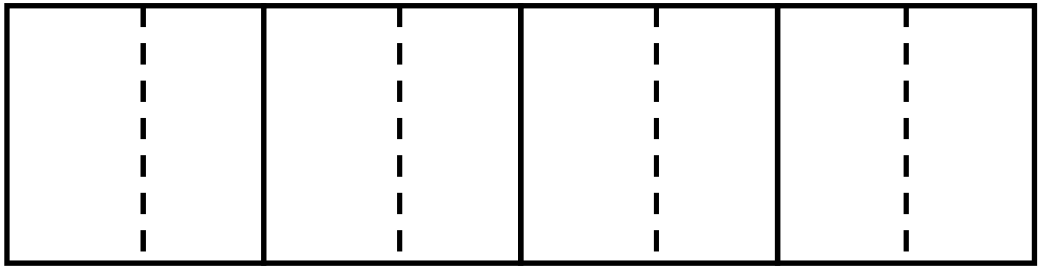
Kate:



Tom paints ¼ of the house in 1 hour.

Tom and John are roommates; they decided to paint their room. Tom can paint the room in 4 hours and John, a perfectionist, in 8 hours. How long would it take them to paint their room if they work together?

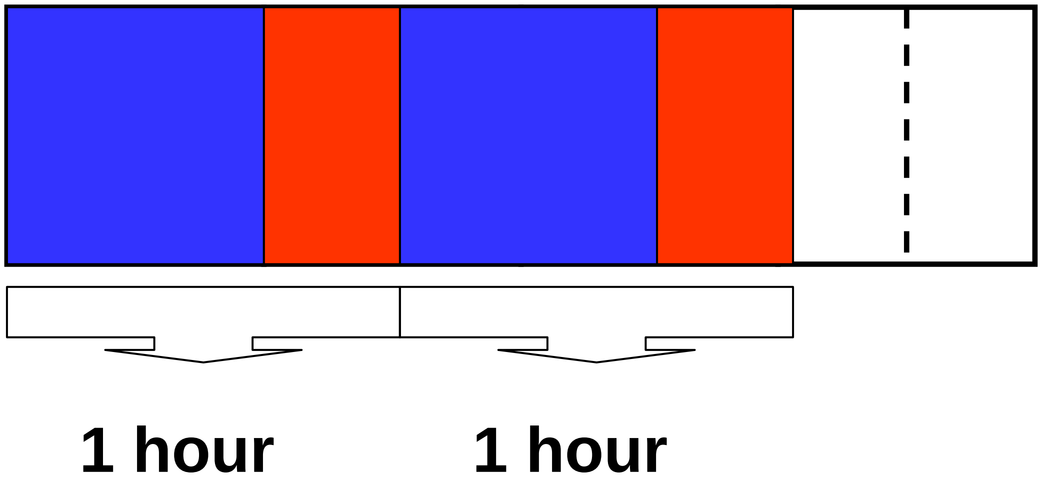
Kate:



Tom paints 1/8 of the house in 1 hour.

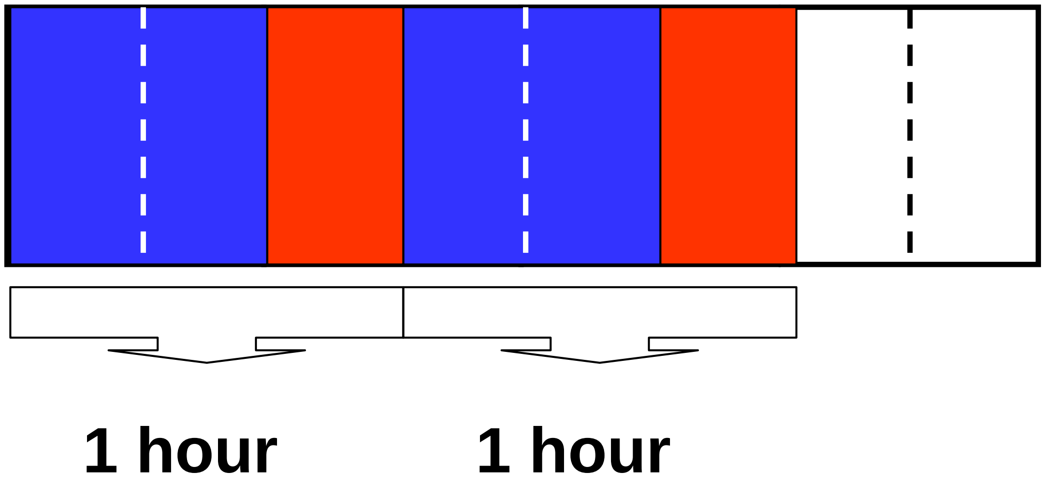
Tom and John are roommates; they decided to paint their room. Tom can paint the room in 4 hours and John, a perfectionist, in 8 hours. How long would it take them to paint their room if they work together?

Kate:



Tom and John are roommates; they decided to paint their room. Tom can paint the room in 4 hours and John, a perfectionist, in 8 hours. How long would it take them to paint their room if they work together?

Kate:



- ¾ of the house is painted in 2 hours, ¼ of the house will be painted in 2/3 of an hour.
- The whole house will be painted in 2 hours and 40 minutes.
- …This can't be right … there is no x …

intellectual need response

social need response

The house will be painted in 2 hours and 40 minutes

…This can’t be right … there is no x …

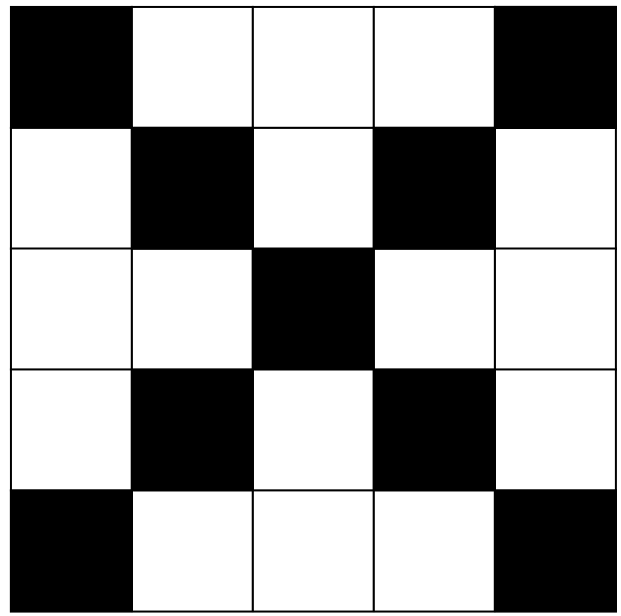
Necessitating Deductive Reasoning

Students and Teachers' Conceptions of Proof: Selected Results

- Students and teachers justify mathematical assertions by examples

- Often students' and teachers' inductive verifications consist of one or two example, rather than a multitude of examples.

- Students' and teachers' conviction in the truth of an assertion is particularly strong when they observe a **pattern**.



Doris Solution

Dimension	Number of White Squares	
1	0	$(1-1)^2$
3	4	$(3-1)^2$
5	16	$(5-1)^2$
7	36	$(7-1)^2$

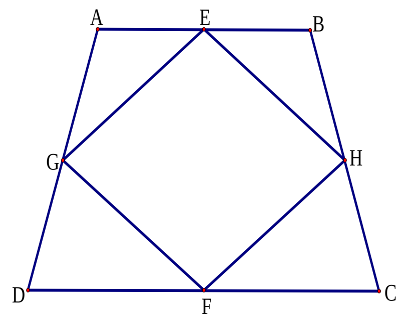
Doris Solution

Dimension	Number of White Squares	
1	0	$(1-1)^2$
3	4	$(3-1)^2$
5	16	$(5-1)^2$
7	36	$(7-1)^2$
x		$(x-1)^2$

- **Their proofs often rely on visual perceptions**

They Impose properties on a figures

The midpoints of an isosceles trapezoid form a rhombus.

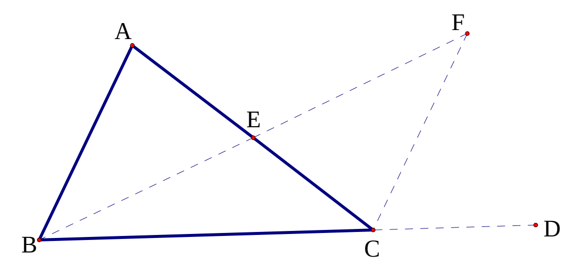


They do not question visual properties:

An Exterior angle of a triangle is greater than the opposite interior angle.

$$\triangle BEA \cong \triangle FEC \Rightarrow \measuredangle A = \measuredangle ACF$$

$$\measuredangle ACF < \measuredangle ACD$$



- How do we know this relation among the angles?
- Why CF is inside $\measuredangle ACD$?

- **Their proofs often rely on perceived number relations**

Problem: Determine whether the following vectors are linearly independent.

$$u = \begin{bmatrix} \sqrt{3} \\ \sqrt{5} \end{bmatrix} \quad v = \begin{bmatrix} -\frac{7}{10} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad z = \begin{bmatrix} \sqrt[3]{61} \\ 1.03 \end{bmatrix}$$

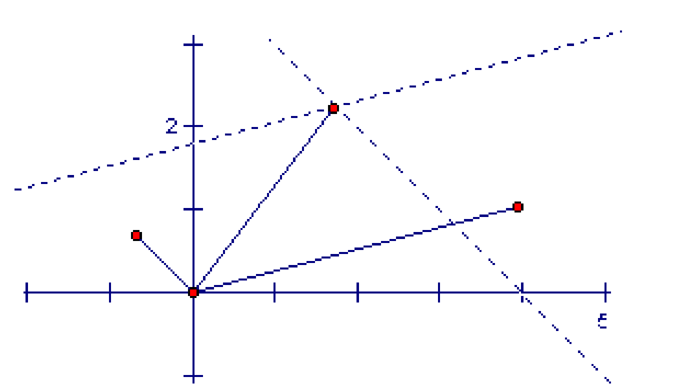
Example of Response: "By looking at the numbers, there is no way one vector can be a linear combination of the others."

- **Their proofs often rely on perceived number relations**

Problem: Determine whether the following vectors are linearly independent.

$$u = \begin{bmatrix} \sqrt{3} \\ \sqrt{5} \end{bmatrix} \quad v = \begin{bmatrix} -\dfrac{7}{10} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} \quad z = \begin{bmatrix} \sqrt[3]{61} \\ 1.03 \end{bmatrix}$$

Response: “The columns cannot be dependent on each other. However, from the [attached] sketch, we see that all lie on the same plane. Thus, the vectors are dependent.”



- **Students view a counterexample as an exception—in their view it does not affect the validity of the statement.**

- **Confusion between empirical proofs and proofs by exhaustion.**

- **Confusion between the admissibility of proof by counterexample with the inadmissibility of proof by example.**

Teaching Actions with Limited Effect

- Raising skepticism as to whether the assertion is true beyond the cases evaluated.
- Showing the limitations inherent in the use of examples through situations such as:

The conjecture " $\sqrt{1141y^2+1}$ is an integer" is false for $1 \leq y \leq 10^{25}$.
The first value for which the statement is true is:
30,693,385,322,765,657,197,397,208

Why showing the limitations inherent in the use of examples is not effective?

- Students do not seem to be impressed by situations such as:
  - The conjecture “ $\sqrt{1141y^2+1}$ is an integer” is false for $1 \le y \le 10^{25}$. The first value for which the statement is true is: 30,693,385,322,765,657,197,397,208

- **Students view a counterexample as an exception—in their view it does not affect the validity of the statement**.

The Role of Causality in Necessitating Deductive Reasoning

Aristotle’s Definition of Science

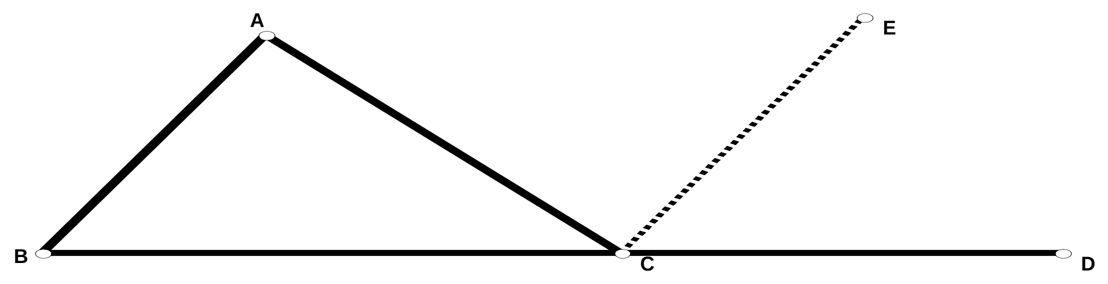
“We do not think we understand something until we have grasped the why of it. … To grasp the why of a thing is to grasp its primary cause.” Aristotle, *Posterior Analytics.*

Mathematics is not a perfect science, argued 16-17th Century philosophers, because an "implication" is not just a logical consequence; it must also demonstrate the ***cause*** of the conclusion.

Euclid’s Proof to Proposition 1.32

The sum of the three interior angles of a triangle is equal to 180^0.

Proof:



“Mathematics is not scientific:” Examples of other arguments

- Proof by contradiction is not a causal proof since it does not provide sufficient insight of how the result was obtained.
- If mathematical proof are scientific (i.e., causal), then equivalent statements (i.e., “A iff B” statements) are an absurdity.

    “A implies B” means “A causes B”

    “B implies A” means “B causes A”

    Hence: “A causes A”—an absurdity.

Theorem:

` Eigenvectors $v_1, v_2, \ldots, v_n$ of a matrix A that correspond to distinct eigenvalues $p_1, p_2, \ldots, p_n$ are linearly independent.

Proof (presented to students):

1. Let $a_1v_1+a_2v_2+\ldots+a_nv_n=0$
2. Take $f_j(x)=(x-p_1)(x-p_2)\ldots(x-p_{j-1})(x-p_{j+1})\ldots(x-p_n)$

   By the Spectral Mapping Theorem, for $j=1, 2, \ldots, n$

   $$0=a_1 f_j(A)v_1+a_2 f_j(A)v_2+\ldots+a_n f_j(A)v_n = a_1 f_j(p_1)v_1+a_2 f_j(p_2)v_2+\ldots+a_n f_j(p_n)v_n = a_j f_j(p_j)v_j$$

3. Hence, $a_j=0$ for $j=1, 2, \ldots, n$

Interviews with students (always the more able ones) indicate:

1. Students understand the proof’s steps
2. Students’ difficulty is *not* the “invention” of the Lagrange Polynomials
2. Students argue that “something is wrong with the proof” in that the desired property is dependent on the choice of the polynomials: “If you take different polynomials …, they [the eigenvectors] may not be linearly independent.”

Proposition:

A set of $n+1$ vectors in R^n is linearly dependent.

Proof (presented to students):

1. Let $v_1, v_2, \ldots, v_{n+1}$ be vectors in R^n, and let $A=[v_1\ v_2\ \ldots\ v_{n+1}]$.

2. The homogeneous system $Ax=0$ has at least one free variable and, therefore, it has a non-zero solution, $x=[x_1\ x_2\ \ldots\ x_{n+1}]^t$.

3. Since $x_1v_1+x_2v_2+\cdots+x_{n+1}v_{n+1}=0$ and $x_1, x_2, \ldots, x_{n+1}$ are not all zero, one of the vectors $v_1, v_2, \ldots, v_{n+1}$ must be a linear combination of the others.

Interviews with students (always the more able ones) indicate:

1. Students understand the proof's steps
2. Students' difficulty is *not* the "invention" of the homogeneous system
2. Students think "something is wrong with the proof": The linear independence property of the $v_1, v_2, \ldots, v_{n+1}$ seems to them a result of a particular choice—that the system we started with was homogenous: "but what if the system you started with is not homogeneous?"

Attempts to Conform to the Aristotelian Theory of Science

1. Reject Proof by Contradiction—Use Ostensive (Causal) Proofs

Descartes appealed to a priori proofs against proofs by contradiction

Cavalieri’s geometry of indivisibles

2. Adopt G*enetic definitions* and *motion-based proofs in geometry*

Genetic definition:

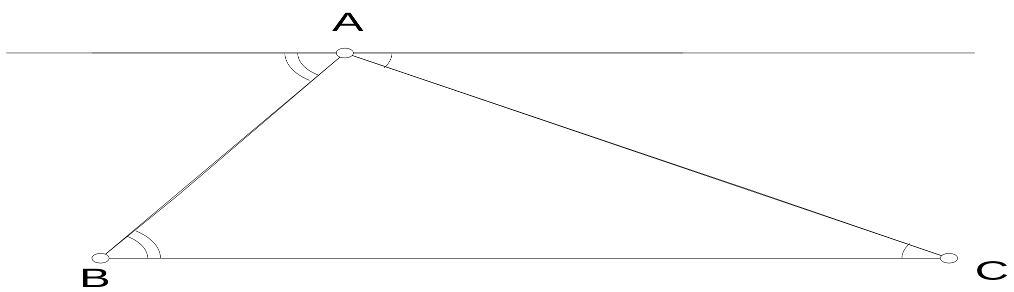
– A *sphere* as an object generated by the rotation of a semicircle around a segment taken as axis.

Non-genetic definition:

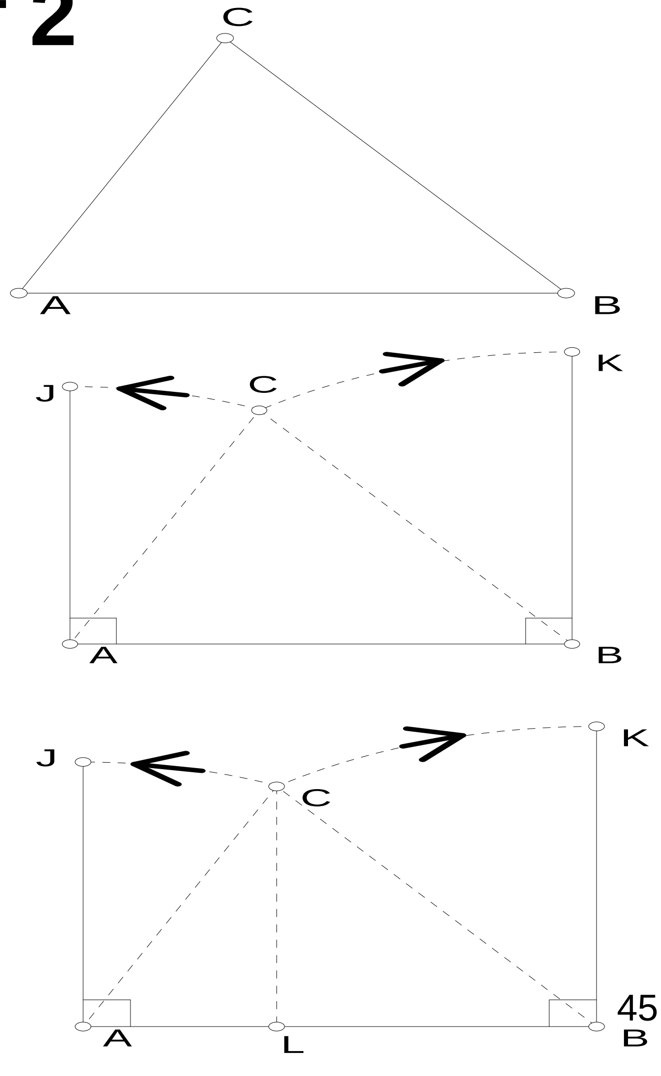
– A *sphere* as the set of all points in space that are equidistant from a given point.

Theorem: For every triangle, the sum of the measures of its interior angles is 180^0.

Proof 1



Proof 2



Is there an upper bound for the sequence

$$\sqrt{2},\quad \sqrt{2+\sqrt{2}},\quad \sqrt{2+\sqrt{2+\sqrt{2}}},\ldots$$

$$\sqrt{2} = 1.41421 < 2$$

$$\sqrt{2+\sqrt{2}} = 1.\ 8478 < 2$$

$$\sqrt{2+\sqrt{2+\sqrt{2}}} = 1.\ 9616 < 2$$

Therefore

Every term in the sequence is less than 2.

$\sqrt{2}$ is less than 2

Therefore

$2+\sqrt{2}$ is less than 4

Therefore

$\sqrt{2+\sqrt{2}}$ is less than 2

Therefore

$2+\sqrt{2+\sqrt{2}}$ is less than 4

Therefore

$\sqrt{2+\sqrt{2+\sqrt{2}}}$ is less than 2

- Find an upper bound to the sequence $\sqrt{2}, \quad \sqrt{2+\sqrt{2}}, \quad \sqrt{2+\sqrt{2+\sqrt{2}}}, \ldots$

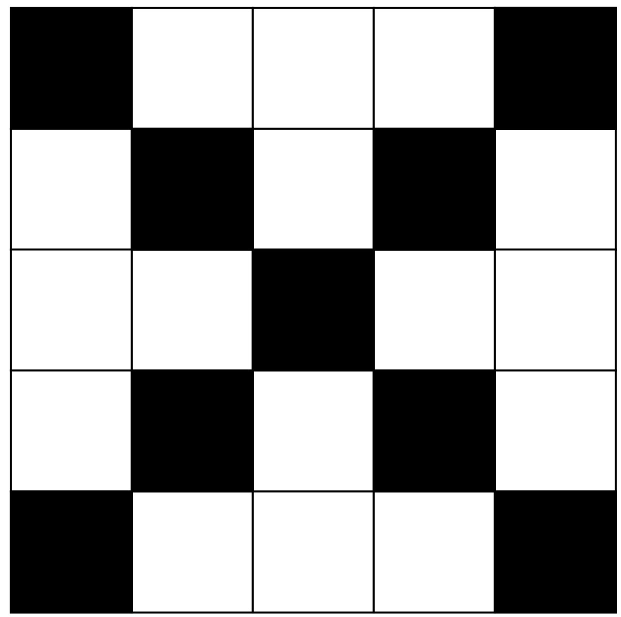
- You are given 3^n coins, all identical except for one which is heavier. Using a balance, prove that you can find the heavy coin in *n* weighings.

- Let *n* be a positive integer. Show that any 3^nx3^n chessboard with one square removed can be tiled using L-shaped pieces, where each pieces covers three squares.

- Three pegs are stuck in a board. On one of these pegs is a pile of disks graduated in size, the smallest being on top. The object of this puzzle is to transfer the pile to one of the other two pegs by moving the disks one at a time from one peg to another in such a way that a disk is never placed on top of smaller disk. How many moves are needed to transfer a pile of *n* disks?

- Prove that every third Fibonacci number is even.

- Into how many regions is the plane cut by *n* lines, assuming no two lines are parallel and no three intersect at a point.



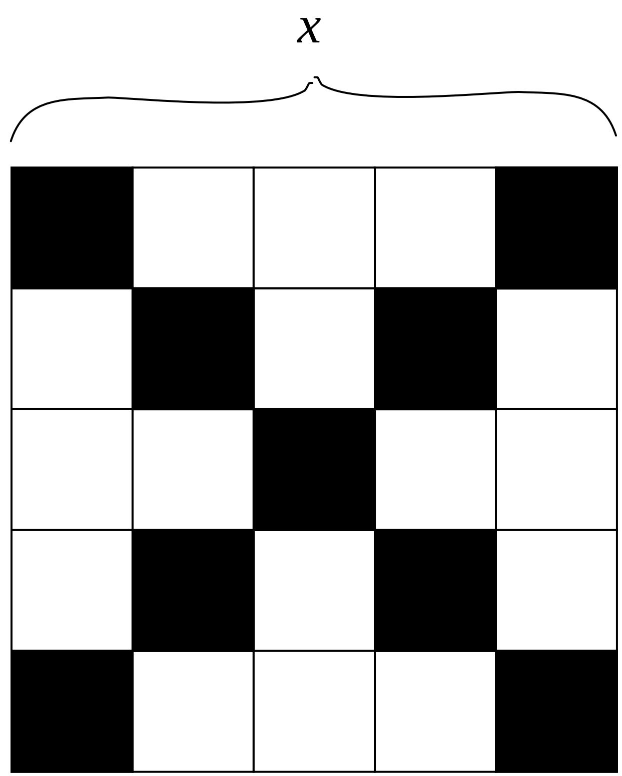
Doris Solution

Dimension	Number of White Squares	
1	0	$(1-1)^2$
3	4	$(3-1)^2$
5	16	$(5-1)^2$
7	36	$(7-1)^2$

Doris Solution

Dimension	Number of White Squares	
1	0	$(1-1)^2$
3	4	$(3-1)^2$
5	16	$(5-1)^2$
7	36	$(7-1)^2$
x		$(x-1)^2$

John’s Solution



$$\left(\frac{x-1}{2}\right)^2 \quad \left(\frac{x-1}{2}\right)^2$$

$$\left(\frac{x-1}{2}\right)^2 \quad \left(\frac{x-1}{2}\right)^2$$

$$4\left(\frac{x-1}{2}\right)^2 = (x-1)^2$$

Local Necessity versus Global Necessity

First Course in Linear Algebra

- Focus: Linear systems: scalar and differential
- The problem: Given a linear system,
  1. How do we solve it?
     - Is there an algorithm to solve such a system?
  2. Is it possible to determine whether a system has a solution without actually solving it?
  3. If a system is solvable, can we determine how many solutions does it have without solving it?
  4. If the system has infinitely many solutions, can we describe them—can we represent all the solutions in terms of a finite number of solutions?

The concept of *linear independence* was necessitated to answer the question:

Under what conditions if a system AX=b has a solution, the solution is unique?

The concept of *orthogonal projection* was necessitated to:

Find a "best" solution to an inconsistent system Ax = b.

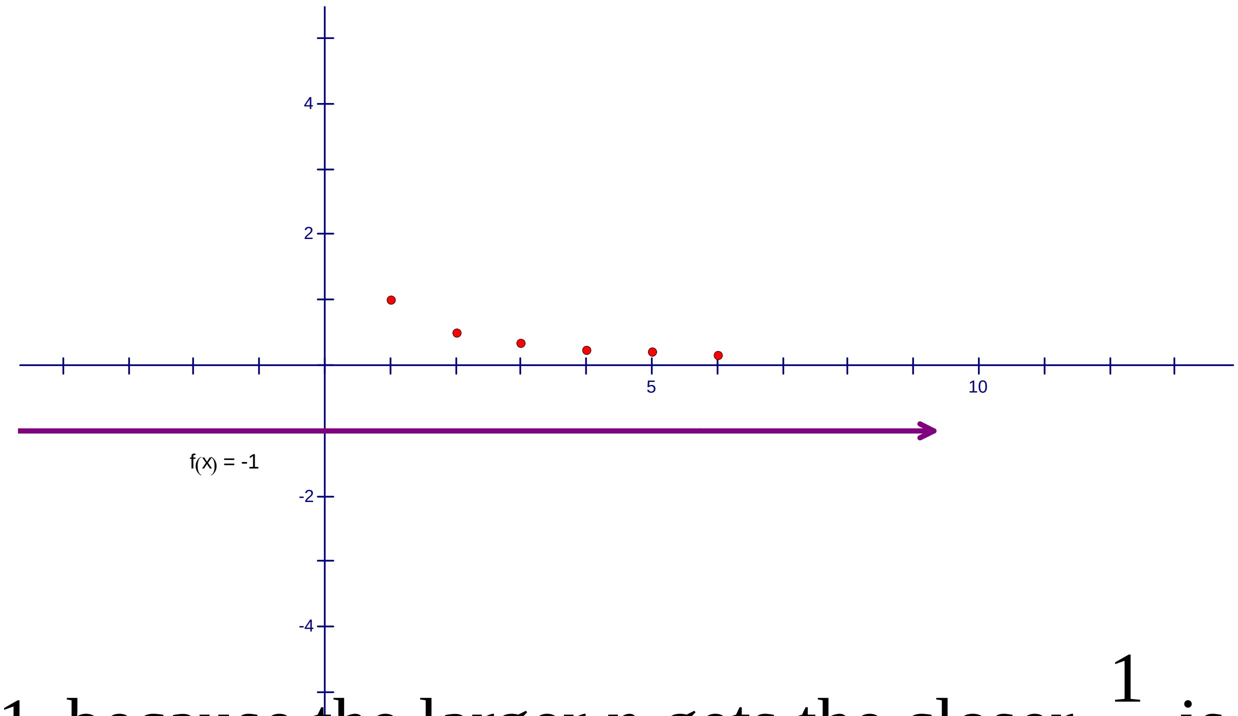
Goal: To necessitate the e-N definition of limit

Teacher: What is $\lim_{n\to\infty}\frac{1}{n}$ **and why?**

Students:

$$\lim_{n\to\infty}\frac{1}{n}=0 \text{ because the larger } n \text{ gets the closer } \frac{1}{n} \text{ is to zero.}$$

Teacher:



$$\lim_{n\to\infty}\frac{1}{n}=-1 \text{ because the larger } n \text{ gets the closer } \frac{1}{n} \text{ is to -1.}$$

Pedagogical Tools Violating the NP

The Case of Premature Introduction of New Knowledge

Multiplication and Division

(Junior-high school students)

A cheese weighs 5 pounds. 1 pound costs $12. Find out the the price of the cheese. Which operation would you have to have to perform?

(a) **12÷5** (b) **5÷12**

(c) **5×12** (d) **12+12+12+12+12**

Correct responses (c and d): **83%**

A cheese weighs 0.823 pounds. 1 pound costs $10.50. Find Find out the price of the cheese. Which operation would would you have to perform?

(a) **10.50+0.823** (b) **10.50×0.823**

(c) **10.50÷0.823** (d) **10.50-0.823**

Correct responses: **29%**

Teachers teach the following solution strategy:

A cheese weighs 0.823 pounds. 1 pound costs \$10.50. Find out the price of the cheese.

1. **Replace the decimal numbers (0.823 & 10.50) by any whole numbers (say, 8 & 10):**

   A cheese weighs 8 pounds. 1 pound costs \$10. Find out the price of the cheese.

2. **Find the expression that solves the *new* problem:**

   8×10

3. **Replace the numbers in this expression with the original decimal numbers:**

   0.823×10.50

A cheese weighs 0.823 pounds. 1 pound costs $10.50. $10.50. Find out the price of the cheese. Which operation would you have to perform?

(a) **10.50 + 0.823** (b) **10.50 × 0.823**

(c) **10.50 ÷ 0.823** (d) **10.50 - 0.823**

Rita: None [none of the given choices is a solution to the problem].

T: How would you solve the problem?

Rita: (A long pause)

One thousandth of a pound costs 10.50 divided by 1000. Then I times that by 823.

Teacher: One pound of candy cost $7. How much would 3 pounds cost?

Tammy: Three times seven: 21.

Dan: I agree, 3 times 7.

Teacher: How much would I pay if I buy only 0.31 of a pound?

Tammy: It is the same. You only changed the number. 0.31 times 7.

Dan: No way! It isn’t the same. … Can’t be. It isn’t times.

Teacher: How would you, Dan, solve the problem?

Dan: Divide 1 by 0.31. Take that number, whatever that number is, and divide 7 by it.

Ed’s Background

1. Second grader
2. Reported by his teacher to be slightly above average in mathematics and average in other subjects
3. Ed’s formal instruction had included addition and subtraction without regrouping
4. Ed was informally exposed to the basic notion of multiplication as repeated addition and the meaning of division as sharing equally
5. Ed was never taught any division strategy

Ed’s Strategy for Solving Division Problems

Interviewer: How much is 42 divided by 7?

Ed: … That’s easy …

40 divided by 10 is 4

3 plus 3 plus 3 plus 3 is 12

12 plus 2 is 14

14 divided by 2 is 7

2 plus 4 is 6

The answer is 6! (**triumphantly**)

Interviewer (to himself):

Okay, miracles do happen …

Interviewer : How about 56 divided by 8?

Ed: You do the same thing (**Impatiently**)

50 divided by 10 is 5

5 times 2 is 10

10 plus 6 is 16

16 divided by 2 is 8

5 plus 2 is 7

The answer is 7.

Interviewer : … And 72 divided by 9?

Ed: Are you going to ask me every single problem?

Ed: **… okay …** 72 divided by 9 …

70 divided by 10 is 7

7+2 is 9

1 and 7 is 8

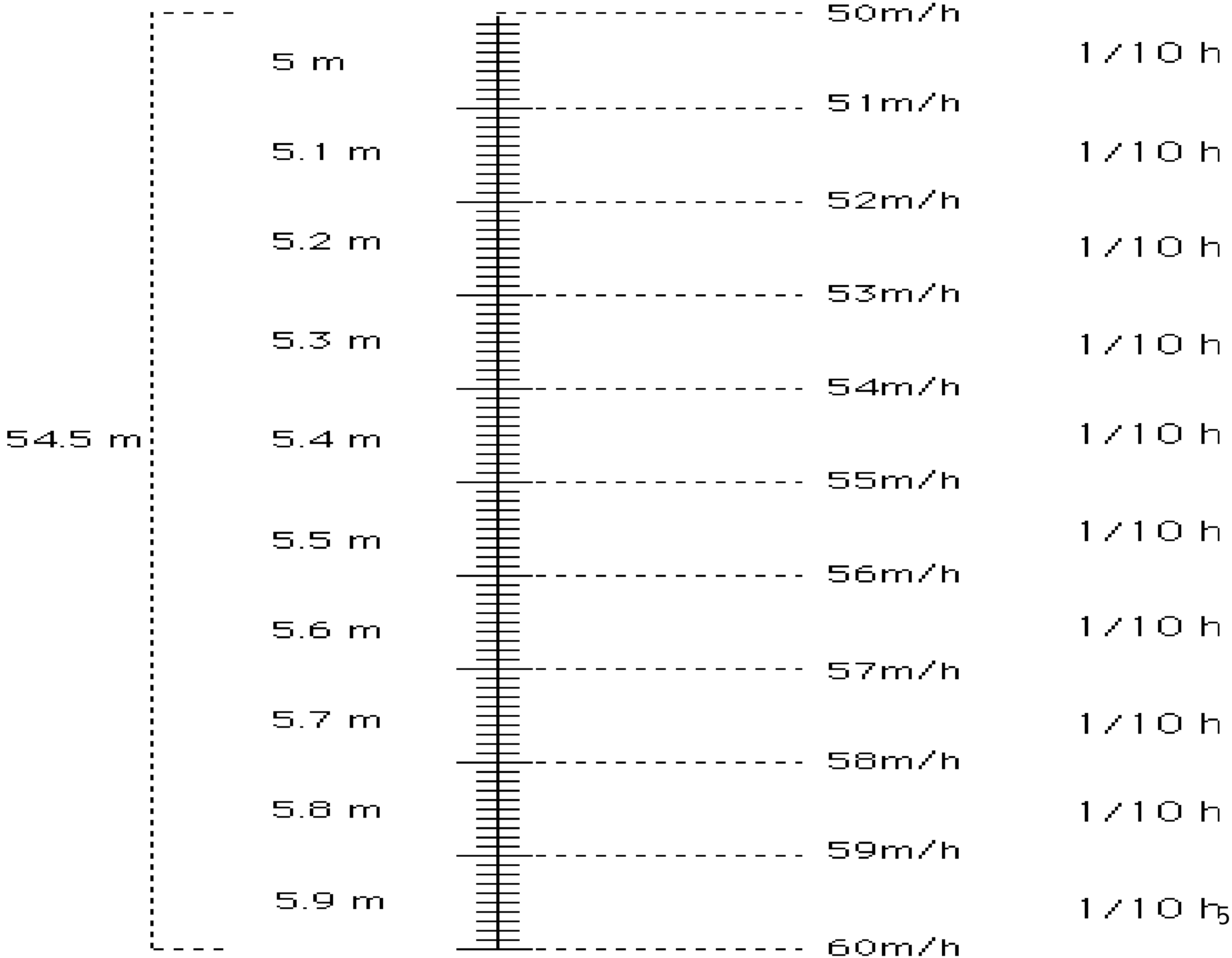
The answer is 8.

I: Imagine this. I'm driving my car at 50 mi/hr. I speed up smoothly to 60 mi/hr, and it takes me one hour to do it. About how far did I go in that hour?

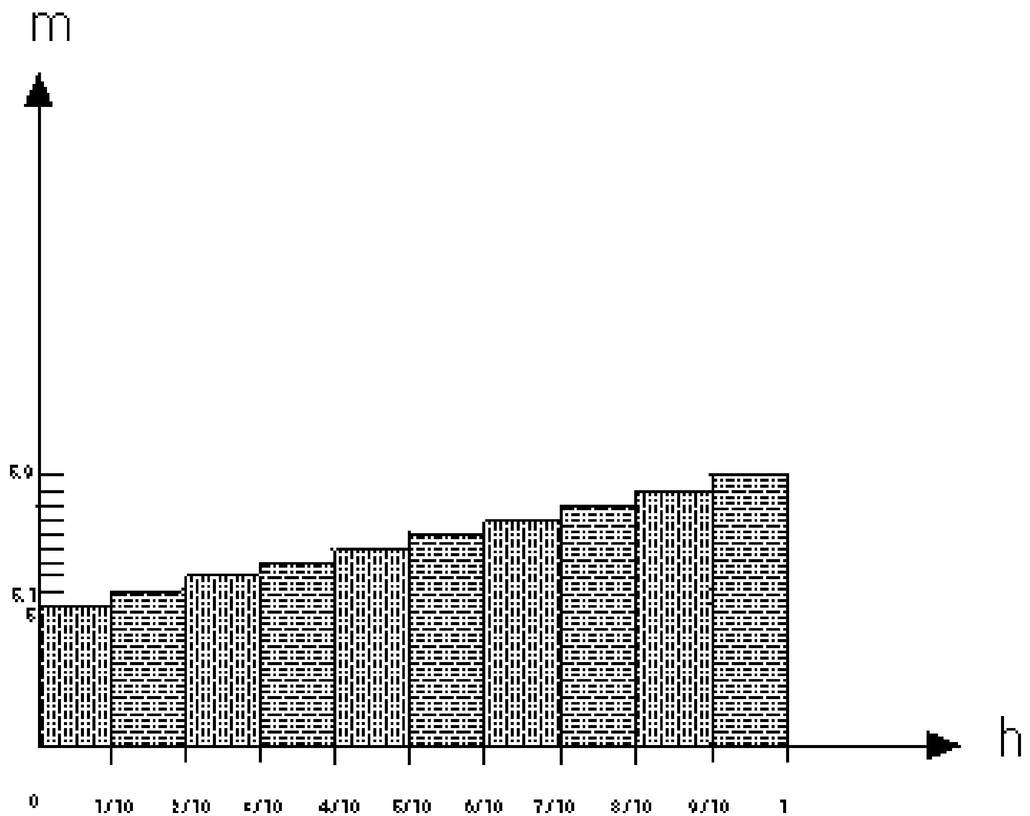
S: *(Long pause. Begins drawing a number line.)*

I: What are you doing?

S: I figure that if you speed up 10 miles per hour in one hour, that you speeded up 1 mile per hour every 6 minutes. So I'll figure how far you went in each of those six minutes and then add them up.



54.5 m
5 m
5.1 m
5.2 m
5.3 m
5.4 m
5.5 m
5.6 m
5.7 m
5.8 m
5.9 m
50m/h
51m/h
52m/h
53m/h
54m/h
55m/h
56m/h
57m/h
58m/h
59m/h
60m/h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h
1/10 h


m
5.9
5.1
5
h
0
1/10
2/10
4/10
5/10
6/10
7/10
8/10
9/10
1

- Prospective teachers who saw Sue's solution commented that, had they been Sue's teacher, they would have had Sue "discover" that she could just multiply the amount of time taken to speed up (1 hour) by the mean of the beginning and ending speeds ($\frac{50+60}{2} = 55$)
- The consensus among these prospective teachers was that Sue's solution does not have to do anything with calculus.
- They viewed Sue's solution as a rather clumsy way to approximate "the correct answer".

Instructor:	What is fundamental about the Fundamental Theorem Calculus?
	Silence
Instructor	What is the Fundamental Theorem of Calculus?
	Silence
Student:	It’s something about integrals and derivatives.

Instructor: What is fundamental about the Fundamental Theorem Calculus?

Student: The Fundamental Theorem of Calculus tells us that in order to find the area of a region under a curve of a function, we do not need to divide the region into squares, count, divide into squares again count, … We only need to find the anti derivative of the function and plug in the end points.