hazards of verification

Daniel Jackson, MIT

Haifa Verification Conference · October 28, 2008

warnings

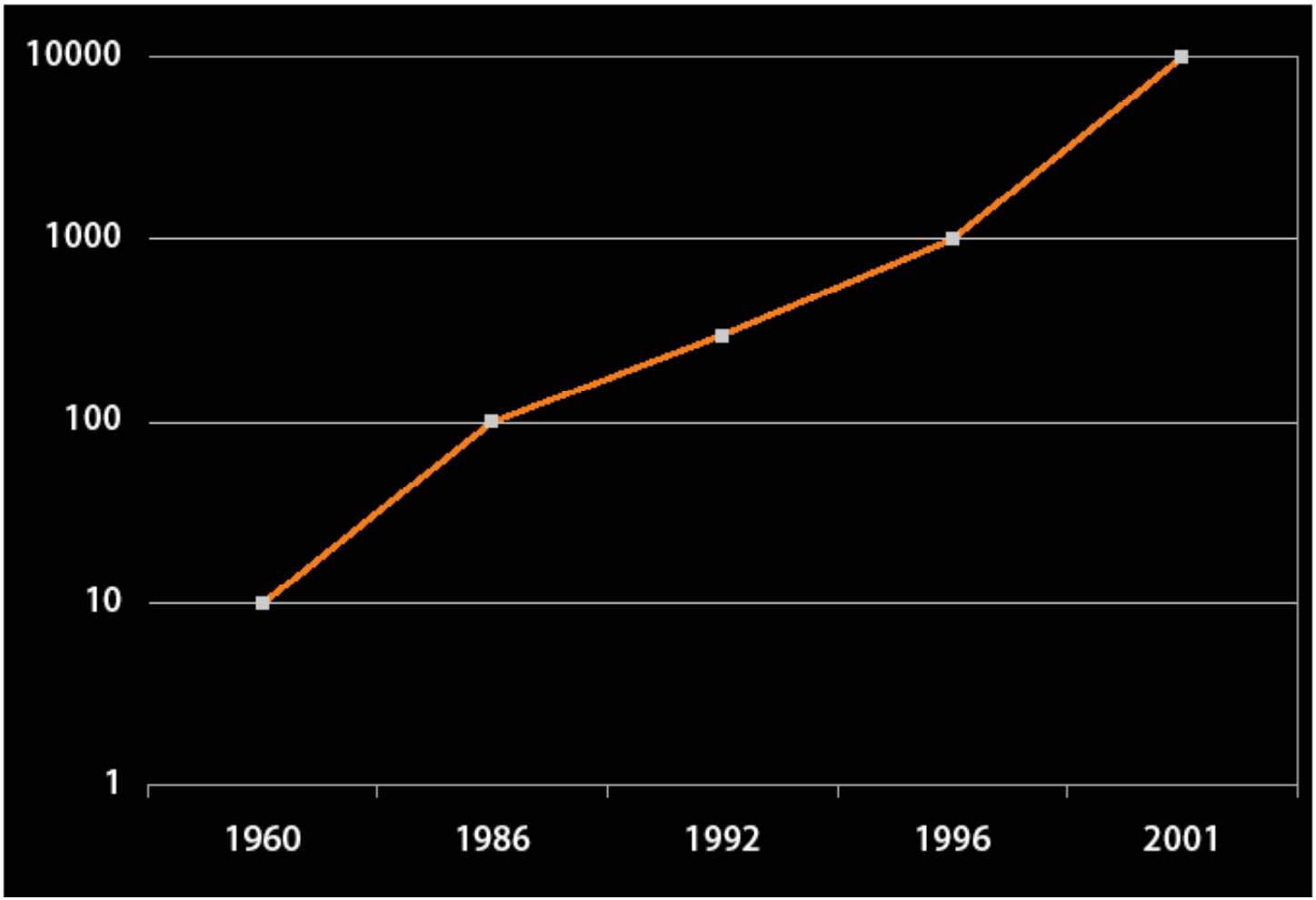
the contents of this talk are

- anecdotal, not analytical
- broad, not focused
- old, not novel

It is insufficiently considered that men more often require to be reminded than informed.
--Samuel Johnson

© Daniel Jackson 2008

how we got here



growth in SAT power (number of variables, data from Sharad Malik)

- one example of why early pessimism about verification was misplaced

© Daniel Jackson 2008

hazards

but will verification made software safe and dependable?

- on the road ahead: much progress, but hazards too

hazards due to

- technical factors
- engineering factors
- social/managerial factors

© Daniel Jackson 2008

technical factors

unsound confirmation

examples

- finite scope & unrolling [KOA, Dennis VSTTE08]
- lack of coverage [CP bug after 8 years, Holzmann]
- abstraction [binary search, Bloch]

```
L:=1; U:=N
loop
     { MustBe(L,U) }
     if L>U then
          P:=0; break
     M := (L+U) div 2
     case
          X[M] < T:   L:=M+1
          X[M] = T:   P:=M; break
          X[M] > T:   U:=M-1
  endloop
```

© Daniel Jackson 2008

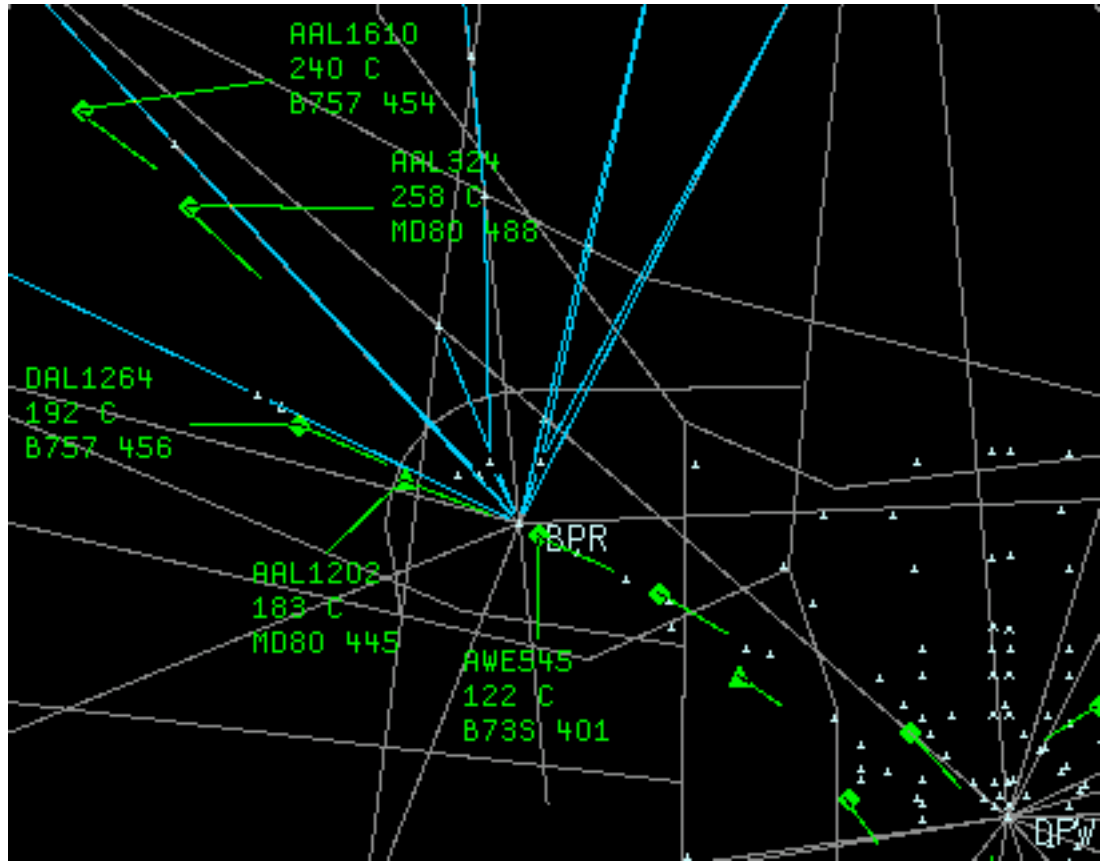
how big a bound?

minimum scope/bitwidth/unrolling to find bugs in voting code

class	method	error	min bound
CandidateListMetadata	init	under	1 / 3 / 1
KiesKring	addDistrict	bug	1 / 3 / 1
VoteSet	addVote(String)	over	1 / 3 / 1
KiesLijst	clear	over	1 / 3 / 3
AuditLog	getCurrentTimeStamp	over	2 / 1 / 1
Candidate	init	under	2 / 3 / 1
CandidateList	addDistrict	under	2 / 3 / 1
CandidateList	addKiesLijst	over	2 / 3 / 1
CandidateList	init	over	2 / 3 / 1
KiesKring	addKiesLijst	bug	2 / 3 / 1
KiesKring	init	under	2 / 3 / 1
KiesKring	make	under	2 / 3 / 1
KiesLijst	addCandidate	over	2 / 3 / 1
KiesLijst	compareTo	bug	2 / 3 / 1
KiesLijst	make	over	2 / 3 / 1
VoteSet	addVote(int)	over	2 / 3 / 1
VoteSet	validateKiesKringNumber	over	2 / 3 / 1
VoteSet	validateRedundantInfo	over	2 / 3 / 1
KiesKring	clear	over	2 / 3 / 3

© Daniel Jackson 2008

unsound counterexamples



examples

- unsound checker finds more bugs [Xie and Aiken 2005]
- most effort on error reporting [Pincus et al, Prefix]

© Daniel Jackson 2008

overconstraint

```
abstract sig MemberEvent extends Event {
   by: Member
   }
   {
   by in before.members
   }

abstract sig MembershipEvent extends MemberEvent {
   }

sig Join extends MembershipEvent {
   }
   {
   by not in before.members
   after.members = before.members + by
   }
```

examples

- declarative models of software (Alloy, Z, VDM, B, etc)
- axioms for code verifiers
- 'unreachable states' in model checking

© Daniel Jackson 2008

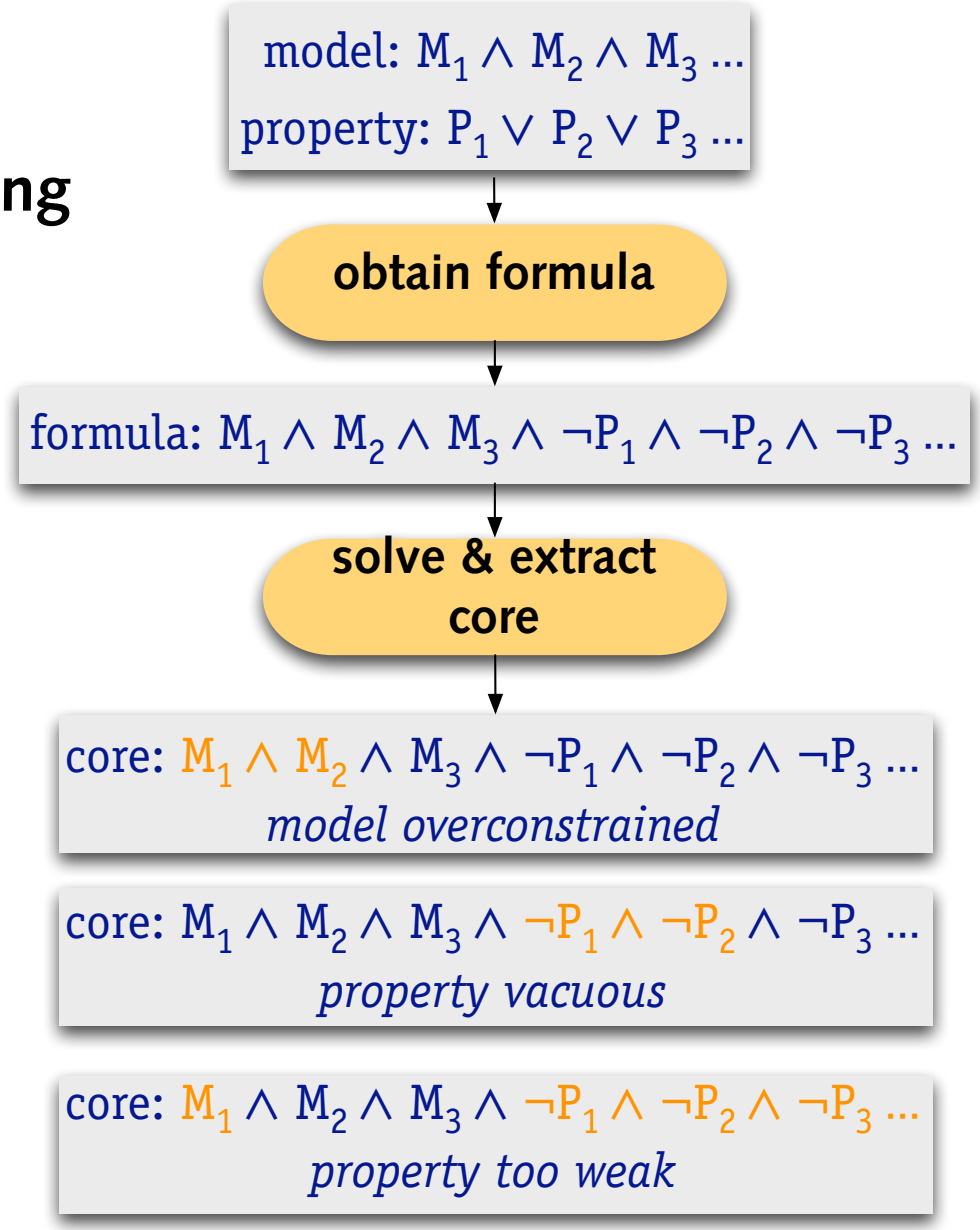
approaches

vacuity and coverage in model checking

- Beer, Ben-David, Eisner, Rodeh
- Chockler, Kupferman, Vardi
- Chechik, Devereux, Gurfinkel

coverage in Alloy

- new algorithm [Torlak, FME08]



© Daniel Jackson 2008

engineering factors

end-to-end

are bugs in code the problem?

- Mackenzie: 3% of software fatalities due to code
- most problems in human/computer interaction

is run-time-error elimination enough?

- 'Sorry no more bugs' -- Greg Nelson, 1980

sad examples

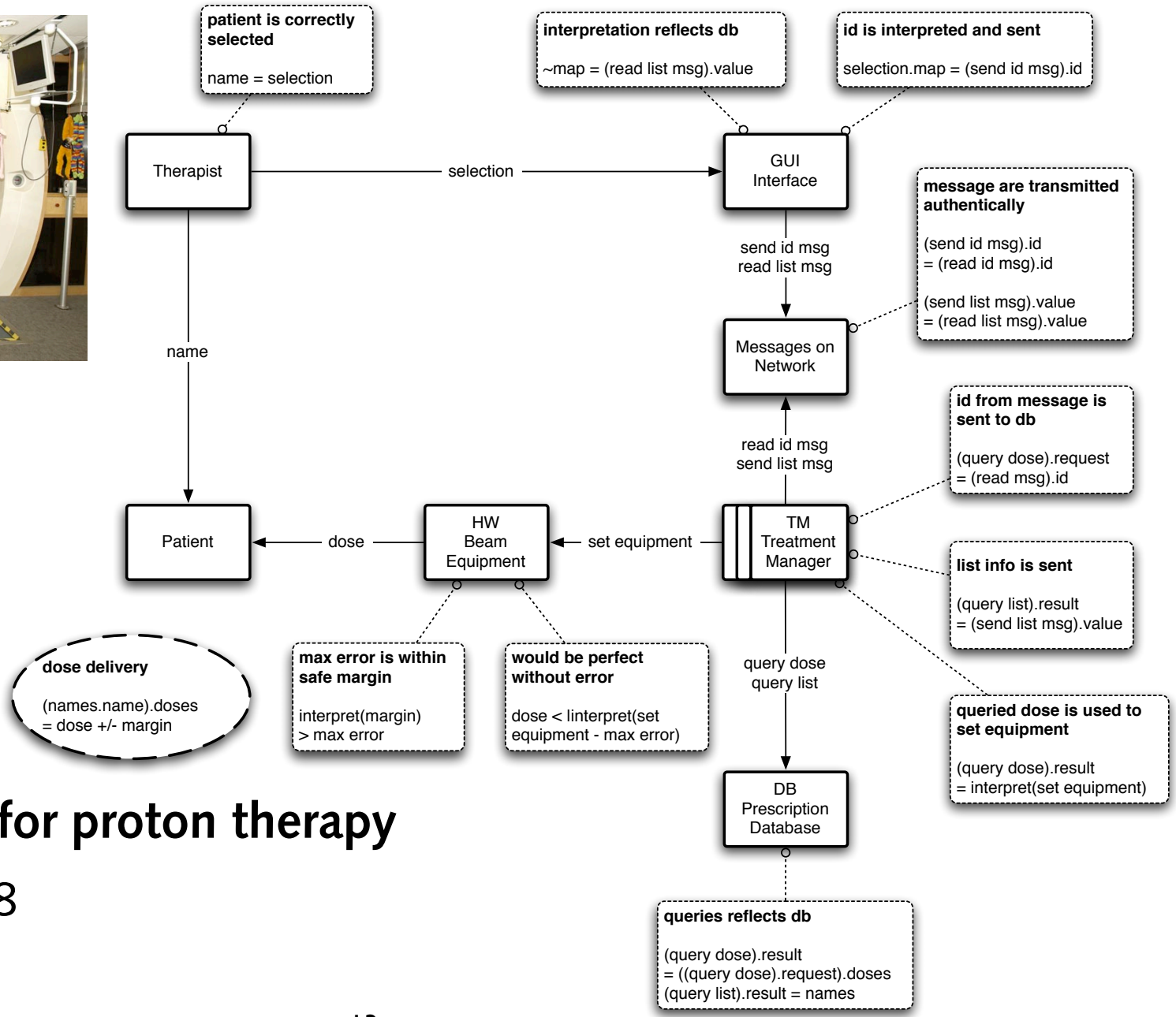
- PLUGR, Afghanistan 2001
- Airbus A320, Warsaw 1993

airborne ⇔ not WheelPulse ⇔ disabled

ENV ✗ MACHINE ✓

© Daniel Jackson 2008

an approach



dependability case for proton therapy

- Robert Seater, 2008

© Daniel Jackson 2008

conservative ≠ good

Korean Air 747, Guam 1997: 200 killed

If the ARTS IIA minimum safe altitude warning system had been operating as initially intended, a visual and aural warning would have activated about 64 seconds before flight 801 impacted terrain --*NTSB report*

© Daniel Jackson 2008

ignoring design

early blender patent

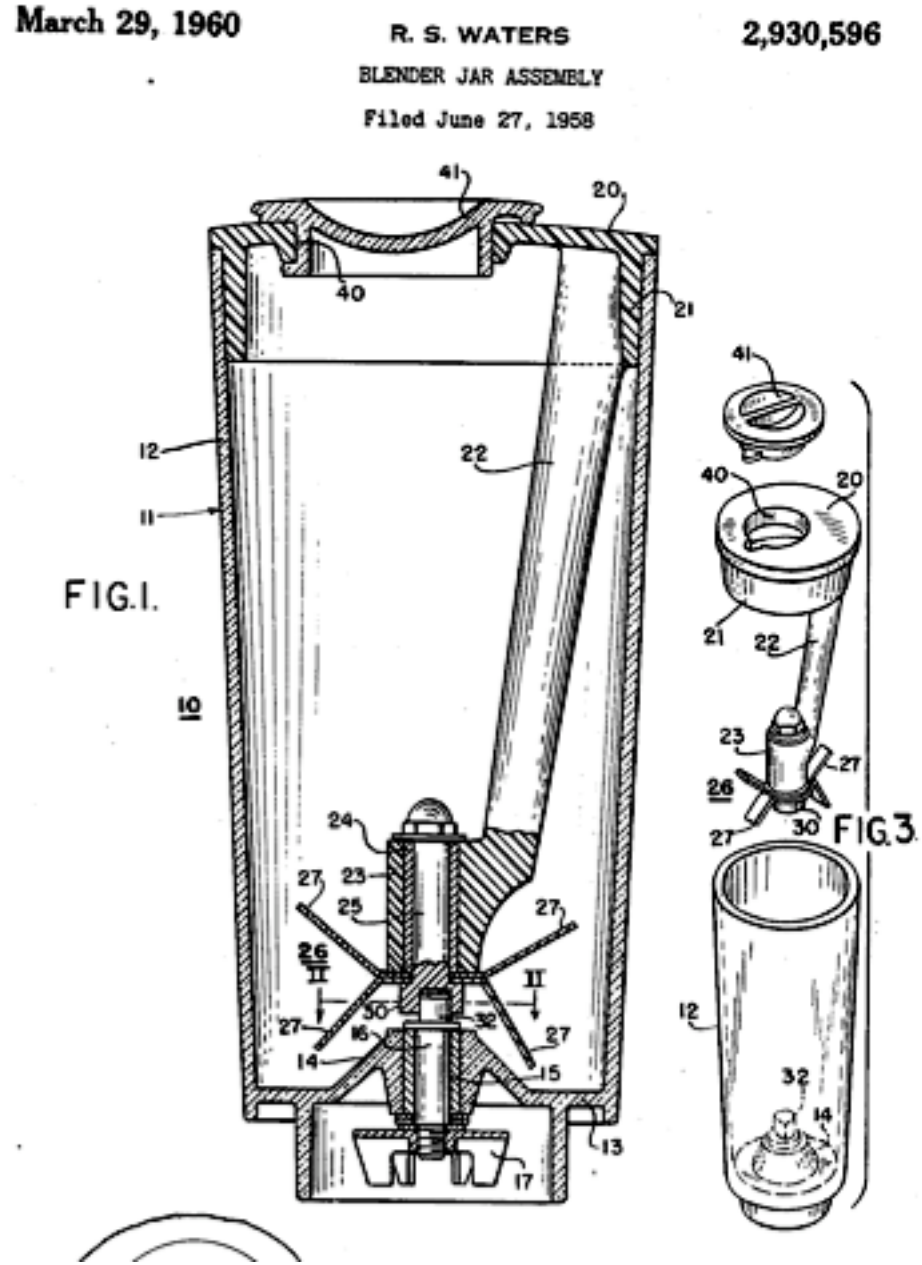
- opening too small for child's hand
- removal of closure disconnects blade

examples

- Therac 25: removed hardware interlock
- voting software: immutable types
- emergency stop: uses message queue

time to think again about

- safety kernels and modularity



© Daniel Jackson 2008

platform risk

IDE risk

- refactoring may not preserve meaning
- >7 such bugs open in Eclipse

language risk

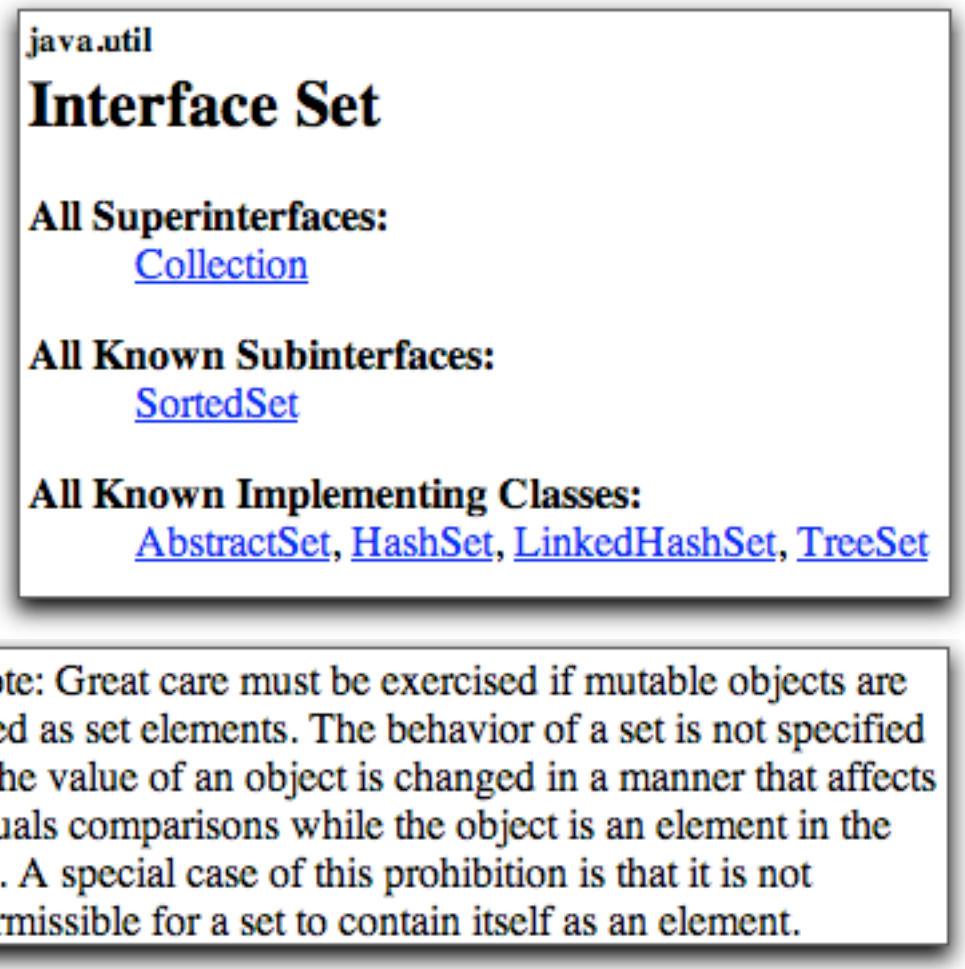
- in Java, eg: memory model, generics

operating system viruses

- time to infection for new PC: 4 mins

configuration problems

- DLLs, classpaths, etc



java.util

Interface Set

All Superinterfaces:
Collection

All Known Subinterfaces:
SortedSet

All Known Implementing Classes:
AbstractSet, HashSet, LinkedHashSet, TreeSet

Note: Great care must be exercised if mutable objects are used as set elements. The behavior of a set is not specified if the value of an object is changed in a manner that affects equals comparisons while the object is an element in the set. A special case of this prohibition is that it is not permissible for a set to contain itself as an element.

© Daniel Jackson 2008

social/managerial factors

process

does process really matter?

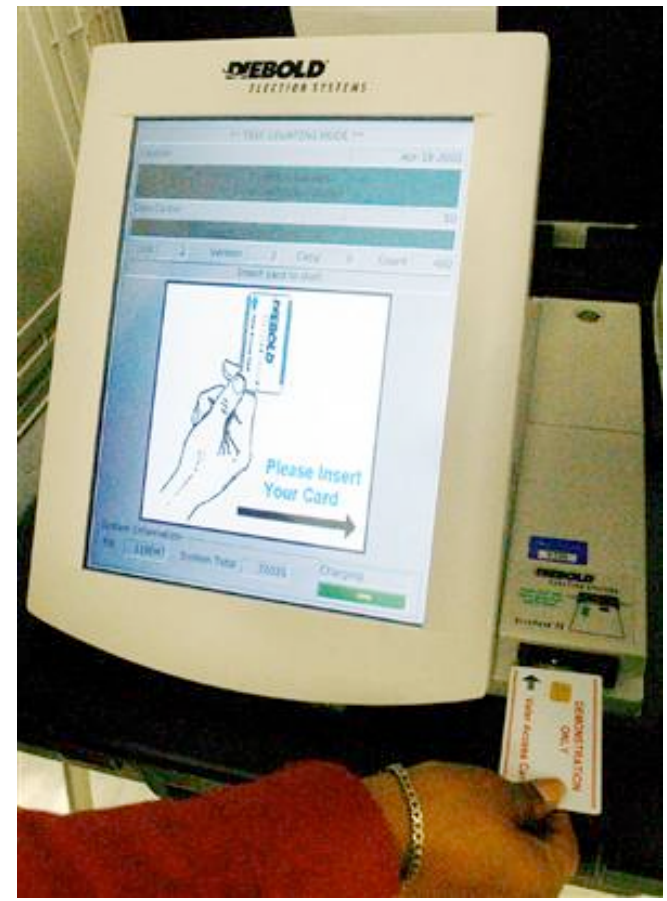
© Scott Adams, Inc./Dist. by UFS, Inc.

© Daniel Jackson 2008

bad process

Alameda County, CA, 2003

- 25% of voting machines crashed on boot
- so Diebold installed uncertified patches



Accuvote-TSx

© Daniel Jackson 2008

bad process

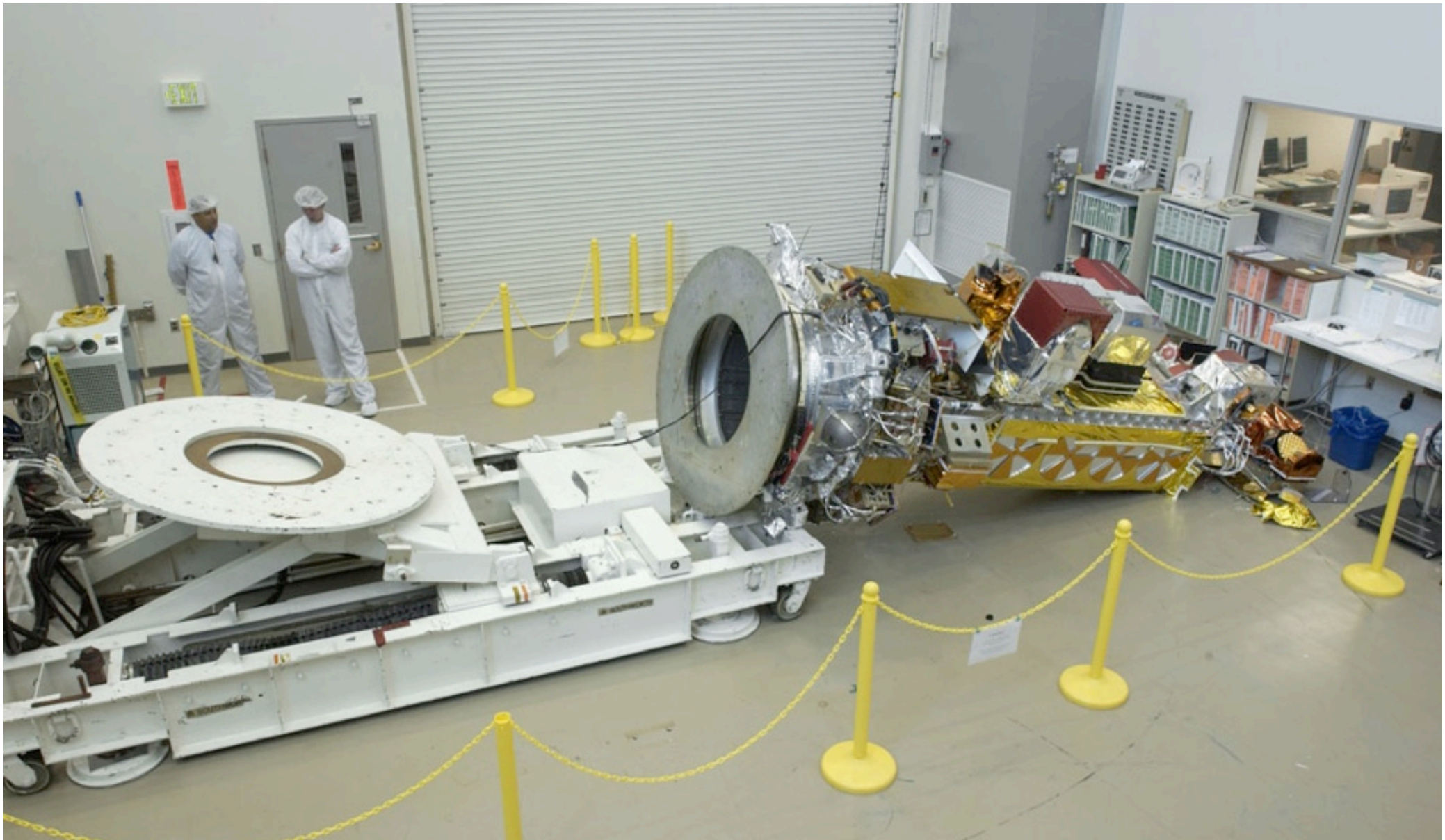
London Ambulance, 1992

- contract awarded to lowest bidder
- report from Arthur Andersen ignored
- no independent QA, software changes on-the-fly
- no incremental deployment, no paper backup
- untested change in operations

© Daniel Jackson 2008

neglecting process

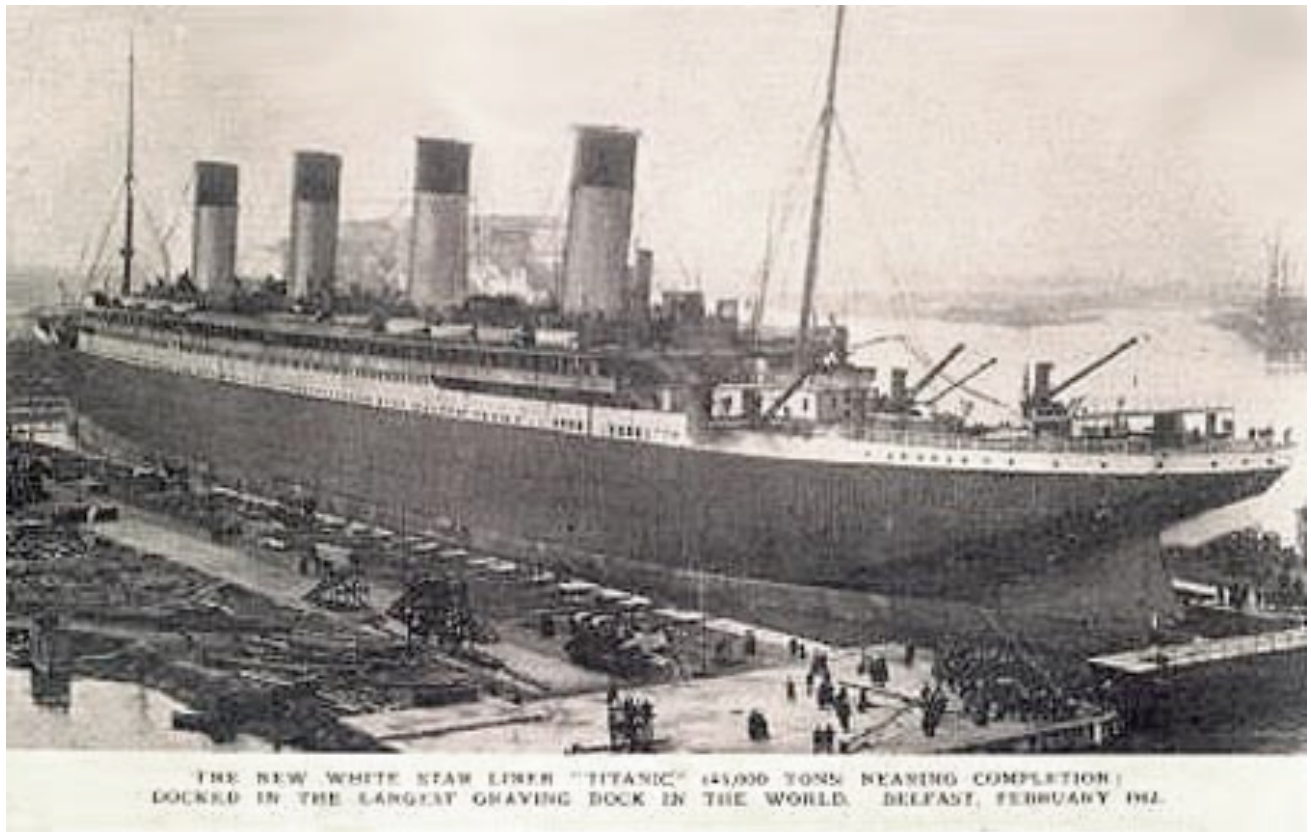
NOAA weather satellite at Lockheed Martin, September 2003



overconfidence

Titanic, 1912

- advanced technology, ‘unsinkable’
- so enough lifeboats not needed



© Daniel Jackson 2008

growing dangers

Texas A&M bonfire

- traditional began in 1928
- by 1990's, crane needed

what happened in 1999

© Daniel Jackson 2008

the risks of dependence

MAR knockout

- major Chicago hospital
- pharmacy database failure
- medication records lost

"Accidents are signals sent from deep within the system about the vulnerability and potential for disaster that lie within"
--Richard Cook and Michael O'Connor

MEDICATION CHART

FACILITY

LEGENDCARE PHARMACY

LEGEND
2330 MC
NORMAN
(405) 321

PATIENT NAME: BERRIOS, ANGELA

DATE OF BIRTH: 2/27/2007 | SEX | MO: 12 | YEAR: 07

HOUR: 1 2 3 4 5 6 7 8 9 10 11 12 13

Rx#: 212379 Dr LEE P. FRYE
DOCUSATE SOD 100MG CA
TAKE 1 CAPSULE TWICE DAILY AT 8:00 A.M. AND 8:00 P.M.
Brand: DSS
08:00 AM
08:00 PM

Rx#: 215090 Dr LEE P. FRYE
DOXYCYCLINE 100MG CAPS
TAKE 1 CAPSULE BY MOUTH THREE TIMES DAILY FOR ACNE
Brand: VIBRAMYCIN
08:00 AM
02:00 PM
08:00 PM

Rx#: 212381 Dr LEE P. FRYE
LORATADINE 10MG TAB
TAKE ONE TABLET BY MOUTH AT 8:00 A.M. AND 8:00 P.M. EVERY DAY
Brand: CLARITIN
08:00 AM

Rx#: 212380 Dr LEE P. FRYE
LORAZEPAM 0.5MG TAB (GEN ATIVAN)
TAKE ONE TABLET BY MOUTH THREE TIMES DAILY FOR ANXIETY
Brand: ATIVAN
08:00 AM
02:00 PM
08:00 PM

DIAGNOSIS & COMMENTS
DEPRESSION

ALLERG
LEVODOPA (DOPAR)

© Daniel Jackson 2008

blame the user

USS Yorktown, 1997

- dead in water for 3 hours

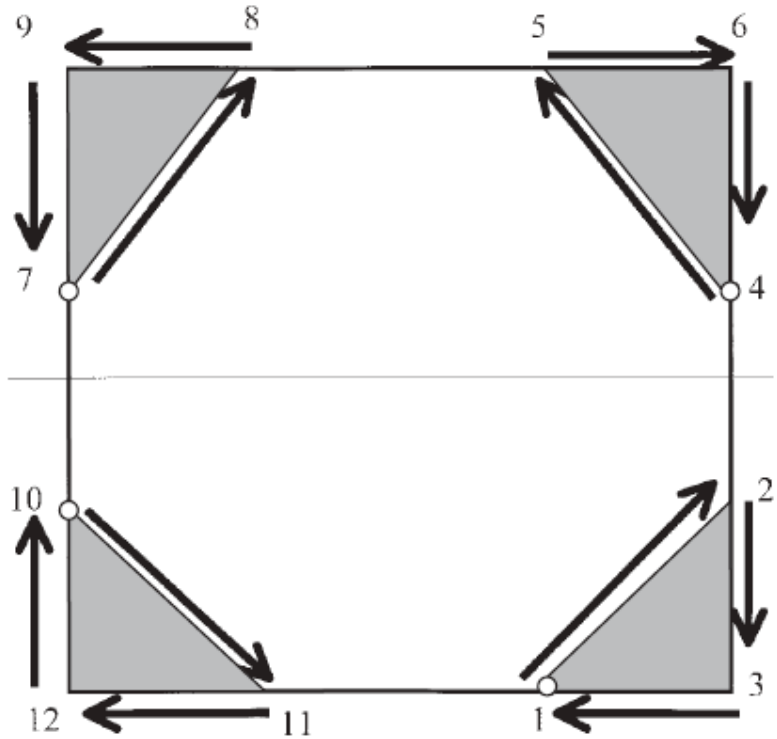
Managers are now aware of the problem of entering zero into database fields and are trained to bypass a bad data field and change the value... ships do go dead in the water... People sometimes make mistakes and systems break. The trick is we have trained our crew...

-- Commanding Officer, USS Yorktown

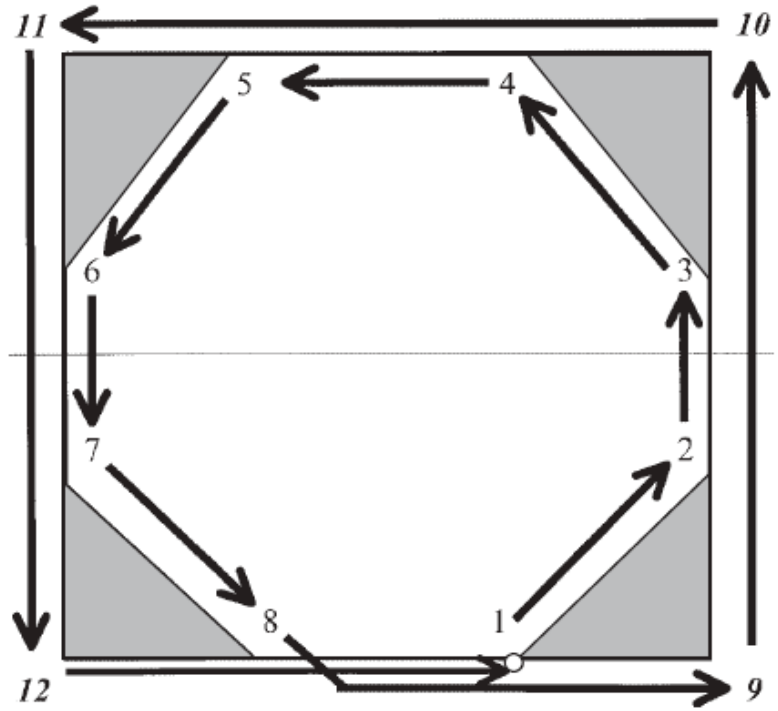
© Daniel Jackson 2008

panama radiation accident

dose = D



dose = 2D



Panama City Hospital, 2001

- Theratronic-780 with therapy planning system by Multidata
- 18 patients killed

© Daniel Jackson 2008

panama consequences

3 Panama physicists tried for second-degree murder

- Olivia Saldaña González paid for her own defence; earns $585/month
- sentenced to four years in prison
- suit by families against Multidata rejected by Panama court

Given [the input] that was given, our system calculated the correct amount, the correct dose. It was an unexpected result. And, if [the staff in Panama] had checked, they would have found an unexpected result.
-- Mick Conley, Multidata

© Daniel Jackson 2008

conclusions

implications for research

if you reward publication, you get

- focus on logic & algorithms
- benchmarks, not real problems
- throwaway implementations

some good strategies

- fund tool development [NSF infrastructure]
- issue challenges [VSR's Mondex, Flash]
- publish case studies [Z, Patterns]

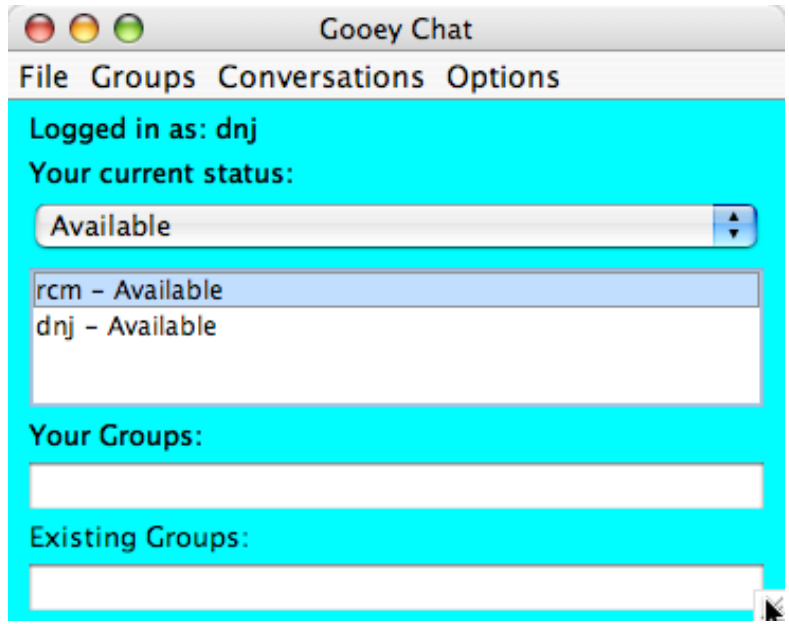
will interdisciplinary work help?

© Daniel Jackson 2008

implications for teaching

what we typically do

- focus on 'respectable' topics (eg, semantics)
- illustrate with small problems
- say hard parts are out-of-scope
- set formal problems that are easy to grade



instead, we might

- explain 'soft' aspects too
- illustrate with substantial case studies
- address the hard parts
- set informal, open-ended problems

© Daniel Jackson 2008

thank you!