Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

6.5 REDUCTIONS

introduction

designing algorithms
establishing lower bounds
classifying problems

intractability

Robert Sedgewick | Kevin Wayne

Algorithms

http://algs4.cs.princeton.edu

Overview: introduction to advanced topics

Main topics. [final two lectures]

- Reduction: relationship between two problems.
- Algorithm design: paradigms for solving problems.

Shifting gears.

- From individual problems to problem-solving models.
- From linear/quadratic to polynomial/exponential scale.
- From implementation details to conceptual frameworks.

Goals.

- Place algorithms and techniques we've studied in a larger context.
- Introduce you to important and essential ideas.
- Inspire you to learn more about algorithms!

6.5 REDUCTIONS

designing algorithms

classifying problems

establishing lower bounds

introduction

intractability

Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

http://algs4.cs.princeton.edu

Desiderata. Classify problems according to computational requirements.

complexity	order of growth	examples
linear	Ν	min, max, median, Burrows-Wheeler transform,
linearithmic	$N \log N$	sorting, element distinctness, closest pair, Euclidean MST,
quadratic	N^2	?
÷	:	
exponential	<i>c ^{<i>N</i>}</i>	?

Frustrating news. Huge number of problems have defied classification.

Desiderata. Classify problems according to computational requirements.

Desiderata'. Suppose we could (could not) solve problem *X* efficiently. What else could (could not) we solve efficiently?



"Give me a lever long enough and a fulcrum on which to place it, and I shall move the world." — Archimedes





Ex 1. [finding the median reduces to sorting]

To find the median of *N* items:

- Sort *N* items.
- Return item in the middle.

cost of sorting cost of reductionCost of finding the median. $N \log N + 1$.



Ex 2. [element distinctness reduces to sorting]

To solve element distinctness on N items:

- Sort *N* items.
- Check adjacent pairs for equality.

Cost of element distinctness. $N \log N + N$.



Novice error. Confusing *X* reduces to *Y* with *Y* reduces to *X*.

ON COMPUTABLE NUMBERS, WITH AN APPLICATION TO THE ENTSCHEIDUNGSPROBLEM

By A. M. TURING.

[Received 28 May, 1936.-Read 12 November, 1936.]



Which of the following reductions have we encountered in this course?

- I. MAX-FLOW reduces to MIN-CUT.
- II. MIN-CUT reduces to MAX-FLOW. 🖌

need to find max st-flow and min st-cut (not simply compute the value)

- A. I only.
- **B.** II only.
- C. Both I and II.
- **D.** Neither I nor II.
- E. I don't know.

6.5 REDUCTIONS

designing algorithms

classifying problems

establishing lower bounds

introduction

intractability

Algorithms

ROBERT SEDGEWICK | KEVIN WAYNE

http://algs4.cs.princeton.edu

Design algorithm. Given an algorithm for *Y*, can also solve *X*.

More familiar reductions.

. . .

- Mincut reduces to maxflow.
- Arbitrage reduces to negative cycles.
- Bipartite matching reduces to maxflow.
- Seam carving reduces to shortest paths in a DAG.
- Burrows-Wheeler transform reduces to suffix sort.

Mentality. Since I know how to solve *Y*, can I use that algorithm to solve *X*?

3-COLLINEAR. Given *N* distinct points in the plane, are there 3 (or more) that all lie on the same line?



Brute force N³. For all triples of points (p, q, r), check if they are collinear.

3-collinear reduces to sorting

Sorting-based algorithm. For each point *p*,

- Compute the slope that each other point *q* makes with *p*.
- Sort the N-1 points by slope.
- Collinear points are adjacent.



Shortest paths on edge-weighted graphs and digraphs

Proposition. Undirected shortest paths (with nonnegative weights) reduces to directed shortest path.



Pf. Replace each undirected edge by two directed edges.



Cost of solving undirected shortest paths. $E \log V + (E + V)$.



6.5 REDUCTIONS

introduction

intractability

Algorithms

establishing lower bounds

designing algorithms

Classifying problems

Robert Sedgewick | Kevin Wayne

http://algs4.cs.princeton.edu

Goal. Prove that a problem requires a certain number of steps. Ex. In decision tree model, any compare-based sorting algorithm requires $\Omega(N \log N)$ compares in the worst case.



Linear-time reductions

Def. Problem *X* linear-time reduces to problem *Y* if *X* can be solved with:

- Linear number of standard computational steps.
- Constant number of calls to Y.

Establish lower bound:

- If *X* takes $\Omega(N \log N)$ steps, then so does *Y*.
- If *X* takes $\Omega(N^2)$ steps, then so does *Y*.

Mentality.

- If I could easily solve *Y*, then I could easily solve *X*.
- I can't easily solve X.
- Therefore, I can't easily solve *Y*.

Which of the following reductions is not a linear-time reduction?

- A. ELEMENT-DISTINCTNESS reduces to SORTING.
- **B.** MIN-CUT reduces to MAX-FLOW.
- C. 3-COLLINEAR reduces to SORTING.
- D. BURROWS-WHEELER-TRANSFORM reduces to SUFFIX-SORTING.
- E. I don't know.

ELEMENT-DISTINCTNESS linear-time reduces to 2D-CLOSEST-PAIR

ELEMENT-DISTINCTNESS. Given *N* elements, are any two equal? 2D-CLOSEST-PAIR. Given *N* points in the plane, find the closest pair.



element distinctness

2d closest pair

ELEMENT-DISTINCTNESS linear-time reduces to 2D-CLOSEST-PAIR

ELEMENT-DISTINCTNESS. Given *N* elements, are any two equal? 2D-CLOSEST-PAIR. Given *N* points in the plane, find the closest pair.

Proposition. ELEMENT-DISTINCTNESS linear-time reduces to 2D-CLOSEST-PAIR. Pf.

- ELEMENT-DISTINCTNESS instance: x_1, x_2, \ldots, x_N .
- 2D-CLOSEST-PAIR instance: $(x_1, x_1), (x_2, x_2), \dots, (x_N, x_N)$.
- The *N* elements are distinct iff distance of closest pair > 0.

and quadratic tests like $(x_i - x_k)^2 + (x_j - x_k)^2 > 4$ ELEMENT-DISTINCTNESS lower bound. In quadratic decision tree model, any algorithm that solves ELEMENT-DISTINCTNESS takes $\Omega(N \log N)$ steps.

allows linear tests like $x_i < x_i$

any argonanti that solves element bistine inclus takes alon 10g 11) steps

Implication. In quadratic decision tree model, any algorithm for 2D-CLOSEST-PAIR takes $\Omega(N \log N)$ steps.

Some linear-time reductions in computational geometry



3-SUM. Given *N* distinct integers, are there three that sum to 0?

3-COLLINEAR. Given *N* distinct points in the plane, are there 3 (or more) that lie on the same line?



3-SUM. Given *N* distinct integers, are there three that sum to 0?

3-COLLINEAR. Given *N* distinct points in the plane, are there 3 (or more) that lie on the same line?

Proposition. 3-SUM linear-time reduces to 3-COLLINEAR. Pf. [next two slides]

lower-bound mentality:
if I can't solve 3-SUM in N^{1.99} time,
I can't solve 3-COLLINEAR
in N^{1.99} time either

Conjecture. Any algorithm for 3-SUM requires $\Omega(N^{2-\varepsilon})$ steps. Implication. No sub-quadratic algorithm for 3-COLLINEAR likely.

our $N^2 \log N$ algorithm was pretty good

3-SUM linear-time reduces to 3-COLLINEAR

Proposition. 3-SUM linear-time reduces to 3-COLLINEAR.

- 3-SUM instance: *x*₁, *x*₂, ..., *x*_N.
- 3-COLLINEAR instance: $(x_1, x_1^3), (x_2, x_2^3), \dots, (x_N, x_N^3)$.

Lemma. If *a*, *b*, and *c* are distinct, then a + b + c = 0if and only if $(a, a^3), (b, b^3)$, and (c, c^3) are collinear.



3-SUM linear-time reduces to 3-COLLINEAR

Proposition. 3-SUM linear-time reduces to 3-COLLINEAR.

- 3-SUM instance: $x_1, x_2, ..., x_N$.
- 3-COLLINEAR instance: $(x_1, x_1^3), (x_2, x_2^3), \dots, (x_N, x_N^3)$.

Lemma. If *a*, *b*, and *c* are distinct, then a + b + c = 0if and only if $(a, a^3), (b, b^3)$, and (c, c^3) are collinear.

Pf. Three distinct points (a, a^3) , (b, b^3) , and (c, c^3) are collinear iff:

$$0 = \begin{vmatrix} a & a^{3} & 1 \\ b & b^{3} & 1 \\ c & c^{3} & 1 \end{vmatrix}$$
$$= a(b^{3} - c^{3}) - b(a^{3} - c^{3}) + c(a^{3} - b^{3})$$

$$= (a-b)(b-c)(c-a)(a+b+c)$$

More geometric reductions and lower bounds



April 2014. Some recent evidence that the complexity might be $N^{3/2}$.

Threesomes, Degenerates, and Love Triangles*

Allan Grønlund MADALGO, Aarhus University Seth Pettie University of Michigan

April 4, 2014

Abstract

The 3SUM problem is to decide, given a set of n real numbers, whether any three sum to zero. We prove that the decision tree complexity of 3SUM is $O(n^{3/2}\sqrt{\log n})$, that there is a randomized 3SUM algorithm running in $O(n^2(\log \log n)^2/\log n)$ time, and a deterministic algorithm running in $O(n^2(\log \log n)^{2/3})$ time. These results refute the strongest version of the 3SUM conjecture, namely that its decision tree (and algorithmic) complexity is $\Omega(n^2)$.

Establishing lower bounds: summary

Establishing lower bounds through reduction is an important tool in guiding algorithm design efforts.

- Q. How to convince yourself no linear-time EUCLIDEAN-MST algorithm exists?
- A1. [hard way] Long futile search for a linear-time algorithm.
- A2. [easy way] Linear-time reduction from element distinctness.



6.5 REDUCTIONS

introduction

intractability

Algorithms

classifying problems

designing algorithms

establishing lower bounds

ROBERT SEDGEWICK | KEVIN WAYNE

http://algs4.cs.princeton.edu

Desiderata. Problem with algorithm that matches lower bound. Ex. Sorting and element distinctness have complexity $N \log N$.

Desiderata'. Prove that two problems X and Y have the same complexity. First, show that problem X linear-time reduces to Y.

- Second, show that *Y* linear-time reduces to *X*.
- Conclude that *X* has complexity N^b iff *Y* has complexity N^b for $b \ge 1$.





Integer arithmetic reductions

Integer multiplication. Given two *N*-bit integers, compute their product. Brute force. N^2 bit operations.



Integer arithmetic reductions

Integer multiplication. Given two *N*-bit integers, compute their product. Brute force. N^2 bit operations.

problem	arithmetic	order of growth
integer multiplication	$a \times b$	M(N)
integer division	$a \mid b, a \mod b$	M(N)
integer square	a ²	M(N)
integer square root	$\lfloor \sqrt{a} \rfloor$	M(N)

integer arithmetic problems with the same complexity as integer multiplication

Q. Is brute-force algorithm optimal?

History of complexity of integer multiplication

year	algorithm	order of growth
?	brute force	N^2
1962	Karatsuba	$N^{1.585}$
1963	Toom-3, Toom-4	$N^{1.465}, N^{1.404}$
1966	Toom-Cook	$N^{1+\epsilon}$
1971	Schönhage-Strassen	$N \log N \log \log N$
2007	Fürer	$N \log N 2^{\log^* N}$
?	?	N

number of bit operations to multiply two N-bit integers

used in Maple, Mathematica, gcc, cryptography, ...

Remark. GNU Multiple Precision Library uses one of five different algorithm depending on size of operands.



Numerical linear algebra reductions

Matrix multiplication. Given two *N*-by-*N* matrices, compute their product. Brute force. *N*³ flops.



 $0.5 \cdot 0.1 + 0.3 \cdot 0.0 + 0.9 \cdot 0.4 + 0.6 \cdot 0.1 = 0.47$

Numerical linear algebra reductions

Matrix multiplication. Given two *N*-by-*N* matrices, compute their product. Brute force. N^3 flops.

problem	linear algebra	order of growth
matrix multiplication	$A \times B$	MM(N)
matrix inversion	A^{-1}	MM(N)
determinant	A	MM(N)
system of linear equations	Ax = b	MM(N)
LU decomposition	A = L U	MM(N)
least squares	$\min \ Ax - b\ _2$	MM(N)

numerical linear algebra problems with the same complexity as matrix multiplication

Q. Is brute-force algorithm optimal?

History of complexity of matrix multiplication

year	algorithm	order of growth
?	brute force	N^3
1969	Strassen	$N^{2.808}$
1978	Pan	$N^{2.796}$
1979	Bini	$N^{2.780}$
1981	Schönhage	$N^{2.522}$
1982	Romani	$N^{2.517}$
1982	Coppersmith-Winograd	$N^{2.496}$
1986	Strassen	$N^{2.479}$
1989	Coppersmith-Winograd	$N^{2.376}$
2010	Strother	N ^{2.3737}
2012	Williams	N ^{2.372873}
2014	de Gall	$N^{2.372864}$
?	?	$N^{2+\epsilon}$

number of floating-point operations to multiply two N-by-N matrices

6.5 REDUCTIONS

designing algorithms

classifying problems

establishing lower bounds

Algorithms

intractability

introduction

Robert Sedgewick | Kevin Wayne

http://algs4.cs.princeton.edu

Def. A problem is intractable if it can't be solved in polynomial time. Desiderata. Prove that a problem is intractable.

Two problems that provably require exponential time.

- Given a constant-size program, does it halt in at most K steps?
- Given *N*-by-*N* checkers board position, can the first player force a win?

using forced capture rule

input size = c + lg K





Frustrating news. Very few successes.

SAT. Given a system of boolean equations, find a solution.

Ex.

$\neg x_1$	or	x_2	or	x_3			=	true	
x_1	or	$\neg x_2$	or	<i>x</i> ₃			=	true	
$\neg x_1$	or	$\neg x_2$	or	$\neg x_3$			=	true	
$\neg x_1$	or	$\neg x_2$	or		or	X_4	=	true	X1 X2 X3 X4
		$\neg x_2$	or	<i>x</i> ₃	or	X_4	=	true	TTFT
			ins	tance l					solution S

3-SAT. All equations of this form (with three variables per equation).

Key applications.

- Automatic verification systems for software.
- Mean field diluted spin glass model in physics.
- Electronic design automation (EDA) for hardware.

• ...

Satisfiability is conjectured to be intractable

- **Q.** How to solve an instance of 3-SAT with *N* variables?
- A. Exhaustive search: try all 2^N truth assignments.



Q. Can we do anything substantially more clever?

Conjecture ($P \neq NP$). 3-SAT is intractable (no poly-time algorithm).

Polynomial-time reductions

Problem *X* poly-time (Cook) reduces to problem *Y* if *X* can be solved with:

- Polynomial number of standard computational steps.
- Polynomial number of calls to Y.



Establish intractability. If 3-SAT poly-time reduces to *Y*, then *Y* is intractable. (assuming 3-SAT is intractable)

Mentality.

- If I could solve *Y* in poly-time, then I could also solve 3-SAT in poly-time.
- 3-SAT is believed to be intractable.
- Therefore, so is Y.

ILP. Given a system of linear inequalities, find an integral solution.



Context. Cornerstone problem in operations research.

Remark. Finding a real-valued solution is tractable (linear programming).

3-SAT. Given a system of boolean equations, find a solution.

$\neg x_1$	or	x_2	or	x_3			=	true
x_1	or	$\neg x_2$	or	<i>x</i> ₃			=	true
$\neg x_1$	or	$\neg x_2$	or	$\neg x_3$			=	true
$\neg x_1$	or	$\neg x_2$	or		or	<i>x</i> ₄	=	true
		$\neg x_2$	or	x_3	or	<i>x</i> ₄	=	true

ILP. Given a system of linear inequalities, find a 0-1 solution.

$(1 - x_1)$	+	x_2	+	x_3			≥	1
x_1	+	$(1 - x_2)$	+	<i>x</i> ₃			≥	1
$(1 - x_1)$	+	$(1 - x_2)$	+	$(1 - x_3)$			≥	1
$(1 - x_1)$	+	$(1 - x_2)$	+		+	<i>x</i> ₄	≥	1
		$(1 - x_2)$	+	<i>x</i> ₃	+	<i>x</i> ₄	≥	1

solution to this ILP instance gives solution to original 3-SAT instance

Reductions: quiz 3

Suppose that Problem *X* poly-time reduces to Problem *Y*. Which of the following can you infer?

- A. If *X* can be solved in poly-time, then so can *Y*.
- **B.** If *X* cannot be solved in cubic time, *Y* cannot be solved in poly-time.
- **C.** If *Y* can be solved in cubic time, then *X* can be solved in poly-time.
- **D.** If *Y* cannot be solved in poly-time, then neither can *X*.
- E. I don't know.

More poly-time reductions from 3-satisfiability



Implications of poly-time reductions from 3-satisfiability

Establishing intractability through poly-time reduction is an important tool in guiding algorithm design efforts.

Q. How to convince yourself that a new problem is (probably) intractable?A1. [hard way] Long futile search for an efficient algorithm (as for 3-SAT).A2. [easy way] Reduction from 3-SAT.

Caveat. Intricate reductions are common.



Search problem. Problem where you can check a solution in poly-time.

Ex 1. 3-SAT.

$\neg x_1$	Oľ	x_2	or	<i>x</i> ₃			=	true	
x_1	or	$\neg x_2$	or	<i>x</i> ₃			=	true	
$\neg x_1$	or	$\neg x_2$	or	$\neg x_3$			=	true	
$\neg x_1$	or	$\neg x_2$	or		or	<i>x</i> ₄	=	true	$x_1 x_2 x_3 x_4$
		$\neg x_2$	or	<i>x</i> ₃	or	<i>x</i> ₄	=	true	ТТЕТ
			ins	tance I					solution S

Ex 2. FACTOR. Given an *N*-bit integer *x*, find a nontrivial factor.

instance I	solution S
147573952589676412927	193707721

P vs. NP

P. Set of search problems solvable in poly-time.

Importance. What scientists and engineers can compute feasibly.

NP. Set of search problems (checkable in poly-time). Importance. What scientists and engineers aspire to compute feasibly.

Fundamental question.



Consensus opinion. No.

A problem is **NP-COMPLETE** if

- It is in NP.
- All problems in NP poly-time to reduce to it.

Cook-Levin theorem. 3-SAT is NP-COMPLETE. Corollary. 3-SAT is tractable if and only if P = NP.

Two worlds.



Implications of Cook-Levin theorem



Implications of Karp + Cook-Levin



Reductions: quiz 4

Suppose that X is **NP-COMPLETE**, Y is in **NP**, and X poly-time reduces to Y. Which of the following statements can you infer?

- I. Y is NP-COMPLETE.
- II. If *Y* cannot be solved in poly-time, then $P \neq NP$.
- III. If $P \neq NP$, then neither X nor Y can be solved in poly-time.
- A. I only.
- **B.** II only.
- C. I and II only.
- D. I, II, and III.
- E. I don't know.

Desiderata. Classify problems according to computational requirements.

complexity	order of growth	examples		
linear	Ν	min, max, median, Burrows-Wheeler transform,		
linearithmic	$N \log N$	sorting, element distinctness,		
quadratic	N^{2}	?		
÷	÷	• • •		
exponential	<i>c ^{<i>N</i>}</i>	?		

Frustrating news. Huge number of problems have defied classification.

Desiderata. Classify problems according to computational requirements.

complexity	order of growth	examples
linear	Ν	min, max, median, Burrows-Wheeler transform,
linearithmic	N log N	sorting, element distinctness,
M(N)	?	integer multiplication, division, square root,
MM(N)	?	matrix multiplication, Ax = b, least square, determinant,
÷	÷	:
NP-complete	probably not N^{b}	3-SAT, IND-SET, ILP,

Good news. Can put many problems into equivalence classes.

Complexity class. Set of problems sharing some computational property.



https://complexityzoo.uwaterloo.ca

Bad news. Lots of complexity classes (496 animals in zoo).

Summary

Reductions are important in theory to:

- Design algorithms.
- Establish lower bounds.
- Classify problems according to their computational requirements.

Reductions are important in practice to:

- Design algorithms.
- Design reusable software modules.
 - stacks, queues, priority queues, symbol tables, sets, graphs
 - sorting, regular expressions, suffix arrays
 - MST, shortest paths, maxflow, linear programming
- Determine difficulty of your problem and choose the right tool.



