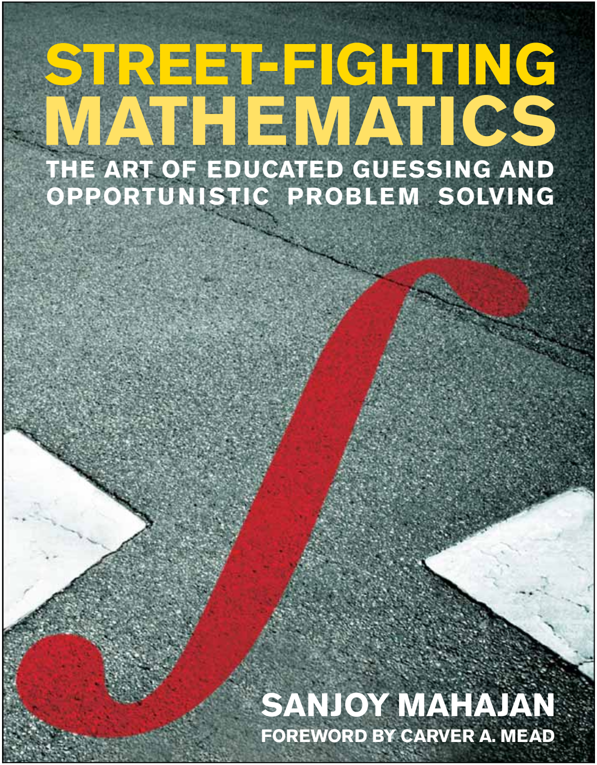
Typesetting a science and engineering textbook

Sanjoy Mahajan

MIT & Olin College

mit.edu/sanjoy/www/
sanjoy@mit.edu

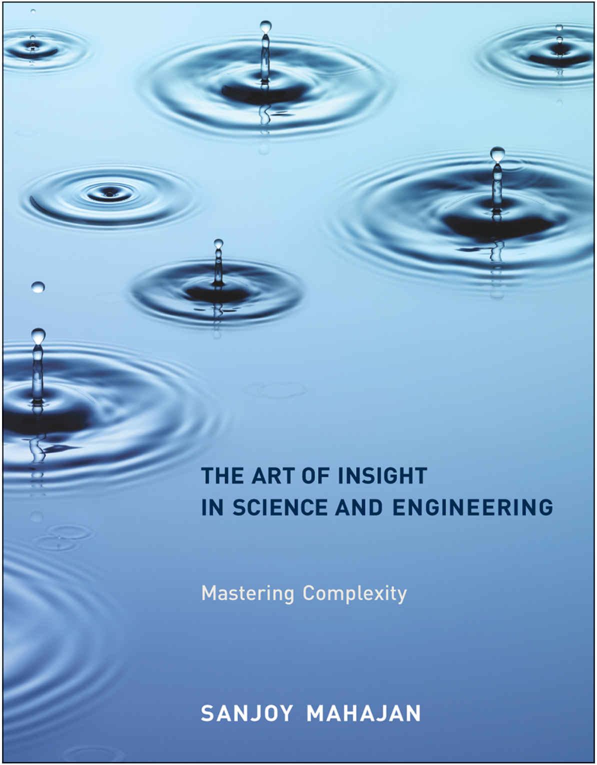
ConTeXt meeting, Bassenge, Belgium, 10 September 2014



My first book

ConTeXt Mark II

(MIT Press, 2010)



My second book

ConTeXt Mark IV

(MIT Press, 2014)

The Art of Insight in Science and Engineering

Mastering Complexity

Sanjoy Mahajan

The MIT Press
Cambridge, Massachusetts
London, England

book.tex

```
\startproduct book
\project project

\startfrontmatter
\setuppagenumbering[style=normal]
\setupheadertexts[\setups{justchnameheadline}][pagenumber]%
                 [pagenumber][\setups{justchnameheadline}]

\setupuserpagenumber[numberconversion=romannumerals]
\setupheader[style=normal]

\component titlepages
\component dedication
\component contents
\component preface
\component constants
```

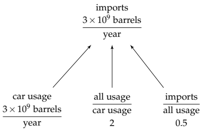
```
\stopfrontmatter

\startbodymatter

\setupheadertexts[\setups{secheadline}][pagenumber]%
                 [pagenumber][\setups{chheadline}]

\setuppagenumbering[style=bold]
\setupuserpagenumber[numberconversion=numbers]
\setuppagenumber[number=1]  % make =1 if preface has even no. of pages

% part 1
\component organizing-complexity-part
\component divide-and-conquer
\component abstraction

% part 2
\component discarding-spurious-complexity-part
\component symmetry-and-conservation
```

```
\component proportional-reasoning
\component dimensions

% part 3
\component discarding-actual-complexity-part
\component lumping
\component probabilistic-reasoning
\component easy-cases
\component springs

\stopbodymatter

\startbackmatter

\writebetweenlist [part] {\EnglishRule}

\component how-to-learn

\setupheadertexts[][pagenumber][pagenumber][]
```

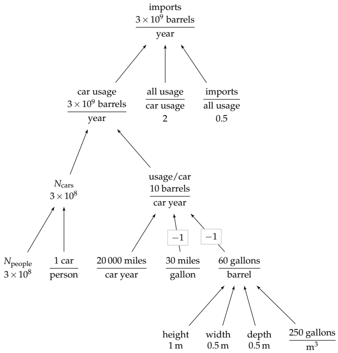
```
\chapter{Bibliography}
{\setuplayout[footer=48bp]      %hack to avoid slightly short pages
\raggedbottom
\switchtobodyfont[9pt]
\raggedright
\newfrenchspacing
\placepublications[criterium=text]}

\chapter{Index}
\setuplayout[footer=60bp]       %restore old value
\input index-xrefs
{\it An italic page number refers to a problem on that page.}
\blank[3*big]
{%\setuplayout[width=5.25in]
\setuptolerance[vertical,tolerant]
\switchtobodyfont[9pt]
\placeindex[criterium=all]}
```

Lots of tree diagrams

This estimate is itself a subtree in the tree representing oil imports. Because the two adjustment factors contribute a factor of 2×0.5, which is just 1, the oil imports are also 3 billion barrels per year.



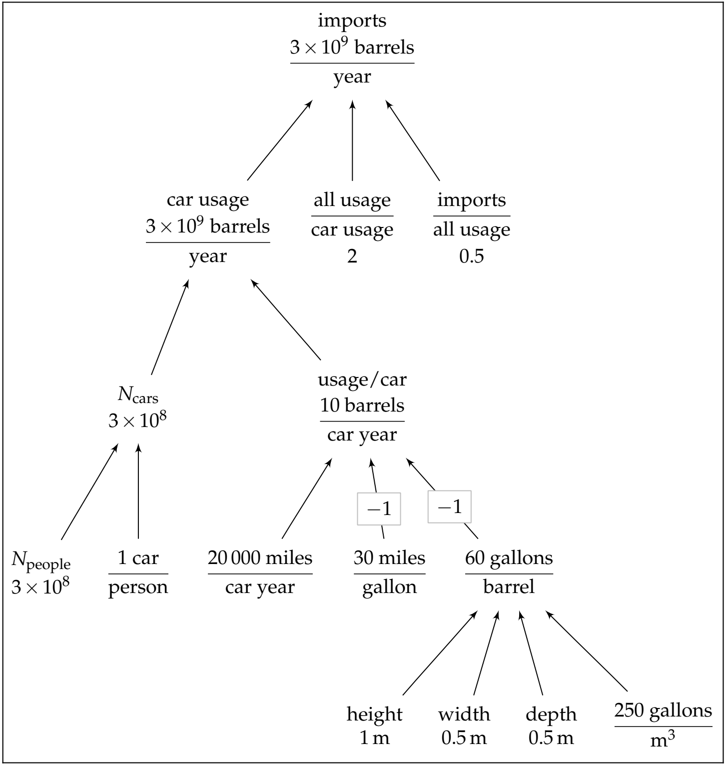
Here is the full tree, which includes the subtree for the total car usage of oil:



Problem 1.6 Using metric units

As practice with metric units (if you grew up in a nonmetric land) or to make the results more familiar (if you grew up in a metric land), redo the calculation using the metric values for the volume of a barrel, the distance a car is driven per year, and the fuel consumption of a typical car.

▶ *How close is our estimate to official values?*


imports
3 × 10^9 barrels / year
car usage
3 × 10^9 barrels / year
all usage / car usage
2
imports / all usage
0.5
N_cars
3 × 10^8
usage/car
10 barrels / car year
−1
−1
N_people
3 × 10^8
1 car / person
20 000 miles / car year
30 miles / gallon
60 gallons / barrel
height
1 m
width
0.5 m
depth
0.5 m
250 gallons / m^3

Tree minilanguage

```
# nodesep=0.1
# ranksep=0.45
imports|$\hoverh{$3\edot9$ barrels}{year}$
  car usage|$\hoverh{$3\edot9$ barrels}{year}$
    $N_{\rm cars}$|$3\edot8$
      $N_{\rm people}$|$3\edot8$
      $\hoverh{1 car}{person}$
    usage/car|$\hoverh{10 barrels}{car year}$
      $\hoverh{20\,000 miles}{car year}$
      [-1]$\hoverh{30 miles}{gallon}$
      [-1]$\hoverh{60 gallons}{barrel}$
        height|$1\m$
        width|$0.5\m$
        depth|$0.5\m$
        $\hoverh{250 gallons}{m$^3$}$
  $\hoverh{all usage}{car usage}$|2
  $\hoverh{imports}{all usage}$|0.5
```

Compiler to convert tree to dot language

```
#!/usr/bin/env python

# turn indented outline representation of a tree into metapost boxes
cmds
# e.g.
#   A
#     B
#     C
# is a tree with A at the root and B, C as the kids.
#
# lines beginning with # are directives; for now, only ones are
# dir=forward or dir=back or dir=none

from re import split, match, sub
from sys import stdin, argv, stderr

fontsize = r"\footnotesize "
```

```
graph_attrs = {'ratio' : 'compress'}
node_attrs  = {'label': r'',
               'shape' : 'plaintext'}
edge_attrs  = {'dir' : 'back',
               'lblstyle' : 'draw=lightgray,fill=white',
               'labelangle' : '0',
               'labelfloat' : 'true'}
...
```

Dot (graphviz) file

```
strict digraph {
graph [ranksep="0.45", ratio="compress", nodesep="0.1"];
edge  [labelfloat="true", lblstyle="draw=lightgray,fill=white", dir="
labelangle="0"];
node  [shape="plaintext", label=""];

node_0 [texlbl="\footnotesize \2{\strut imports}{\strut $\hoverh{$3\e
barrels}{year}$}"];
  node_0_0 [texlbl="\footnotesize \2{\strut car usage}{\strut $\hover
barrels}{year}$}"];
    node_0_0_0 [texlbl="\footnotesize \2{\strut $N_{\rm cars}$}{\stru
$3\edot8$}"];
      node_0_0_0_0 [texlbl="\footnotesize \2{\strut $N_{\rm people}$}
$3\edot8$}"];
      node_0_0_0_1 [texlbl="\footnotesize \strut $\hoverh{1 car}{pers
...
```

Revision control

```
$ hg log -r 1:tip
changeset:   1:792a81645ae3
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Mon Apr 08 10:25:05 2013 -0400
summary:     first version that compiles

changeset:   2:162a877c400b
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Mon Apr 08 10:30:15 2013 -0400
summary:     paste in preface

changeset:   3:3df2bc462f16
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Mon Apr 08 20:00:20 2013 -0400
summary:     util/tree2mp.py: 2010 version: horizontal growth, exponen

changeset:   4:1765a8f1e689
```

```
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Mon Apr 08 20:00:45 2013 -0400
summary:     divide and conquer: trees and text

...

changeset:   955:2e29d1f98611
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Fri Jun 27 17:01:06 2014 -0400
summary:     tag as v0.9 the version for MIT Press to start productio

...

changeset:   1057:78ca0ee9dfae
tag:         v0.10
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Tue Sep 02 06:51:35 2014 -0400
summary:     preface: small fix
```

```
changeset:   1058:503ec3115158
user:        Sanjoy Mahajan <sanjoy@olin.edu>
date:        Tue Sep 02 06:58:25 2014 -0400
summary:     tag as v0.10: w/ fixes from DJCM's comments, give to MIT
Press
```

Make controls everything

```
$ make -n
cd fig/ && asy -noprc -render=0 -tex lualatex -f pdf air-slab.asy
cd fig/ && asy -noprc -render=0 -tex lualatex -f pdf gm.asy
cd fig/ && asy -noprc -render=0 -tex lualatex -f pdf rc-teacup.asy
cd fig/ && asy -noprc -render=0 -tex lualatex -f pdf mouse-eaten-boar
cd fig/ && asy -noprc -render=0 -tex lualatex -f pdf parabola-max.asy
...
convert fig/British_Isles_United_Kingdom.svg fig/British_Isles_United_
...
cd fig/ && asy -noprc -render=0 -tex pdflatex -f pdf spring-approx-or
...
util/tree2dot.py < fig/oil-imports-three-leaves-with-numbers.tree >
fig/oil-imports-three-leaves-with-numbers.dot
dot2tex --debug -t raw -f tikz --tikzedgelabel -c --autosize --docpre
"\input /home/sanjoy/sfse-context-mtg-talk/fig/dot_template " fig/oil-
> fig/oil-imports-three-leaves-with-numbers.dottex
sed -i '1,10s/\[x11names, rgb\]/\[x11names, gray\]/' fig/oil-imports-
```

```
...
cd fig/ && pdflatex oil-imports-three-leaves-with-numbers.dottex
...
cat fig/template.mp fig/defs.mp > fig/house-circuit.mp
echo "beginfig(1)" >> fig/house-circuit.mp
cat fig/house-circuit.mpfig >> fig/house-circuit.mp
echo -e "endfig;\nend" >> fig/house-circuit.mp
cd fig/ && mpost -mem=metafun house-circuit.mp | grep -v 'duplicates
ignored'
mptopdf fig/house-circuit.1
...
cd ./ && PATH=/home/sanjoy/context/2014.05.17-delta/tex/texmf-linux/b
context  --nonstopmode --passon="-file-line-error -recorder -interact
nonstopmode" --mode=print --result="book-print.pdf" book.tex > all.lo
|| ( /home/sanjoy/sfse-context-mtg-talk/util/tex-errors.sh < all.log
; false )
```

Directory organization

```
abstraction.tex
book.tex
contents.tex
COPYING
dimensions.tex
easy-cases.tex
env.tex
fig/
  asy_defs.tex
  barbell-rolling-down-plane.asy
  British_Isles_United_Kingdom.svg
  defs.mp
  driven-capacitor.mpfig
  l_vap-water-energy-subdivided.tree
index-xrefs.tex
lumping.tex
Makefile
```

```
preface.tex
PREMISE
premise.tex
probabilistic-reasoning.tex
project.tex
proportional-reasoning.tex
README
refs.bib
SOMEDAY
springs.tex
symmetry-and-conservation.tex
TeXGyrePagellaMath-Regular.otf
titlepages.tex
TODO
util/
  check-various.py*
  overfull-awk-script
  tex-errors.sh*
  tree2dot.py*
```

1.7 Physical estimates

Your gut understands not only the social world but also the physical world. If you trust its feelings, you can tap this vast reservoir of knowledge. For practice, we'll estimate the salinity of seawater (Section 1.7.1), human power output (Section 1.7.2), and the heat of vaporization of water (Section 1.7.3).

1.7.1 Salinity of seawater

To estimate the salinity of seawater, which will later help you estimate the conductivity of seawater (Problem 8.10), do not ask your gut directly: "How do you feel about, say, 200 millimolar?" Although that kind of question worked for estimating population density (Section 1.6), here, unless you are a chemist, the answer will be: "I have no clue. What is a millimolar anyway? I have almost no experience of that unit." Instead, offer your gut concrete data—for example, from a home experiment: adding salt to a cup of water until the mixture tastes as salty as the ocean.

This experiment can be a thought or a real experiment—another example of using multiple methods (Section 1.5). As a thought experiment, I ask my gut about various amounts of salt in a cup of water. When I propose adding 2 teaspoons, it responds, "Disgustingly salty!" At the lower end, when I propose adding 0.5 teaspoons, it responds, "Not very salty." I'll use 0.5 and 2 teaspoons as the lower and upper endpoints of the range. Their midpoint, the estimate from the thought experiment, is 1 teaspoon per cup.

I tested this prediction at the kitchen sink. With 1 teaspoon (5 milliliters) of salt, the cup of water indeed had the sharp, metallic taste of seawater that I have gulped after being knocked over by large waves. A cup of water is roughly one-fourth of a liter or 250 cubic centimeters. By mass, the resulting salt concentration is the following product:

$$\frac{\cancel{1\,\text{tsp salt}}}{\cancel{1\,\text{cup water}}} \times \frac{\cancel{1\,\text{cup water}}}{250\,\text{g water}} \times \frac{\cancel{5\,\text{cm}^3\,\text{salt}}}{\cancel{1\,\text{tsp salt}}} \times \underbrace{\frac{2\,\text{g salt}}{\cancel{1\,\text{cm}^3\,\text{salt}}}}_{\rho_{\text{salt}}}. \tag{1.19}$$

The density of 2 grams per cubic centimeter comes from my gut feeling that salt is a light rock, so it should be somewhat denser than water at 1 gram per cubic centimeter, but not too much denser. (For an alternative method, more accurate but more elaborate, try Problem 1.10.) Then doing the arithmetic gives a 4 percent salt-to-water ratio (by mass).

Using Metafun strike-through lines to check calculation

Using Metafun strike-through lines to check calculation

```
% units
\startuniqueMPgraphic{strikeme}
   path p ; p := OverlayBox topenlarged -ExHeight bottomenlarged -ExHe
;
   draw llcorner p -- urcorner p withcolor 0.5white withpen pencircle
scaled 0.630pt; % transparent(1,0.5,black) ;
\stopuniqueMPgraphic

\defineoverlay[strikeme][\uniqueMPgraphic{strikeme}]

\def\strike#1{\inframed[background=strikeme,frame=off]{#1}}
\def\mstrike#1{\strike{$#1$}}
```

Doubling r quadruples the amount of paper used to make the cone and therefore its mass m. It also quadruples its cross-sectional area A_{cs}. According to the proportionality, the two effects cancel: When r doubles, v should remain constant. All cones of the same shape (and made from the same paper) should fall at the same speed!

This result always surprises me. So I tried the experiment. I printed the cone template in Section 3.5.2 at 400-percent magnification (a factor of 4 increase in length), cut it out, taped the two straight edges together, and raced the small and big cones by dropping them from a height of about 2 meters. After a roughly 2-second fall, they landed almost simultaneously—within 0.1 seconds of each other. Thus, their terminal speeds are the same, give or take 5 percent.

Proportional reasoning triumphs again! Surprisingly, the proportional-reasoning result is much more accurate than the drag-force estimate $\rho_{air} A_{cs} v^2$ on which it is based.

▶ *How can predictions based on proportional reasoning be more accurate than the original relations?*

To see how this happy situation arose, let's redo the calculation but include the dimensionless prefactor in the drag force. With the dimensionless prefactor (shaded in gray), the drag force is

$$F_{drag} = \boxed{\frac{1}{2}c_d}\, \rho_{air} A_{cs} v^2, \tag{4.45}$$

where c_d is the drag coefficient (introduced in Section 3.2.1). The prefactor carries over to the terminal speed:

$$v = \sqrt{\frac{m}{\boxed{\frac{1}{2}c_d}\, A_{cs}}}. \tag{4.46}$$

Ignoring the prefactor decreases v by a factor of $\sqrt{2/c_d}$. (For nonstreamlined objects, $c_d \sim 1$, so the decrease is roughly by a factor of $\sqrt{2}$.)

In contrast, in the ratio of terminal speeds v_{big}/v_{small}, the prefactor drops out. Here is the ratio with the prefactors shaded in gray:

$$\frac{v_{big}}{v_{small}} = \sqrt{\frac{m_{big}}{\boxed{\frac{1}{2}c_d^{big}}\, A_{cs}^{big}}} \Bigg/ \sqrt{\frac{m_{small}}{\boxed{\frac{1}{2}c_d^{small}}\, A_{cs}^{small}}}. \tag{4.47}$$

Emphasizing key parts of an equation

Emphasizing key parts of an equation

```
\definemathframed
  [graymath]
  [frame=off,location=mathematics,
    background=color,backgroundcolor=verylight,offset=2pt]
```

Small, light equation numbers

The drag coefficient is

$$c_{\mathrm{d}} \equiv \frac{F_{\mathrm{drag}}}{\frac{1}{2}\rho_{\mathrm{air}} v^2 A_{\mathrm{cs}}} \tag{5.52}$$

The cone falls at its terminal speed, so the drag force is also its weight W:

$$F_{\mathrm{drag}} = W = A_{\mathrm{paper}}\sigma_{\mathrm{paper}}\, g, \tag{5.53}$$

where σ_{paper} is the areal density (mass per area) of paper and A_{paper} is the area of the cone template. The drag coefficient is then

$$c_{\mathrm{d}} = \frac{\overbrace{A_{\mathrm{paper}}\sigma_{\mathrm{paper}}\, g}^{F_{\mathrm{drag}}}}{\frac{1}{2}\rho_{\mathrm{air}} A_{\mathrm{cs}} v^2}. \tag{5.54}$$

As we showed in Section 3.5.2,

$$A_{\mathrm{cs}} = \frac{3}{4} A_{\mathrm{paper}}. \tag{5.55}$$

This proportionality means that the areas cancel out of the drag coefficient:

$$c_{\mathrm{d}} = \frac{\sigma_{\mathrm{paper}} g}{\frac{1}{2}\rho_{\mathrm{air}} \times \frac{3}{4} v^2}. \tag{5.56}$$

To compute c_{d}, plug in the areal density $\sigma_{\mathrm{paper}} \approx 80$ grams per square meter and the measured speed $v \approx 1$ meter per second:

$$c_{\mathrm{d}} \approx \frac{\overbrace{8\times 10^{-2}\,\mathrm{kg\,m^{-2}}}^{\sigma_{\mathrm{paper}}} \times \overbrace{10\,\mathrm{m\,s^{-2}}}^{g}}{\frac{1}{2} \times \underbrace{1.2\,\mathrm{kg\,m^{-3}}}_{\rho_{\mathrm{air}}} \times \frac{3}{4} \times \underbrace{1\,\mathrm{m^2\,s^{-2}}}_{v^2}} \approx 1.8. \tag{5.57}$$

Because no quantity in this calculation depends on the cone's size, both cones have the same drag coefficient. (Our estimated drag coefficient is significantly larger than the canonical drag coefficient for a solid cone, roughly 0.7, and is approximately the drag coefficient for a wedge.)

Thus, the drag coefficient is independent of Reynolds number—at least, for Reynolds numbers between 3500 and 7000. The giant-cone experiment of Problem 4.16 shows that the independence holds even to Re $\sim 14\,000$. Within this range, the dimensionless function f in

$$\text{drag coefficient} = f_{\mathrm{cone}}(\text{Reynolds number}) \tag{5.58}$$

is a constant. What a simple description of the complexity of fluid flow!

Small, light equation numbers

```
\newdimen\mathindent \mathindent=20pt
\setupformulas[indentnext=auto, spacebefore=none, spaceafter=none]
\setupformulae[align=right,leftmargin=\mathindent,
  left={\startcolor[eqnumbercolor]\relax (},right={)\stopcolor},
  numberstyle=small,
]
```

Chapter page with local table of contents

7

Probabilistic reasoning

7.1	Probability as degree of belief: Bayesian probability	235
7.2	Plausible ranges: Why divide and conquer works	239
7.3	Random walks: Viscosity and heat flow	249
7.4	Transport by random walks	263
7.5	Summary and further problems	276

Our previous tool, lumping, helps us simplify by discarding less important information. Our next tool, probabilistic reasoning, helps us when our information is already incomplete—when we've discarded even the chance or the wish to collect the missing information.

7.1 Probability as degree of belief: Bayesian probability

The essential concept in using probability to simplify the world is that probability is a degree of belief. Therefore, a probability is based on our knowledge, and it changes when our knowledge changes.

7.1.1 Is it my telephone number?

Here is an example from soon after I had moved to England. I was talking to a friend on the phone, of the old-fashioned variety with wires connecting it to the wall. David needed to call me back. However, having just moved to the apartment, I was unsure of my phone number; plus, for anyone used to American phone numbers, British phone numbers have a strange and hard-to-remember format. I had a reasonably likely guess, which I gave David so that he could call me back. After I hung up, I tested my guess by picking up my phone and dialing my guess—and got a busy signal.

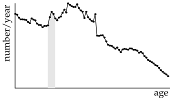
Local table of contents

```
\newbox\sectionlistbox
% include local contents (sections) iff there are any
\def\localcontents{%
\setbox\sectionlistbox=\vbox{\placelist[section,summarysection]}
\ifdim\ht\sectionlistbox>1em    %heuristic to identify whether not ju
whitespace
  \startframedtext[frame=off, background=color,
    offset=6pt,
    backgroundcolor=verylight, width=broad]\relax
    {\setuplist[section][margin=0em]
     \switchtobodyfont[9pt]
     \placelist[section,summarysection]}
  \stopframedtext
  \blank[3*big]
\fi
}
```

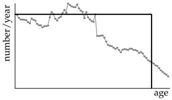
```
\setuphead[chapter,title][prefix=+,
  before={\bgroup\switchtobodyfont[24pt]\relax
    \hyphenpenalty=10000},
  textcommand={\smallskip\setupinterlinespace[line=2ex]%
    \doit},
  deepnumbercommand=\addch,
  style={\ss\bf},
  after={\vskip1.5in\egroup\localcontents}]
```

Many figures (roughly 310)

The first step is to estimate the number of people in the United States who are 18, 19, 20, or 21 years old. This total provides, at least in the United States, the pool from which most undergraduate students come. Because not all 18-to-21-year-olds go to college, at the end we will multiply the total by the fraction of adults who are college graduates.

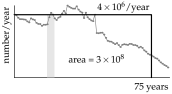


Finding the exact pool size requires the birth date of every person in the United States. Although these data are collected once every decade by the US Census Bureau, they would only overwhelm us. As an approximation to the voluminous data, the Census Bureau also publishes the number of people at each age. For example, the 1991 data are the wiggly line in the graph. The left side of the graph represents the number of infants and toddlers in 1991, and the right side represents the number of older people (also in 1991). The undergraduate pool size, representing all 18-, 19-, 20-, and 21-year-olds, is the shaded area. (The peak around the ages 30–35 represents the baby boomers, born in the period after World War Two.)



Unfortunately, even this graph depends on the huge resources of the Census Bureau, so it is not suited for back-of-the-envelope estimates. It also provides little insight or transfer value. Insight comes from lumping: from turning the complex, wiggly curve into a rectangle. The rectangle's dimensions can be determined without any information from the Census Bureau.

▶ *What are the height and width of this rectangle?*



The rectangle's width is a time, so it must be a characteristic time related to the population. A good guess is the life expectancy, because the age distribution varies significantly over that time. In the United States, the life expectancy is roughly 75 years, which will be the rectangle's width. In this lumping approximation, everyone lives happily until a sudden death at his or her 75th birthday. This all-or-nothing reasoning is the essential characteristic of lumping, making it such a useful approximation.

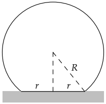
Problem 6.23 Sketching the actual and lumped deflection angle
On axes for cumulative deflection θ versus distance along the beam, sketch (a) the actual curve, (b) the lumped curve, assuming that the deflection happens only while the beam is near the Sun, and (c) the lumped curve, assuming, as in the text, that the deflection happens only at the point of closest approach.

6.4.7 All-or-nothing reasoning: Solid mechanics by lumping

For estimating the bending of light, the heart of the lumping analysis was all-or-nothing reasoning: replacing the complex, varying downward acceleration with a simpler curve that was either zero or a nonzero constant. To practice this idea, we'll apply it to an example from solid mechanics, also a subject fraught with differential equations. In particular, we'll estimate the contact radius of a solid ball resting on the ground.

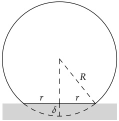
We know a bit about the contact radius: In Problem 5.50, you used dimensional analysis to find that the contact radius r is given by

$$\frac{r}{R} = f\left(\frac{\rho g R}{Y}\right), \tag{6.47}$$



where R is the ball's radius, ρ is its density, Y is its Young's modulus, and f is a dimensionless function. The function f is not determined by dimensional analysis, which is purely mathematical reasoning. Finding f requires a physical model; the easiest way to make and analyze such a model is by making lumping approximations.

Physically, the ground compresses the tip of the ball by a small distance δ, making a flat circle of radius r in contact with the ground. The ball fights back, trying to restore its natural, spherical shape. When the ball rests on the table, the restoring force equals its weight. This constraint will give us enough information to find the dimensionless function f.



The restoring force comes from the stress (or pressure) over the contact surface. To estimate this stress, let's make the lumping approximation that it is constant over the contact surface and equal to a typical or characteristic stress. This approximation is analogous to replacing the varying population curve with a constant value (and making a rectangle). With that approximation, the restoring force is

Is this σ_{paper} consistent with the estimates for a dollar bill in Section 1.1?

There we estimated that the thickness t of a dollar bill, or of paper in general, is approximately 0.01 centimeters. The regular (volume) density ρ would then be 0.8 grams per cubic centimeter:

$$\rho_{\text{paper}} = \frac{\sigma_{\text{paper}}}{t} \approx \frac{80\,\text{g}\,\text{m}^{-2}}{10^{-2}\,\text{cm}} \times \frac{1\,\text{m}^2}{10^4\,\text{cm}^2} = 0.8\,\frac{\text{g}}{\text{cm}^3}. \tag{3.67}$$

This density, slightly below the density of water, is a good guess for the density of paper, which originates as wood (which barely floats on water). Therefore, our estimate in Section 1.1 is consistent with the proposed areal density of 80 grams per square meter.

After putting in the constants, the cone's terminal speed is predicted to be roughly 0.9 meters per second:

$$v_{\text{term}} \sim \left(\frac{4}{3} \times \frac{\overbrace{8\times10^{-2}\,\text{kg}\,\text{m}^{-2}}^{\sigma_{\text{paper}}} \times \overbrace{10\,\text{m}\,\text{s}^{-2}}^{g}}{\underbrace{1.2\,\text{kg}\,\text{m}^{-3}}_{\rho_{\text{air}}}} \right)^{1/2} \approx 0.9\,\text{m}\,\text{s}^{-1}. \tag{3.68}$$

To test the prediction and, with it, the analysis justifying it, I held the cone slightly above my head, from about 2 meters high. After I let the cone go, it fell for almost exactly 2 seconds before it hit the ground—for a speed of roughly 1 meter per second, very close to the prediction. Box models and conservation triumph again!

3.5.3 Cycling

In introducing the analysis of drag, I said that drag is one of the most important physical effects in everyday life. Our analysis of drag will now help us understand the physics of a fantastically efficient form of locomotion—cycling (for its efficiency, see Problem 3.34).

What is the world-record cycling speed?

The first task is to define the kind of world record. Let's analyze cycling on level ground using a regular bicycle, even though faster speeds are possible riding downhill or on special bicycles. In bicycling, energy goes into rolling resistance, friction in the chain and gears, and air drag. The importance of drag rises rapidly with speed, due to the factor of v^2 in the drag force, so at high-enough speeds drag is the dominant consumer of energy.

Annotating equations with over- and under-braces

Annotating equations with over- and underbraces

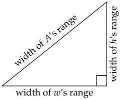
```
\setupmathstackers[vfenced][hoffset=3pt]
\def\tallest#1{\def\thetallest{#1}}
\def\support{\vphantom{\thetallest}}

\placeformula\startformula
\tallest{\hbox{energy density}}
E\fuel \sim \underbrace{\hbox{energy density}}_{\Edensity\fuel} \x
            \underbrace{\support\hbox{fuel mass}}_{\rho\fuel\,V\fuel}
= \Edensity\fuel\,\rho\fuel V\fuel.
\stopformula
```

Font hacking

(These decibels are slightly more general than the acoustic decibels introduced in Problem 3.10: Acoustic decibels measure energy flux relative to a reference value, usually 10^{-12} watts per square meter. Both kinds of decibels measure factors of 10, but the decibels here have no implicit reference value.)

In decibels, bels, or any logarithmic unit, the half width (the σ) of the product's range is the Pythagorean sum of the individual half widths (the σ values). Using σ_x to represent the half width of the plausible range for the quantity x, the recipe is

$$\sigma_A = \sqrt{\sigma_h^2 + \sigma_w^2}. \tag{7.15}$$



Let's apply this recipe to our example. The plausible range for the height (h) was 800 kilometers give or take a factor of 1.2. On a logarithmic scale, distances are measured by ratios or factors, so think of a range as "give or take a factor of" rather than as "plus or minus" (a description that would be appropriate on a linear scale). A factor of 1.2 is about ± 0.8 decibels:

$$10 \log_{10} 1.2 \approx 0.8. \tag{7.16}$$

Therefore, $\sigma_h \approx 0.8$ decibels.

The plausible range for the width (w) was roughly 310 kilometers give or take a factor of 1.3. A factor of 1.3 is ± 1.1 decibels:

$$10 \log_{10} 1.3 \approx 1.1. \tag{7.17}$$

Therefore, $\sigma_w \approx 1.1$ decibels.

The Pythagorean sum of σ_h and σ_w is approximately 1.4 decibels:

$$\sqrt{0.8^2 + 1.1^2} \approx 1.4. \tag{7.18}$$

As a factor, 1.4 decibels is, coincidentally, approximately a factor of 1.4:

$$10^{1.4/10} \approx 1.4. \tag{7.19}$$

Because the midpoint of the plausible range is 250 000 kilometers, the UK land area should be 250 000 square kilometers give or take a factor of 1.4. Retaining a bit more accuracy, it is a factor of 1.37.

$$\underbrace{180\,000}_{/1.37} \ldots \underbrace{250\,000}_{\text{midpoint}} \ldots \underbrace{340\,000}_{\times 1.37} \text{ km}^2. \tag{7.20}$$

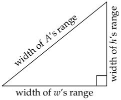
As a probability bar, the range is

Font hacking

(These decibels are slightly more general than the acoustic decibels introduced in Problem 3.10: Acoustic decibels measure energy flux relative to a reference value, usually 10^{-12} watts per square meter. Both kinds of decibels measure factors of 10, but the decibels here have no implicit reference value.)

In decibels, bels, or any logarithmic unit, the half width (the σ) of the product's range is the Pythagorean sum of the individual half widths (the σ values). Using σ_x to represent the half width of the plausible range for the quantity x, the recipe is

$$\sigma_A = \sqrt{\sigma_h^2 + \sigma_w^2}. \quad (7.15)$$



Let's apply this recipe to our example. The plausible range for the height (h) was 800 kilometers give or take a factor of 1.2. On a logarithmic scale, distances are measured by ratios or factors, so think of a range as "give or take a factor of" rather than as "plus or minus" (a description that would be appropriate on a linear scale). A factor of 1.2 is about ±0.8 decibels:

$$10 \log_{10} 1.2 \approx 0.8. \quad (7.16)$$

Therefore, $\sigma_h \approx 0.8$ decibels.

The plausible range for the width (w) was roughly 310 kilometers give or take a factor of 1.3. A factor of 1.3 is ±1.1 decibels:

$$10 \log_{10} 1.3 \approx 1.1. \quad (7.17)$$

Therefore, $\sigma_w \approx 1.1$ decibels.

The Pythagorean sum of σ_h and σ_w is approximately 1.4 decibels:

$$\sqrt{0.8^2 + 1.1^2} \approx 1.4. \quad (7.18)$$

As a factor, 1.4 decibels is, coincidentally, approximately a factor of 1.4:

$$10^{1.4/10} \approx 1.4. \quad (7.19)$$

Because the midpoint of the plausible range is 250 000 kilometers, the UK land area should be 250 000 square kilometers give or take a factor of 1.4. Retaining a bit more accuracy, it is a factor of 1.37.

$$\underbrace{180\,000}_{/1.37} \ldots \underbrace{250\,000}_{\text{midpoint}} \ldots \underbrace{340\,000}_{\times 1.37} \text{ km}^2. \quad (7.20)$$

As a probability bar, the range is

Careful checking

(These decibels are slightly more general than the acoustic decibels introduced in Problem 3.10: Acoustic decibels measure energy flux relative to a reference value, usually 10^{-12} watts per square meter. Both kinds of decibels measure factors of 10, but the decibels here have no implicit reference value.)

In decibels, bels, or any logarithmic unit, the half width (the σ) of the product's range is the Pythagorean sum of the individual half widths (the σ values). Using σ_x to represent the half width of the plausible range for the quantity x, the recipe is

$$\sigma_A = \sqrt{\sigma_h^2 + \sigma_w^2}. \tag{7.15}$$

Let's apply this recipe to our example. The plausible range for the height (h) was 800 kilometers give or take a factor of 1.2. On a logarithmic scale, distances are measured by ratios or factors, so think of a range as "give or take a factor of" rather than as "plus or minus" (a description that would be appropriate on a linear scale). A factor of 1.2 is about ±0.8 decibels:

$$10 \log_{10} 1.2 \approx 0.8. \tag{7.16}$$

Therefore, $\sigma_h \approx 0.8$ decibels.

The plausible range for the width (w) was roughly 310 kilometers give or take a factor of 1.3. A factor of 1.3 is ±1.1 decibels:

$$10 \log_{10} 1.3 \approx 1.1. \tag{7.17}$$

Therefore, $\sigma_w \approx 1.1$ decibels.

The Pythagorean sum of σ_h and σ_w is approximately 1.4 decibels:

$$\sqrt{0.8^2 + 1.1^2} \approx 1.4. \tag{7.18}$$

As a factor, 1.4 decibels is, coincidentally, approximately a factor of 1.4:

$$10^{1.4/10} \approx 1.4. \tag{7.19}$$

Because the midpoint of the plausible range is 250 000 kilometers, the UK land area should be 250 000 square kilometers give or take a factor of 1.4. Retaining a bit more accuracy, it is a factor of 1.37.

$$\underbrace{180\,000}_{/1.37} \ldots \underbrace{250\,000}_{\text{midpoint}} \ldots \underbrace{340\,000}_{\times 1.37}\ \text{km}^2. \tag{7.20}$$

As a probability bar, the range is

compare-pdfs.sh

```
#! /bin/bash

# Usage: $0 file1.pdf file2.pdf
#
#   compares file1.pdf and file2.pdf by rendering each page and using
#   the 'compare' ImageMagick utility
#
# Copyright 2007-2014 Sanjoy Mahajan.  Licensed under the GNU GPL vers
3
# or (at your option) any later version.
#
# HISTORY
#   2014-06-22: Use mudraw instead of pdftoppm.  GPL v3+
#   2009-09-30: Fix capture of dB output; don't use a viewer; use pdf
#   2007-01-15: First version
#
```

```
if [ -z "$DPI" ]; then
    DPI=72
fi
ext=png

if [ -z "$1" -o -z "$2" ]; then
  echo "Usage: $0 file1.pdf file2.pdf"
  exit 3
fi

# generate the many page images in a temporary directory
d=`mktemp -d`
mkdir -p $d/a $d/b
mudraw -r $DPI -g -o $d/a/%03d.$ext $1 &
mudraw -r $DPI -g -o $d/b/%03d.$ext $2
wait

# find the union of the page numbers (in case one pdf has more pages)
```

```
pages=`ls $d/{a,b}/*.$ext | sed "s%.*/\([0-9][0-9]*\).$ext%\1%" | sort
-un`
# compare each page
for p in $pages ; do
  if ! [ -e "$d/a/$p.$ext" ] ; then
    echo "$p: missing from $1"
    continue
  fi
  if ! [ -e "$d/b/$p.$ext" ] ; then
    echo "$p: missing from $2"
    continue
  fi
  echo -n "$d/diff-$p.$ext $p "
  compare -metric mae $d/{a,b}/$p.$ext $d/diff-$p.$ext 2>&1
done
```

As an example, I often prepare a cup of tea but forget to drink it while it is hot. Like a discharging capacitor, the tea slowly cools toward room temperature and becomes undrinkable. Heat flows out through the mug. Its walls provide a thermal resistance; by analogy to an RC circuit, let's denote the thermal resistance R_t. The heat is stored in the water and mug, which form a heat reservoir. This reservoir, of heat rather than of charge, provides the thermal capacitance, which we denote C_t. (Thus, the mug participates in the thermal resistance and capacitance.) Resistance and capacitance are transferable ideas.

The product R_tC_t is, by analogy to the RC circuit, the thermal time constant τ. To estimate τ with a home experiment (the method we used in Section 1.7), heat up a mug of tea; as it cools, sketch the temperature gap between the tea and room temperature. In my extensive experience of tea neglect, an enjoyably hot cup of tea becomes lukewarm in half an hour. To quantify these temperatures, enjoyably warm may be 130 °F (≈ 55 °C), room temperature is 70 °F (≈ 20 °C), and lukewarm may be 85 °F (≈ 30 °C).

Based on the preceding data, what is the approximate thermal time constant of the mug of tea?

In one thermal time constant, the temperature gap falls by a factor of e (just as the voltage gap falls by a factor of e in one electrical time constant). For my mug of tea, the temperature gap between the tea and the room started at 60 °F:

$$\underbrace{\text{enjoyably warm}}_{130\,^\circ\text{F}} - \underbrace{\text{room temperature}}_{70\,^\circ\text{F}} = 60\,^\circ\text{F}. \tag{2.57}$$

In the half hour while the tea cooled in the microwave, the temperature gap fell to 15 °F:

The recipe for using evidence to update probabilities is Bayes' theorem:

$$\Pr(H|E) \propto \Pr(H) \times \Pr(E|H). \qquad (7.3)$$

The new factor, the probability $\Pr(E|H)$—the probability of the evidence given the hypothesis—is called the likelihood. It measures how well the candidate theory (the hypothesis) explains the evidence. Bayes' theorem then says that

$$\underbrace{\text{posterior probability}}_{\Pr(H|E)} \propto \underbrace{\text{prior probability}}_{\Pr(H)} \times \underbrace{\text{explanatory power}}_{\Pr(E|H)}. \qquad (7.4)$$

(The constant of proportionality is chosen so that the posterior probabilities for all the competing hypotheses add to 1.) Both probabilities on the right are necessary. Without the likelihood, we could not change our probabilities. Without the prior probability, we would always prefer the hypothesis with the maximum likelihood, no matter how contrived or post hoc.

In a frequent use of Bayes' theorem, there are only two hypotheses, H and its negation $\overline{H}$. In this problem, $\overline{H}$ is the statement that my guess is wrong. With only two hypotheses, a compact form of Bayes' theorem uses odds instead of probabilities, thereby avoiding the constant of proportionality:

$$\underbrace{\text{posterior odds}}_{O(H|E)} = \underbrace{\text{prior odds}}_{O(H)} \times \frac{\Pr(E|H)}{\Pr(E|\overline{H})}. \qquad (7.5)$$

The odds O are related to the probability p by $O = p/(1-p)$. For example, a probability of $p = 2/3$ corresponds to an odds of 2—often written as 2:1 and read as "2-to-1 odds."

Problem 7.1 Converting probabilities to odds
Convert the following probabilities to odds: (a) 0.01, (b) 0.5, (c) 0.75, and (d) 0.9.

Problem 7.2 Converting odds to probabilities
Convert the following odds to probabilities: (a) 3, (b) 1/3, (c) 1:9, and (d) 4-to-1.

The ratio $\Pr(E|H)/\Pr(E|\overline{H})$ is called the likelihood ratio. Its numerator measures how well the hypothesis H explains the evidence E; its denominator measures how well the contrary hypothesis $\overline{H}$ explains the same evidence. So their ratio measures the relative explanatory power of the two hypotheses. Bayes' theorem, in the odds form, is simple:

$$\text{updated odds} = \text{initial odds} \times \text{relative explanatory power}. \qquad (7.6)$$

Let's use Bayes' theorem to judge my phone-number guess. Before the experiment, I was not too sure of the phone number; $\Pr(H)$ is perhaps 1/2, making $\mathrm{O}(H) = 1$. In the likelihood ratio, the numerator $\Pr(E|H)$ is the probability of getting a busy signal assuming ("given") that my guess is correct. Because I would be dialing my own phone using my phone, I would definitely get a busy signal. Thus, $\Pr(E|H) = 1$: The hypothesis of a correct guess (H) explains the data as well as possible.

The trickier estimate is the denominator $\Pr(E|\overline{H})$: the probability of getting a busy signal assuming that my guess is incorrect. I'll assume that my guess is still a valid phone number (I nowadays rarely get the recorded message saying that I have dialed an invalid number). Then I would be dialing a random person's phone. Thus, $\Pr(E|\overline{H})$ is the probability that a random valid phone is busy. It is probably similar to the fraction of the day that my own phone is busy. In my household, the phone is in use for 0.5 hours in a 24-hour day, and the busy fraction could be 0.5/24.

However, that estimate uses an overly long time, 24 hours, for the denominator. If I do the experiment at 3 am and my guess is wrong, I would wake up an innocent bystander. Furthermore, I am not often on the phone at 3 am. A more reasonable denominator is 10 hours (9 am to 7 pm), making the busy fraction and the likelihood $\Pr(E|\overline{H})$ roughly 0.05. An incorrect guess ($\overline{H}$) is a lousy explanation for the data.

The relative explanatory power of H and $\overline{H}$, which is measured by the likelihood ratio, is roughly 20:

$$\frac{\Pr(E|H)}{\Pr(E|\overline{H})} \sim \frac{1}{0.05} = 20. \qquad (7.7)$$

Because the prior odds were 1 to 1, the updated, posterior odds are 20 to 1:

$$\underbrace{\text{posterior odds}}_{\mathrm{O}(H|E)\sim 20} = \underbrace{\text{prior odds}}_{\mathrm{O}(H)\sim 1} \times \underbrace{\text{likelihood ratio}}_{\Pr(E|H)/\Pr(E|\overline{H})\sim 20} \sim 20. \qquad (7.8)$$

My guess has become very likely—and it turned out to be correct.

Problem 7.8 PKU testing

In most American states and many countries, newborn babies are tested for the metabolic defect phenylketonuria (PKU). The prior odds of having PKU are about 1 to 10 000. The test gives a false-positive result 0.25 percent of the time; it gives a false-negative result 0.8 percent of the time. What are $\Pr(\text{PKU}\,|\,\text{positive test})$ and $\Pr(\text{PKU}\,|\,\text{negative test})$?

Check for unknown references (in pdf file)

```
$ $make preflight
...
pdftotext -layout book-print - | egrep '][A-Za-z]|\[\]|\?\?|\[,|, ]'
...
```

(among 30 or so other preflight checks: on the source files and on the resulting PDF file)

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.edu.

Typeset by the author in 10.5/13.3 Palatino and Computer Modern Sans using ConTEXt and LuaTEX.

Library of Congress Cataloging-in-Publication Data

Mahajan, Sanjoy, 1969- author.
The art of insight in science and engineering : mastering complexity / Sanjoy Mahajan.
pages cm
Includes bibliographical references and index.
ISBN 978-0-262-52654-8 (pbk. : alk. paper) 1. Statistical physics. 2. Estimation theory. 3. Hypothesis. 4. Problem solving. I. Title.
QC174.85.E88M34 2014
501′.9-dc23

2014003652

Printed and bound in the United States of America

10 9 8 7 6 5 4 3 2 1

The benefit is a free license.

Acknowledgments

In addition to the dedication, I would like to thank the following people and organizations for their generosity.

For encouragement, forbearance, and motivation: my family—Juliet Jacobsen, Else Mahajan, and Sabine Mahajan.

For a sweeping review of the manuscript and improvements to every page: Tadashi Tokieda and David MacKay. Any remaining mistakes were contributed by me subsequently!

For advice on the process of writing: Larry Cohen, Hillary Rettig, Mary Carroll Moore, and Kenneth Atchity (author of *A Writer's Time* [1]).

For editorial guidance over many years: Robert Prior.

For valuable suggestions and discussions: Dap Hartmann, Shehu Abdussalam, Matthew Rush, Jason Manuel, Robin Oswald, David Hogg, John Hopfield, Elisabeth Moyer, R. David Middlebrook, Dennis Freeman, Michael Gottlieb, Edwin Taylor, Mark Warner, and many students throughout the years.

For the free software used for typesetting: Hans Hagen, Taco Hoekwater, and the ConTEXt user community (ConTEXt and LuaTEX); Donald Knuth (TEX); Taco Hoekwater and John Hobby (MetaPost); John Bowman, Andy Hammerlindl, and Tom Prince (Asymptote); Matt Mackall (Mercurial); Richard Stallman (Emacs); and the Debian GNU/Linux project.

For the NB document-annotation system: Sacha Zyto and David Karger.

For being a wonderful place for a graduate student to think, explore, and learn: the California Institute of Technology.

For supporting my work in science and mathematics education: the Whitaker Foundation in Biomedical Engineering; the Hertz Foundation; the Gatsby Charitable Foundation; the Master and Fellows of Corpus Christi College, Cambridge; Olin College of Engineering and its Intellectual Vitality program; and the Office of Digital Learning and the Department of Electrical Engineering and Computer Science at MIT.

Thanks to you.

Typesetting a science and engineering textbook

Sanjoy Mahajan

MIT & Olin College

mit.edu/sanjoy/www/
sanjoy@mit.edu

ConTeXt meeting, Bassenge, Belgium, 10 September 2014