

Numbers in medicine: History and Perspectives

Short Course on Innovative Medical Devices

Tsinghua University, Beijing

Andy Adler

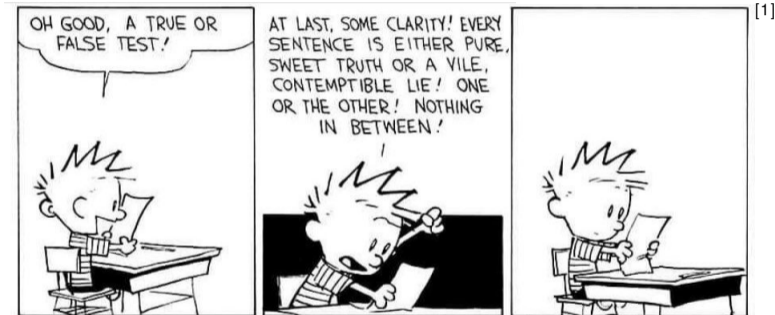
Carleton University, Ottawa, Canada

2026-04-23

Digital numbers: true *or* false

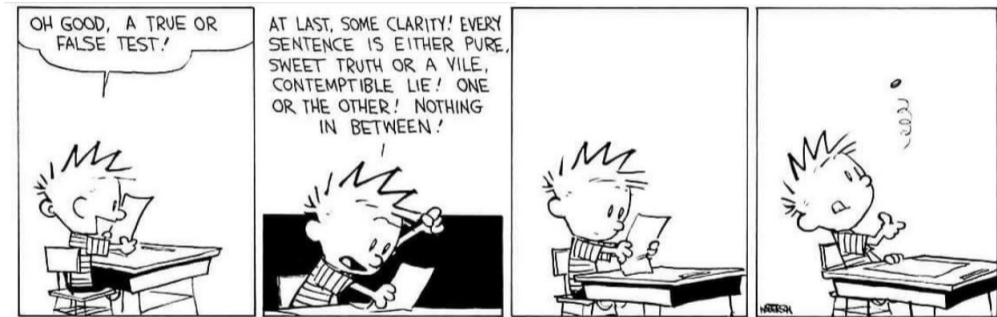
⁰Watterson, "The Days Are Just Packed", 1993

Digital numbers: true or false



¹ Watterson, "The Days Are Just Packed", 1993

Digital numbers: true or false



[1]

¹Watterson, "The Days Are Just Packed", 1993

Knowledge

Knowledge

≡ true belief; in accordance with fact or truth

Knowledge

Knowledge

≡ true belief; in accordance with fact or truth

... distinct from a lucky guess, because of

Justification

Knowledge

Knowledge

≡ true belief; in accordance with fact or truth

... distinct from a lucky guess, because of

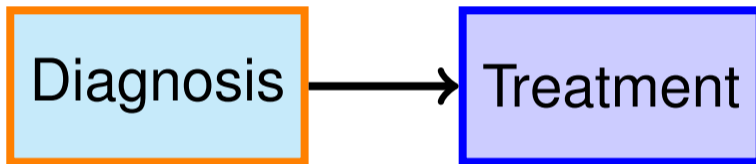
Justification

≡ reasons for a rationally admissible belief

... makes belief *knowledge* instead of opinion.

Medicine: Diagnosis and Treatment

Why do we want **knowledge** . . . to **treat**.



The *only* reason to **measure**=diagnose is if we plan to make a **decision**=treat

Plato's cave: how we gain knowledge

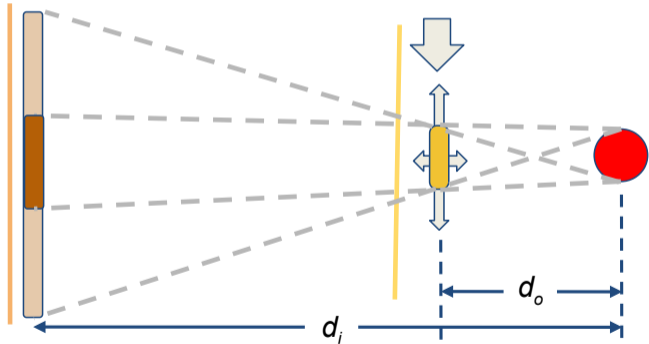


²Plato's cave, [wikimedia.org](https://www.wikimedia.org), inspired by C Groetsch "Linear Inverse Problems" in O Schertzer *Handbook of Mathematical Methods in Imaging*

Sensitivity & Noise in the cave



[3]



³Adler (2026) "Medical Devices and Data: ...", IOP Press

Medical scenario

- We're hiking in the wilderness; a friend falls and gets injured.
- It's late; medical care is three hours away.
- We are prepared to walk out in the dark, but that's risky and inconvenient.
- We would prefer to camp tonight and leave in the morning.

<i>Diagnosis</i>	<i>Treatment</i>	<i>If correct</i>	<i>If wrong</i>
Serious	Hike out at night	Get medical care	Risk of walking at night
Not Serious	Leave in morning	Not risk walk at night	Serious case doesn't get care

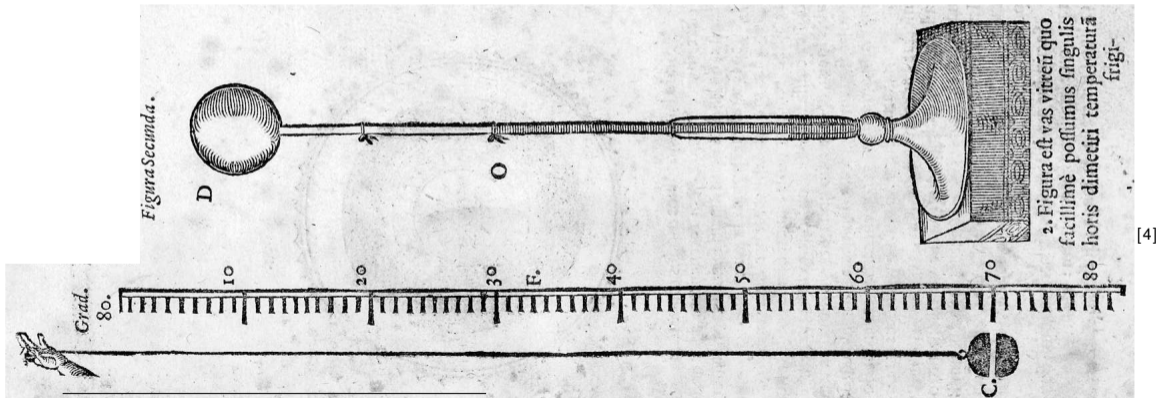
Diagnosis: Does HR stay high?

→ Need to record measurements over time

Pulsilogium

Santorio Santori (1561–1636)

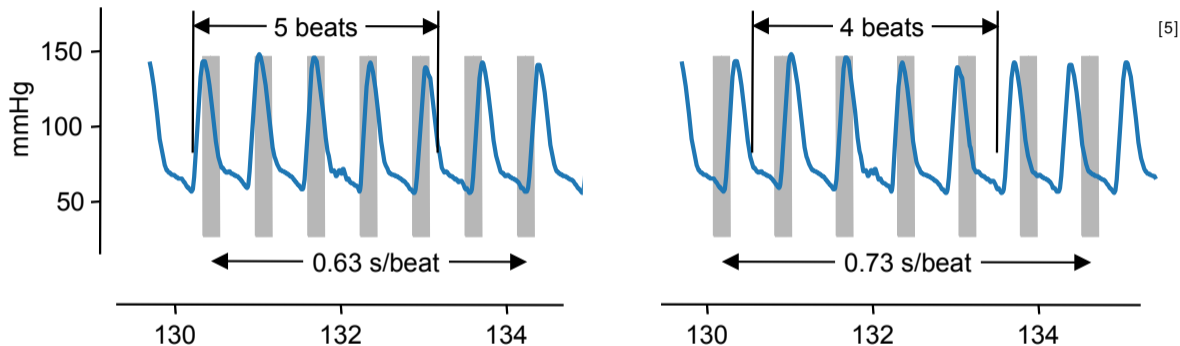
- invented the pulsilogium – the first medical instrument to give a number.
- worked with Galileo and invented several medical instruments.



[4]

⁴Santorius (1626) "Commentaria in primam Fen primi libri Canonis Avicennae" (Wellcome collection)

Heart rate instruments



- Count pulses with a watch, *or* align synchronize the pendulum length
- Interestingly, pulsilogium is more accurate. Why?

⁵Adler (2026) "Medical Devices and Data: ...", IOP Press

For and against measurement

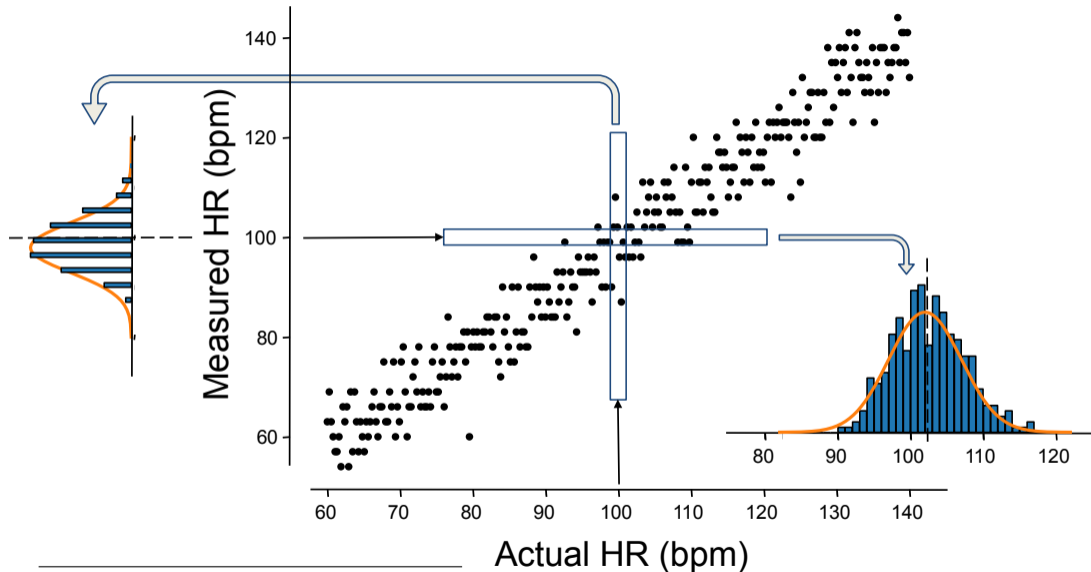
Pro

- *Comparisons*: measurements over time
- *Standardization*: measurements between patients
- *Cost*: measurements by technician
- *Sophisticated analysis*: . . . see more later

Con

- “a physician cannot prescribe by letter; he must feel the pulse.” (Seneca)

But numbers can have errors ...



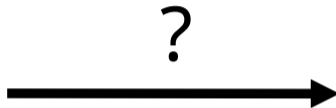
[6]

Thinking about errors

- *True* → *Measure*:
 - start from correct value; ask what could be measured.
 - (vertical distribution) useful for the accuracy, resolution, and precision.
- *Measure* → *True*:
 - start with measured value; ask what true values could have led to it
 - (horizontal distribution). resolution and precision no longer distinct features
 - for hypothesis testing → Bayesian analysis

Interpretation of numbers

Long
Sweet
Yellow



Bayes' rule

Thomas Bayes (c.1701 – 1761) was an English statistician, philosopher and Presbyterian minister. [Some] speculate he was motivated to rebut David Hume's argument against believing in miracles

Diagnosis problem:

Have x measurements

Want $p(C_i|x)$ Probability of class C_i given x

Bayes' rule: formula

Prior probability: chance of measuring features (long, yellow, sweet) if it's a banana

$$p(C_i|x) = \frac{p(x|C_i)p(C_i)}{p(x)} \propto p(x|C_i)p(C_i)$$

Posterior probability: chance of being a banana, given our measured features (long, yellow, sweet)

[what we want to calculate]

Choose largest, so only need \propto

Class probability (prior): probability of it being a banana

$$p(x) = \sum_i p(x|C_i)p(C_i)$$

Bayes' rule: example

Collect data from the world

Type	Long	Not Long	Sweet	Not Sweet	Yellow	Not Yellow	Total
Banana	400	100	350	150	450	50	500
Orange	0	300	150	150	300	0	300
Other	100	100	150	50	50	150	200
Total	500	500	650	350	800	200	1000



Naïve Bayes

A (composite) measurement x is a set of individual measurements (x_1, x_2, x_3) .

Naïve Bayes assumes measurements are independent.

$$P(C_i|x) \propto p(x|C_i)p(C_i) = p(x_1|C_i)p(x_2|C_i)p(x_3|C_i)p(C_i)$$

Bayes' rule: example

A fruit is *Long*, *Sweet*, and *Yellow*. What is it?

	x_1 =Long	x_2 =Sweet	x_3 =Yellow	$p(C_i)$		$p(C_i x)$
C_{Banana}	0.8	0.7	0.9	0.5		0.252
C_{Orange}	0	0.5	0	0.3	\propto	0
C_{Other}	0.5	0.75	0.25	0.2		0.019



Bayes' rule: example #2

A fruit is *Not Long*, *Sweet*, and *Not Yellow*. What is it?

	x_1 =Not Long	x_2 =Sweet	x_3 =Not Yellow	$p(C_i)$		$p(C_i x)$
C_{Banana}	1-0.8	0.7	1-0.9	0.5		0.007
C_{Orange}	1-0	0.5	1-0	0.3	\propto	0.150
C_{Other}	1-0.5	0.75	1-0.25	0.2		0.053



Numbers?

Medical measurements are useful because we can

- *Comparisons*: value over time
- *Standardization*:
- *Cost*: measurements by technician
- *Sophisticated analysis*

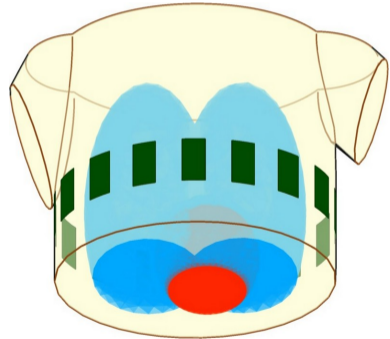
Numbers give us a ways to combine information.

This is machine learning.

Medical Imaging: Electrical Impedance Tomography

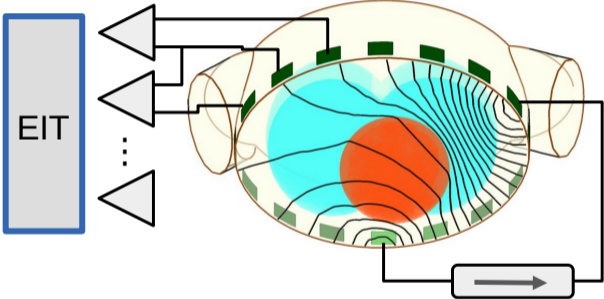


[7]



⁷Heinrich *et al* (2006) **Body and head position effects on regional lung ventilation in infants ...** *Int Care Med* 32:1392–1398

EIT: apply current and measure voltages



EIT: Physiology

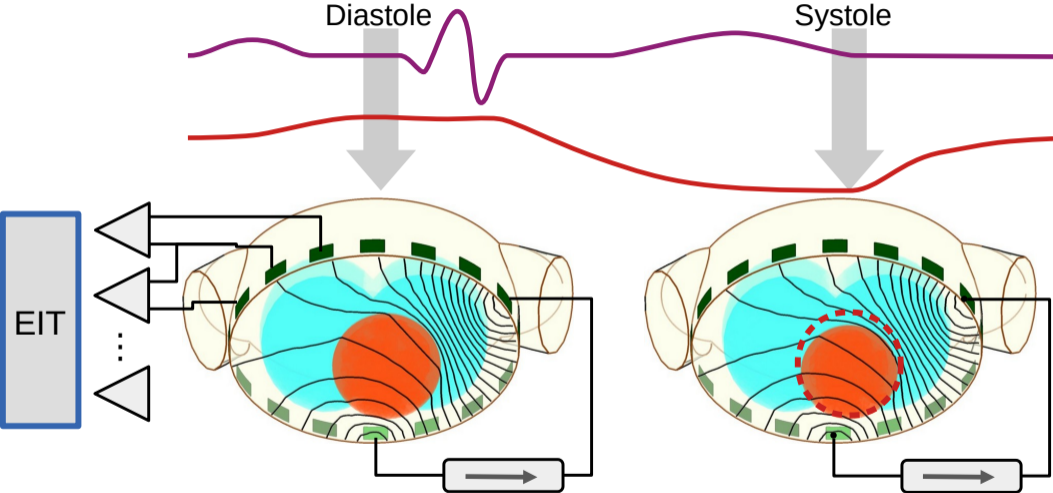


Image Reconstruction

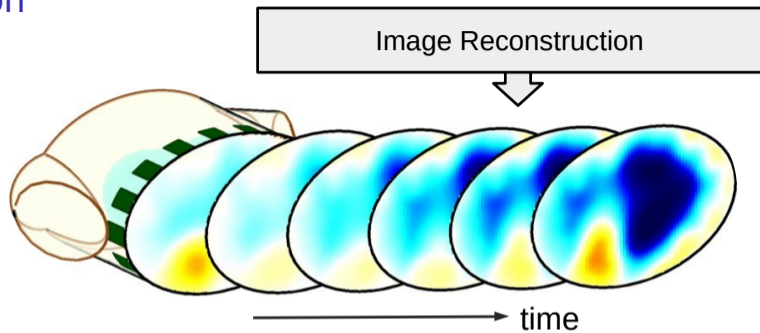
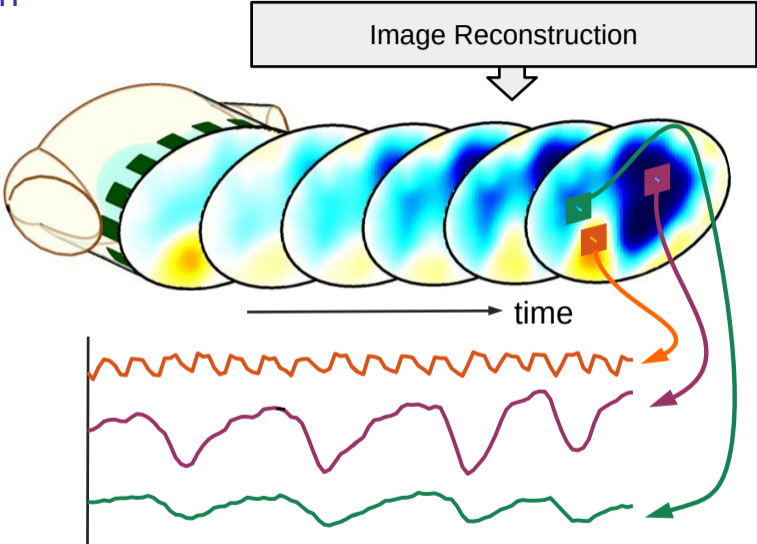
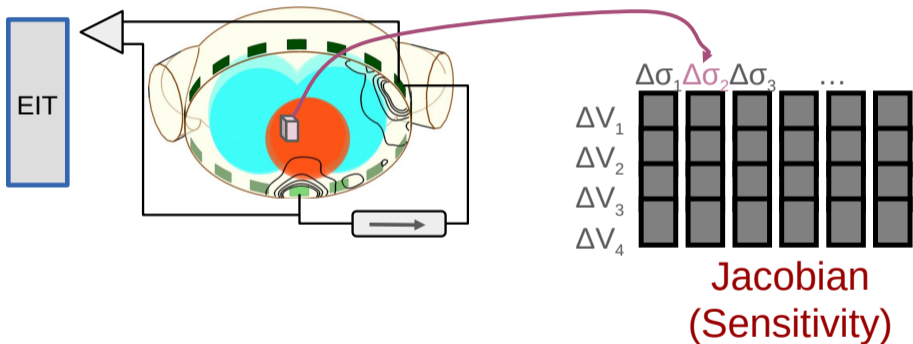


Image Reconstruction

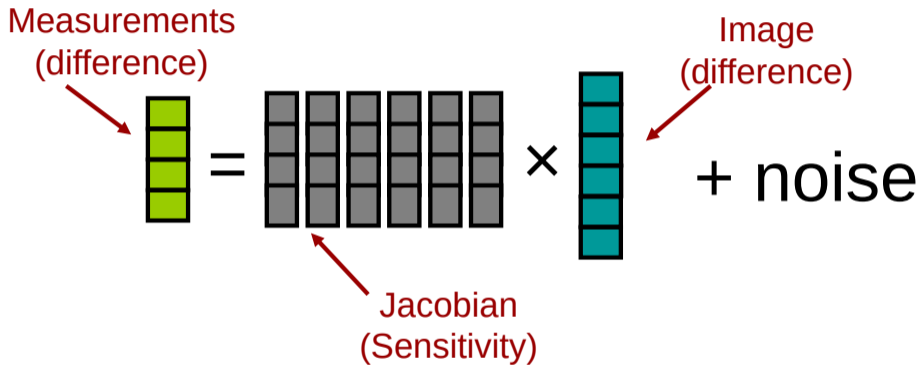


Bayesian analysis

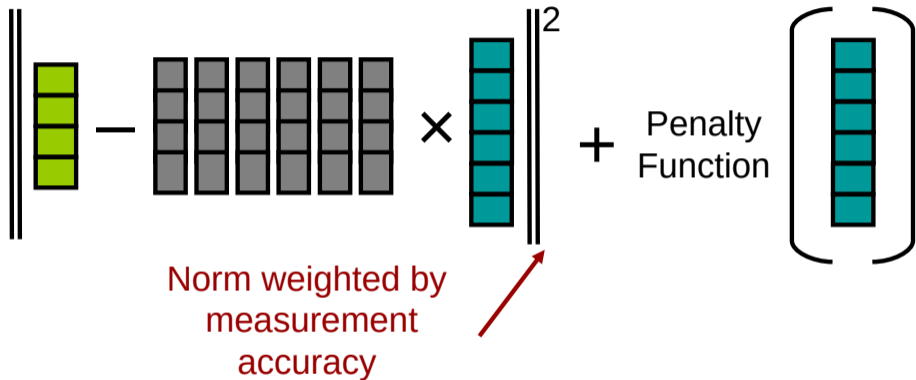
Start with a model to understand sensitivity of measurements (x) to the image



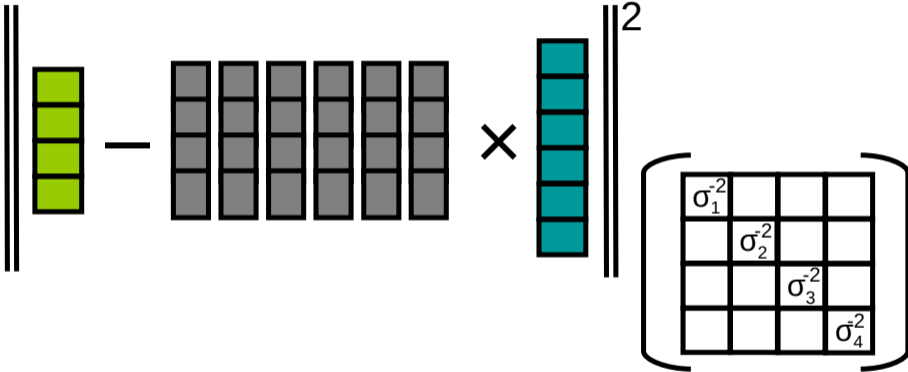
Forward problem: parameters \rightarrow measurements



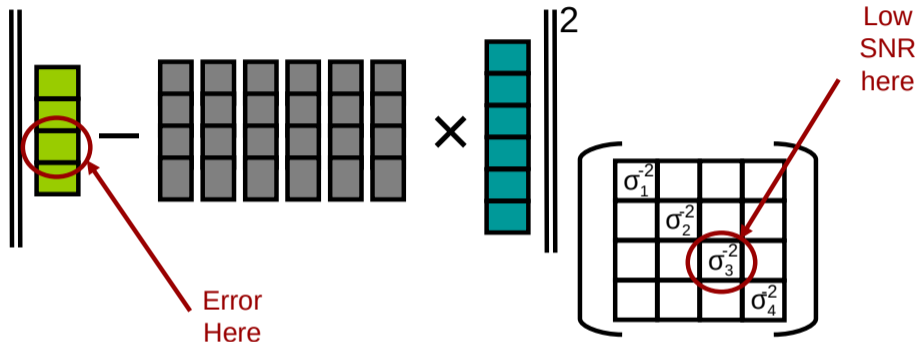
Inverse problem is the image which minimizes error



The *data fidelity* term (= Naïve Bayes)



How to deal with data errors



Translation: if we *know* about our measurements, we can make better images

Summary so far

- Numbers let us do **powerful** processing

Next: what can go wrong?

Image stabilization

≡ techniques to reduce blurring associated with the motion of a camera

- Optical image stabilization
- Digital image stabilization
 - take multiple exposures
 - reject blurred ones
 - transform and average the remaining ones

What went wrong?



[8]

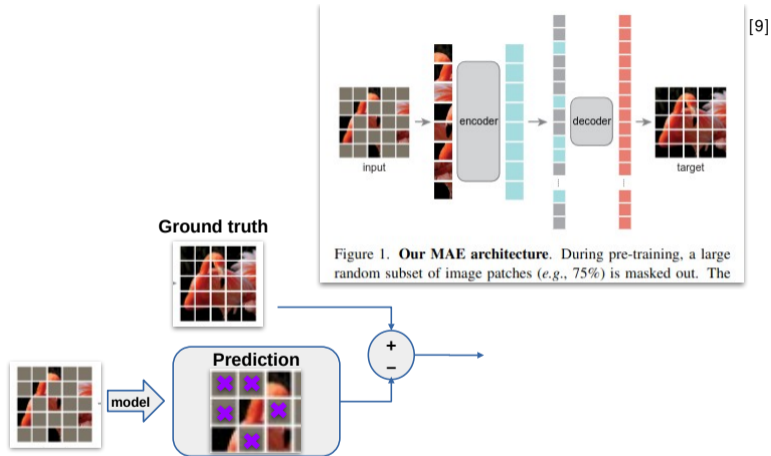
Panning the camera to follow a moving cyclist

⁸[expertphotography.com/image-stabilization.](https://expertphotography.com/image-stabilization/)

Problems

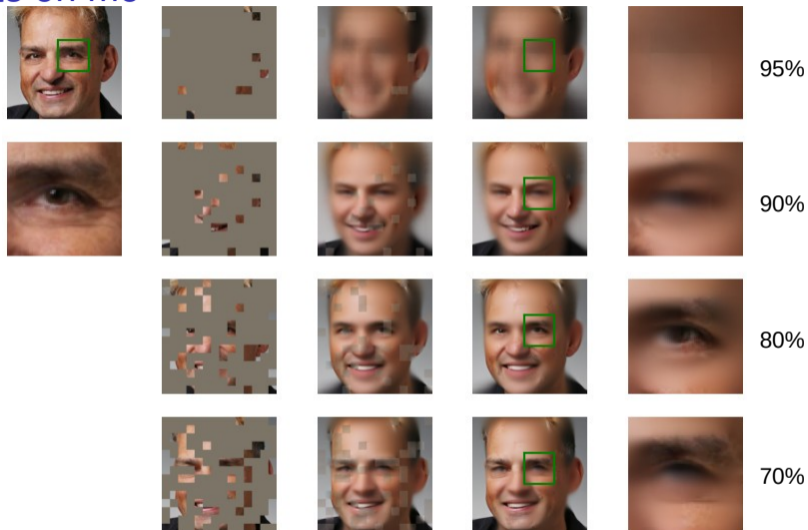
- Algorithm model is wrong:
 - Here algorithm assumes shaking motion; but this motion is continuous
-
-
-

Autoencoder: Facebook AI Research



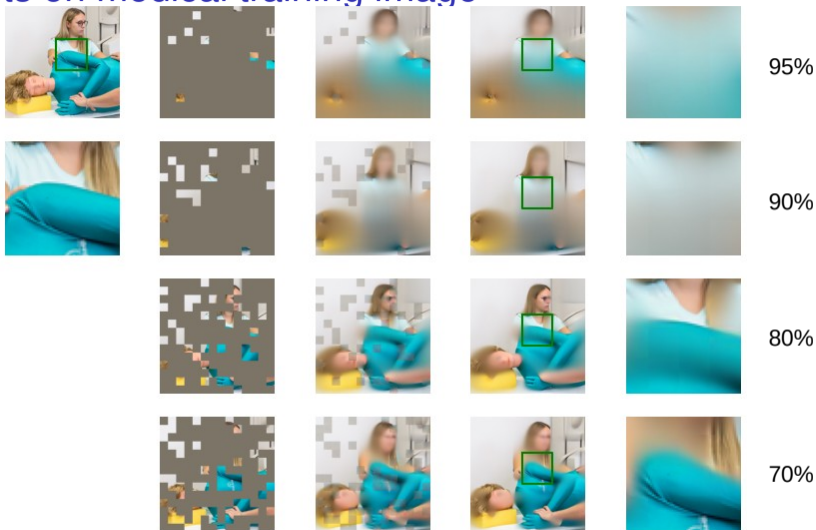
⁹He et al (2021) Masked Autoencoders

Autoencoder: Tests on me



Seems to work well?

Autoencoder: Tests on medical training image



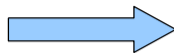
Top image: why is the woman white?

Problems

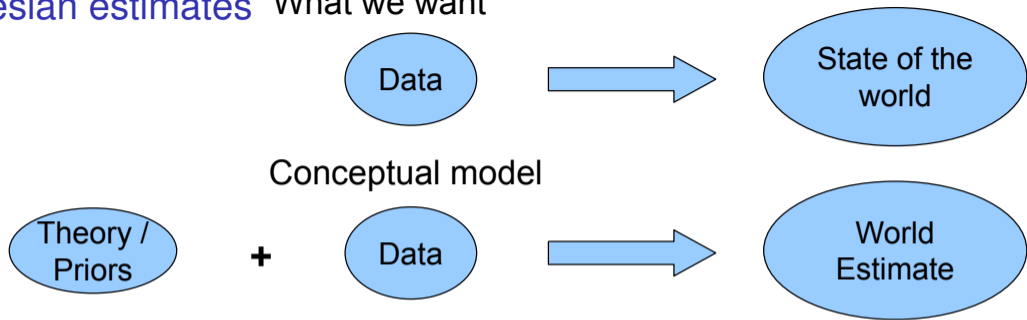
- Algorithm model is wrong:
- Biased training data:
 - Facebook's main users are white women. It assumes this is true in general.
-
-

Bayesian estimates

What we want



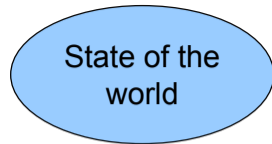
Bayesian estimates What we want



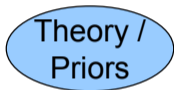
Bayesian estimates

... as implemented

What we want



Conceptual model $p(C|x) \propto p(x|C)p(x)$

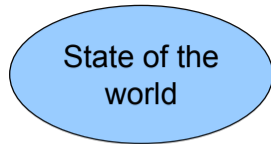


+

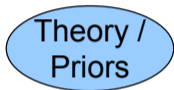


Bayesian estimates ... as implemented

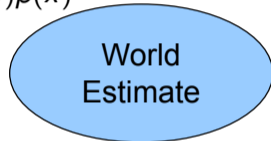
What we want



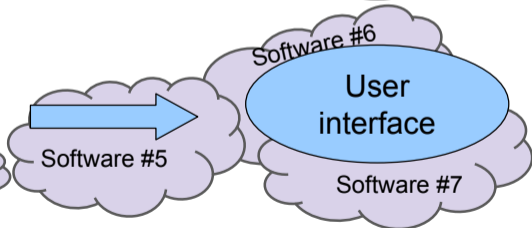
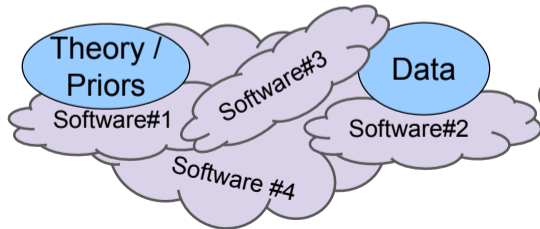
Conceptual model $p(C|x) \propto p(x|C)p(x)$



+



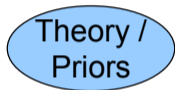
What we have



Bayesian estimates What we want



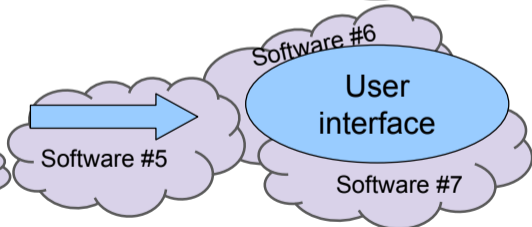
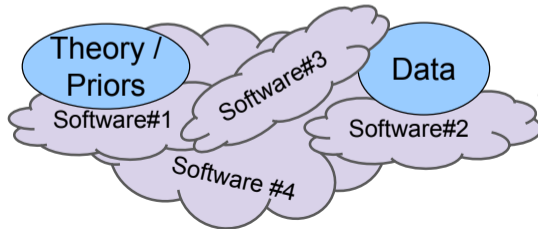
Conceptual model $p(C|x) \propto p(x|C)p(x)$



+



What we have



$p(C|x) \propto p(x|C, \text{Software})p(x|\text{Software})$

Software from a recent project



Python3 Packages:

```
localreg==0.5.0
matplotlib==3.5.1
neurokit2==0.2.4
numpy==1.21.5
pandas==2.0.1
Pillow==9.0.1
Pillow==9.5.0
pyparsing==2.4.7
scipy==1.8.0
skimage==0.0
wfdb==4.1.1
```

libpython3.10-stdlib:

```
libpython3.10-minimal
libbz2-1.0
libc6 (>= 2.34)
libcrypt1 (>= 1:4.1.0)
libdb5.3
libffi8 (>= 3.4)
liblzma5 (>= 5.1.1alpha+20120614)
libmpdec3
libncursesw6 (>= 6.1)
libnsl2 (>= 1.0)
libreadline8 (>= 7.0~beta)
libsqlite3-0 (>= 3.14.0)
libtinfo6 (>= 6)
libtirpc3 (>= 1.0.2)
libuuid1 (>= 2.20.1)
```

Numpy Packages: python3-numpy

```
python3-pkg-resources
python3 (<< 3.11)
python3 (>= 3.10~)
libblas3 -
libatlas3-base -
libblas3 -
libblis3-openmp -
libblis3-pthread -
libblis3-serial -
libopenblas0-openmp
libopenblas0-pthread
libopenblas0-serial
libc6 (>= 2.35)
liblapack3: <liblapack.so.3>
libatlas3-base
libc6 (>= 2.29)
libgcc-s1 (>= 4.0)
libgfortran5 (>= 8)
liblapack3
libopenblas0-openmp
libopenblas0-pthread
libopenblas0-serial
libc6 (>= 2.29)
libgcc-s1 (>= 4.0)
libgfortran5 (>= 8)
```

ATLAS
(Automatically
Tuned Linear
Algebra Software)

BLAS (Basic Linear
Algebra
Subprograms)

BLAS: Basic Linear Algebra Subprograms

Basic Linear Algebra Subprograms

(**BLAS**) is a [specification](#) that prescribes a set of low-level routines for performing common [linear algebra](#) operations such as [vector](#) addition, [scalar multiplication](#), [dot products](#), linear combinations, and [matrix multiplication](#).

BLAS implementations are often optimized for speed on a particular machine,

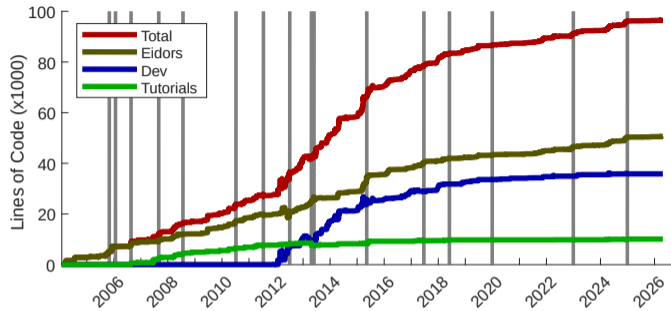
It originated as a Fortran library in 1979 and its interface was standardized by the BLAS Technical (BLAST) Forum

file	blas.tgz	Fortran77 reference implementation
prec		single, double, complex, doublecomplex
lang		fortran
#		DOCUMENTATION AND TEST SUITES
file	blasqr.ps	quick reference guide for the BLAS.
file	blas3-paper.ps	details on the Level 3 BLAS
file	blas2test.f	original test driver for the real level two blas
file	sblat1.f	Level 1 BLAS Test Suite (prec. single)
file	dblats1.f	Level 1 BLAS Test Suite (prec. double)
...

Tests

¹⁰ [wikipedia.org](#) and [netlib.org](#)

My open-source experience: EIDORS.org



- Contributions from community
- Managing growth
- Testing!!
- Toolchain rot (sourceforge, mailing lists, matlab, perl, html)

EIDORS: contributions and testing

- Contributions are great!
- Contributors will solve their own problem, but often won't test all cases
- Who checks?

On Thu, Feb 16, 2023 XXXX wrote:

> Dear Prof. Adler,

>


> I am a phd-student, who is dealing with EIT.

> Recently I published an article ...

>

> I would like to ask you whether you think it is a good idea to modify the code of my paper such that it can be integrated directly into EIDORS?

...



```
function data = fwd_solve_2p5d_1st_order(fwd_model, img)
% FWD_SOLVE_2P5D_1ST_ORDER: data= fwd_solve_2p5d_1st_order( img)

% -----

function do_unit_test
% -----

unit_test_cmp('h2a 2p5d values TEST', uvh, [ 0.0060; 0.0080; 0.0110; 0.0160; 0.0320 ], 1/scale);
tol = norm(vh.meas)*0.025; % 2.5% error tolerance
unit_test_cmp('2D (h2a) vs analytic TEST', norm(vh.meas - v2.meas), 0, -tol);
unit_test_cmp('2.5D (h2a + 2.5D, k=0) vs 2D TEST', norm(v2.meas - v2p5k0.meas), 0, tol);
unit_test_cmp('2.5D (h2a + 2.5D trapz, 2*tol) vs analytic TEST', norm(vh.meas - v2p5t.meas), 0, 2*tol);
unit_test_cmp('2.5D (h2a + 2.5D quadv**) vs analytic TEST', norm(vh.meas - v2p5q.meas), 0, tol);

plot([vh.meas, v2p5q.meas, v3.meas, v2.meas], ':');
legend('analytic', 'FEM 2.5D', 'FEM 3D', 'FEM 2D', 'Location', 'Best');
% -----
```

Problems

- Algorithm model is wrong:
- Biased training data:
- **Dependent on lots of software:**
 - Most of it is unknown. I'm familiar \approx 10% of code in my packages
-

Reproducibility Project: Psychology

△ a crowdsourced collaboration . . . to repeat 100 published experimental and correlational psychological studies . . . led by the Center for Open Science¹¹

Comments:

- Overall reproducibility is poor
- Let's not criticise psychology . . . they're brave
- Open science is the right way

¹¹en.wikipedia.org/wiki/Reproducibility_Project

Can we implement a formula correctly?

- In APA reporting format, inferential statistics must report: the test statistic, the degrees of freedom for the test, and the p -value.
- values redundant with one another.
- can check consistency by evaluating whether they agree¹²
- p -values for more than 250,000 psychology papers checked; around half contained at least one incorrect p -value!¹³

¹²Frank *et al* *Experimentology: An Open Science Approach to Experimental Psychology Methods*, experimentology.io/

¹³Nuijten *et al*, "The prevalence of statistical reporting errors in psychology (1985–2013)", (2016) *Behavior Research Methods*. 48:1205–1226

How about reproducing complicated results?

- Hardwicke *et al*^{14,15} identified datasets with reusable data (not all complete and comprehensible)
- attempted to reproduce the main statistical results from 60 articles
- Very labour-intensive (5–10 hours of work each).
- Results
 - about 1/3 reproducible without help from the original authors
 - 62% reproduced after – sometimes extensive – correspondence
 - Many remaining appeared to have some irreproducible results

¹⁴Hardwicke *et al* “Data availability, reusability, and analytic reproducibility: evaluating the impact of a mandatory open data policy at the journal *Cognition*” (2018) Royal Soc. open sci 5180448180448

¹⁵Hardwicke *et al* “Analytic reproducibility in articles receiving open data badges at the journal *Psychological Science: an observational study*” (2021) Royal Soc. open sci 8201494201494

Open science: Reproducible science¹⁶



¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different		

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	

¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Open science: Reproducible science¹⁶

		Data	
		Same	Different
Code	Same	Reproducible	Replicable
	Different	Robust	Generalizable

- Data availability: zenodo, dataverse
- Code availability: github, or with data

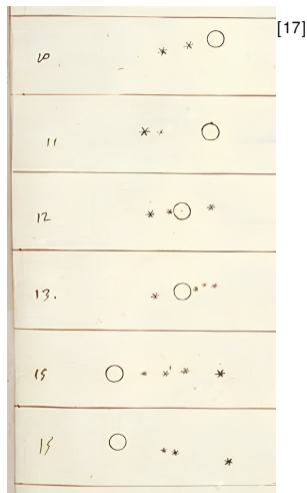
¹⁶Whitaker, "Open Science in Practice Summer School" (2017)

Problems

- Algorithm model is wrong:
- Biased training data:
- Dependent on lots of software:
- **Untested software:**
 - But complicated projects are hard to test!

Why numbers? Back to the Renaissance

- Math's accurate predications made astronomy work
- Galileo debated the philosophers at the University of Padua, who refuted *any "new" planets on what they regarded as logical grounds. Galileo challenged them to come and look in his telescope for themselves, but they declined, saying that they would not look in his telescope, because it showed merely appearances, which were illusions, not to be trusted when they conflicted with logic and reason.*



¹⁷Galileo's sketch of Jupiter's moons, observations Jan 10–15, 1610.

Summary

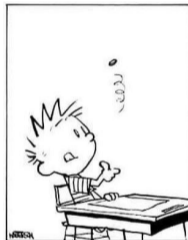
- We are so familiar with numbers in medicine, we don't think about how it wasn't always like this
- Numbers let us do math. This means we can combine numbers on different days and equipment.
- Numbers aren't just facts. They have a psychological effect
- As processing gets more complicated, verification gets *much* harder.

Questions for discussion

- Modern ML means we can do so much more. Are there limits?
- How do we better manage errors?
- Can we help people understand / trust algorithms?
- How do we get algorithms to collaborate with people?

Answer Correct?

How do you know?



Thank you, 谢谢

Abstract: This talk reviews the role of numerical techniques in medical data. Starting from a historical perspective, we look at the earliest medical instruments. We then look at medical images and data and the role of data of numerical techniques in data interpretation. An an example, we review cardiovascular and lung imaging and the author's technology of Electrical Impedance Tomography. Finally, we look at the opportunities and challenges in modern data analysis, including the difficulties in validating complex algorithms.