CASE-BASED REASONING VS PARAMETRIC MODELS FOR SOFTWARE QUALITY OPTIMIZATION

Adam Brady : adam.m.brady@gmail.com
Tim Menzies : tim@menzies.us

Lane Dept. of Computer Science and Electrical Engineering
West Virginia University, USA

Is "effort estimation" the wrong question?

- Don't tell me what *is*, tell me what to *change*."
  - -- An irate business user, May'00

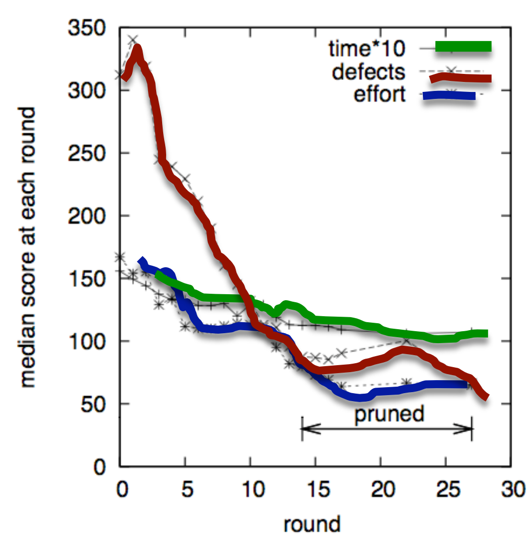
- After prediction…
  - Comes the *planning* that *improves* that prediction

- SQO= software quality optimization
  - adjusting a software project to improve attributes e.g.
    - defects (number of delivered defects),
    - months (calendar time to delivery)
    - effort (staff time, in person months, required for that delivery).

"Model-based" SQO

- Use a predictive model (e.g. COCOMO) as a sub-routine in a what-if analysis
  - Guide the what-if analysis with some AI algorithm
    - A*,
    - Beam,
    - simulated annealing,
    - ISAMP,
    - MaxWalkSat,
    - …
  - As done by Menzies et al. in ASE'07, ICSP'08, ICSE'09, PROMISE'09, SPIP 2009, ASE'09

- Repeat 1000 times
  - p = a possible project
  - t = a possible model tuning
  - Out = value(model(p, t))
- Search for ranges in p that maximize Out
- Test those ranges by calling the model again



This talk: "case-based" SQO?

- "W"
  - "dubya": the decider
  - a very greedy algorithm
  - Meant to have been a straw-man baseline tool
    - But it out-performs ASE'07, ICSP'08, ICSE'09, PROMISE'09, SPIP 2009, ASE'09
- No need for a process model
  - embarrassingly simple

Definitions

Model-based, Case-based Reasoning

Case-based, model-based reasoning

Model-based

- Training data + learner ➔ model
- Test case + model ➔ prediction
- Most of current PROMISE papers

Case-based

- Test case + Training data ➔ neighbors
- Neighbors ➔ prediction

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

3

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

3

4a

BEST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12
08	5	3	2	3	3	2	4	3	3	2	4	3	3	3	36
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42

4b

REST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
12	5	3	4	3	3	2	4	3	3	2	4	4	3	3	48
11	4	3	4	3	3	2	4	3	3	2	4	4	3	3	60
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60
19	4	2	4	4	3	5	4	5	5	2	5	3	3	2	62
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300
07	5	3	4	3	3	2	4	3	3	2	4	5	3	3	360
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

3

4a

BEST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12
08	5	3	2	3	3	2	4	3	3	2	4	3	3	3	36
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42

4b

REST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
12	5	3	4	3	3	2	4	3	3	2	4	4	3	3	48
11	4	3	4	3	3	2	4	3	3	2	4	4	3	3	60
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60
19	4	2	4	4	3	5	4	5	5	2	5	3	3	2	62
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300
07	5	3	4	3	3	2	4	3	3	2	4	5	3	3	360
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648

5

Treatment	$b = freq(x\|best)$	$r = freq(x\|rest)$	*like* (Eq 2) $(b/5)^2/(b/5 + r/15)$
pmat=3	5	10	60%'
sced=3	5	13	54%
tool=3	5	14	52%
acap=3	4	7	51%
data=3	4	9	46%
rely=4	3	6	36%
time=3	3	7	34%
pvol=4	2	2	30%
stor=3	3	10	28%
cplx=5	2	3	27%
stor=5	2	3	27%
cplx=4	3	12	26%
time=5	2	4	24%
pvol=3	2	5	22%
data=2	2	5	22%
rely=3	2	9	16%
pvol=2	1	9	5%

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

3

4a

BEST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12
08	5	3	2	3	3	2	4	3	3	2	4	3	3	3	36
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42

4b

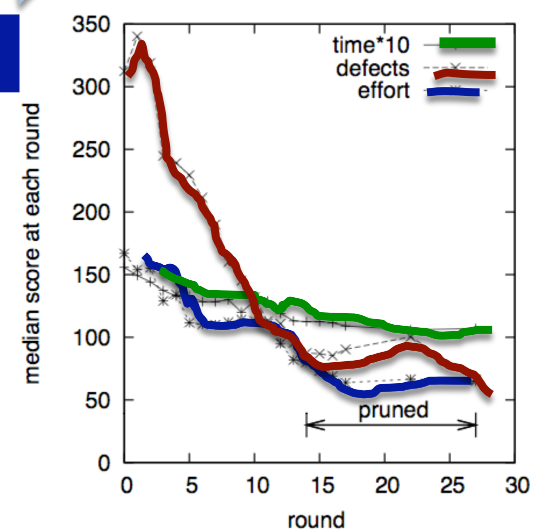
REST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
12	5	3	4	3	3	2	4	3	3	2	4	4	3	3	48
11	4	3	4	3	3	2	4	3	3	2	4	4	3	3	60
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60
19	4	2	4	4	3	5	4	5	5	2	5	3	3	2	62
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300
07	5	3	4	3	3	2	4	3	3	2	4	5	3	3	360
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648

5

Treatment	$b = freq(x\|best)$	$r = freq(x\|rest)$	*like* (Eq 2) $(b/5)^2/(b/5+r/15)$
pmat=3	5	10	60%'
sced=3	5	13	54%
tool=3	5	14	52%
acap=3	4	7	51%
data=3	4	9	46%
rely=4	3	6	36%
time=3	3	7	34%
pvol=4	2	2	30%
stor=3	3	10	28%
cplx=5	2	3	27%
stor=5	2	3	27%
cplx=4	3	12	26%
time=5	2	4	24%
pvol=3	2	5	22%
data=2	2	5	22%
rely=3	2	9	16%
pvol=2	1	9	5%

6



1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

3

4a

BEST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12
08	5	3	2	3	3	2	4	3	3	2	4	3	3	3	36
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42

4b

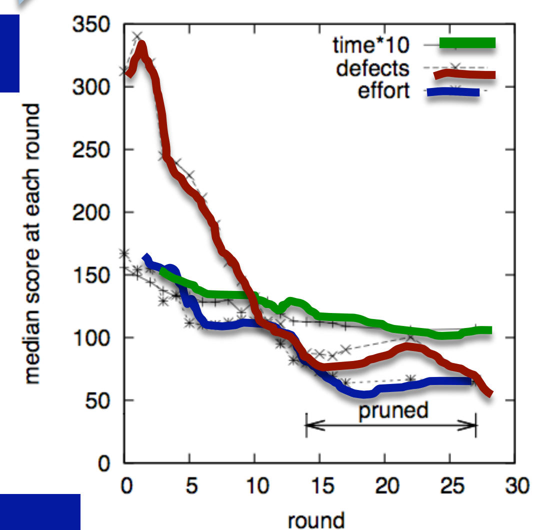
REST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
12	5	3	4	3	3	2	4	3	3	2	4	4	3	3	48
11	4	3	4	3	3	2	4	3	3	2	4	4	3	3	60
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60
19	4	2	4	4	3	5	4	5	5	2	5	3	3	2	62
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300
07	5	3	4	3	3	2	4	3	3	2	4	5	3	3	360
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648

5

Treatment	$b = freq(x\|best)$	$r = freq(x\|rest)$	*like* (Eq 2) $(b/5)^2/(b/5+r/15)$
pmat=3	5	10	60%'
sced=3	5	13	54%
tool=3	5	14	52%
acap=3	4	7	51%
data=3	4	9	46%
rely=4	3	6	36%
time=3	3	7	34%
pvol=4	2	2	30%
stor=3	3	10	28%
cplx=5	2	3	27%
stor=5	2	3	27%
cplx=4	3	12	26%
time=5	2	4	24%
pvol=3	2	5	22%
data=2	2	5	22%
rely=3	2	9	16%
pvol=2	1	9	5%

6



Repeat 20 times

1

```
@project brookslaw
@attribute apex 2
@attribute plex 1 2
@attribute ltex 1 2 3
@attribute ?pmat 2 3
@attribute ?rely 3 4 5
@attribute ?data 2 3
@attribute ?cplx 4 5
@attribute ?time 4 5
@attribute ?stor 3 4 5
@attribute ?pvol 2 3 4
@attribute ?acap 3 4 5
@attribute ?pcap 3 4 5
@attribute ?tool 3 4
@attribute ?sced 2 3
```

A membership function returning 0 to 14

Simplest case of the KEYS2: see Gay, Menzies et al ASE journal, Dec'10

2

Relevancy Filtering: Brooks' Law Query, NASA93 Dataset

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort	overlap
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38	13
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12	13
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480	13
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648	13
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370	12
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114	12
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215	12
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636	11
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42	11
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60	11
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42	11
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210	11
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90	11
47	3	4	4	4	4	3	5	4	4	2	4	3	3	3	703	10
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
43	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300	10
41	4	4	4	2	4	3	4	3	5	2	4	4	3	2	576	10
36	3	2	3	4	3	4	5	3	3	2	4	5	3	2	278	10
34	4	3	4	2	3	4	4	5	3	3	4	4	3	3	155	10
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8	10
(39 cases omitted)																
54	4	4	4	4	5	4	5	6	6	3	4	4	3	3	8211	7
52	4	4	4	4	5	4	5	6	6	3	4	4	3	3	1645.9	7
51	4	4	4	4	5	4	5	6	6	3	4	4	3	3	4178.2	7

3

4a

BEST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
56	3	2	2	3	4	3	5	5	5	4	3	3	3	3	12
08	5	3	2	3	3	2	4	3	3	2	4	3	3	3	36
57	3	2	2	3	4	3	5	5	5	4	3	3	3	3	38
22	3	3	3	3	4	3	4	3	3	3	3	3	3	3	42
25	3	3	3	3	3	3	4	3	3	3	3	3	3	3	42

4b

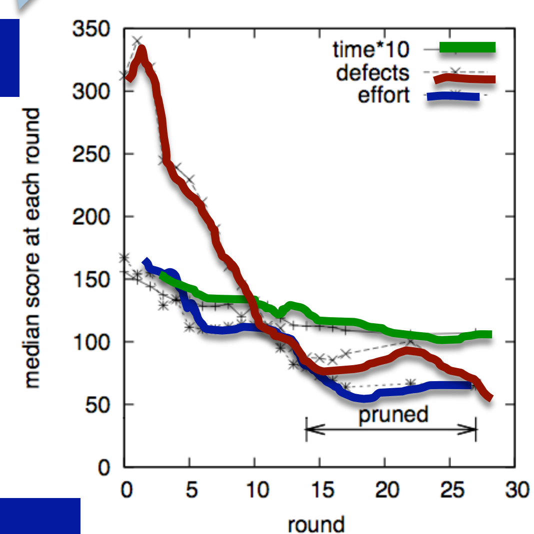
REST Set

row	*apex*	*plex*	*ltex*	pmat	rely	data	cplx	time	stor	pvol	acap	pcap	tool	sced	effort
12	5	3	4	3	3	2	4	3	3	2	4	4	3	3	48
11	4	3	4	3	3	2	4	3	3	2	4	4	3	3	60
23	3	3	3	3	3	3	4	3	3	3	3	3	3	3	60
19	4	2	4	4	3	5	4	5	5	2	5	3	3	2	62
16	4	3	3	3	4	3	3	4	3	3	3	4	3	3	90
33	4	3	4	2	3	4	4	5	3	3	4	4	3	3	98.8
26	3	3	3	3	3	3	4	4	3	3	3	3	3	3	114
17	4	3	3	3	4	3	4	3	3	3	3	4	3	3	210
09	4	2	1	3	3	2	4	3	3	4	4	4	3	3	215
44	4	4	4	2	3	3	4	3	5	2	4	4	3	2	300
07	5	3	4	3	3	2	4	3	3	2	4	5	3	3	360
35	4	3	3	2	4	3	4	4	4	2	3	3	3	3	370
55	3	2	2	3	4	3	5	5	5	4	3	3	3	3	480
40	4	3	4	3	4	3	4	4	3	2	4	4	3	3	636
53	2	1	2	2	5	2	5	5	6	2	4	3	4	3	648

5

Treatment	$b = freq(x\|best)$	$r = freq(x\|rest)$	*like* (Eq 2) $(b/5)^2/(b/5+r/15)$
pmat=3	5	10	60%'
sced=3	5	13	54%
tool=3	5	14	52%
acap=3	4	7	51%
data=3	4	9	46%
rely=4	3	6	36%
time=3	3	7	34%
pvol=4	2	2	30%
stor=3	3	10	28%
cplx=5	2	3	27%
stor=5	2	3	27%
cplx=4	3	12	26%
time=5	2	4	24%
pvol=3	2	5	22%
data=2	2	5	22%
rely=3	2	9	16%
pvol=2	1	9	5%

6



Repeat 20 times

"W" = CBR + query adaption

Standard CBR = "what is"

- Q= query
- C= cases
- Predict = adapt(Q $\wedge$C)

"W" = "what to change"

- Q0 = query
- N times repeat
  - C = Cases = Train + Prune
  - Predict0 = adapt(Q0 $\wedge$Train)
  - Q1 = Q0 + treatment
  - Predict1 = adapt(Q1 $\wedge$Prune)
- Best treatment:
  - maximizes Predict1/Predict0
  - And is stable across the repeats
- Q: Why just "Prune"
  - And not "Prune" and "Test"?
  - A: Some very small data sets

Experimental Comparisons

Model-based vs CBR

Is “W” beaten by other methods?

- STAR, NOVA:
  - model-based what-if exploration of software process models
  - Oracle = COCOMO (effort); COQUALMO (defects)
  - ASE’07, ICSP’08, ICSE’09, PROMISE’09, SPIP 2009, ASE’09
- STAR:
  - Simulated annealing (SA) to explore project options
  - Tuning options (T) sampled Monte Carlo
  - SA explored project options (p) looking for changes
    - That caused stable outputs
    - Despite tuning uncertainties

$$\arg\max_{x}\left(\overbrace{r_x \subseteq p}^{AI\ search},\underbrace{t \subseteq T, value(model(r_x, t))}_{Monte\ Carlo}\right)$$

- NOVA:
  - SA, beam, issamp, keys, maxwalksat, A*, SEESAW,….

“W” vs NOVA

- NOVA’s best search method
  - SEESAW (ASE’09)
- Project descriptions from NASA
  - Flight, ground systems
  - Two flight guidance system: OSP, OSP2
- Case study data = NASA93
  - 50 times,
    - divided into “train” and “test” (66% / 33%)

	ranges			values	
project	feature	low	high	feature	setting
OSP2	prec	3	5	flex	3
	pmat	4	5	resl	4
	docu	3	4	team	3
	ltex	2	5	time	3
	sced	2	4	stor	3
	KSLOC	75	125	data	4
				pvol	3
				ruse	4
				rely	5
				acap	4
				pcap	3
				pcon	3
				apex	4
				plex	4
				tool	5
				cplx	4
				site	6
JPL ground software	rely	1	4	tool	2
	data	2	3	sced	3
	cplx	1	4		
	time	3	4		
	stor	3	4		
	acap	3	5		
	apex	2	5		
	pcap	3	5		
	plex	1	4		
	ltex	1	4		
	pmat	2	3		
	KSLOC	11	392		

Testing the treatments

- Query= NASA project descriptions
- Learn a treatment
- Testing:
  - Sort test data by overlap with Query + Treatment
  - Look at up to 20 cases with greatest overlap
  - (In practice, Query +Treatment selects for less than 20 test cases).
- Treatment1: CBR ("W")
  - Learn treatment1 with "W" from "train", assuming query
  - Assess on "test"
- Treatment2: model-based (SEESAW)
  - Constrain inputs to just the query
  - Learn treatment2 with NOVA
  - Assess on "test"

Results

- Learner= (SEESAW or W)
- Project= (Flight, ground, OSP, or OSP2)
- Goal= (defects, months, or effort)
- Report= (median or spread)
  - Median = 50th percentile
  - Spread = (75-25)th percentile

- Finding1:
  - In 94% of runs, reductions observed
  - So SQO works

EFFORT (total staff months)

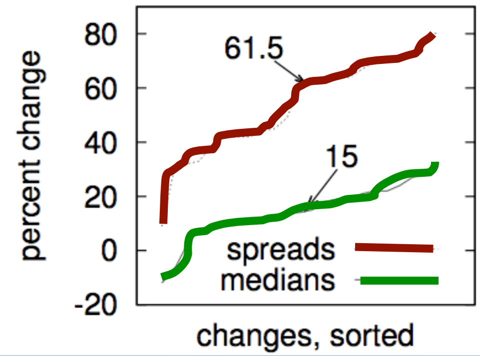
		Algorithm	Before b	After a	Change (b-a)/a
Medians	Ground	SEESAW	269	197	27%
		W	269	184	32%
	Flight	SEESAW	258	252	2%
		W	258	208	19%
	OSP	SEESAW	270	195	28%
		W	270	210	22%
	OSP2	SEESAW	291	269	8%
		W	291	227	22%

EFFORT (total staff months)

		Algorithm	Before b	After a	Change (b-a)/a
Spreads	Ground	SEESAW	526	99	81%
		W	526	125	76%
	Flight	SEESAW	567	204	64%
		W	567	165	71%
	OSP	SEESAW	350	114	67%
		W	350	100	71%
	OSP2	SEESAW	824	418	49%
		W	824	299	64%

Results (more)

- Finding2:
  - SEESAW's / W's recommendations reduced the spread
  - So SQO increases certainty



- Finding3:
  - In absolute terms …
  - Improvements seen by SEESAW/W very similar

- 19 vs 18% change (for Ground)
- 14 vs 9% change (for Flight)
- 11 vs 12% change (for OSP)
- 12 vs 14% change (for OSP2)

- Finding4:
  - In relative terms …
  - W wins a little more: Mann-Whitney, 95%

Algorithm	Wins	Losses	Ties
W	6	0	18
SEESAW	0	6	18

Other observations

- Simple to code: 200 lines of AWK vs 5000 lines of LISP
- Faster to run: less than second (in awk!!!)
- Simpler to maintain
  - in CBR, “maintenance” = “add more cases”
- No use of an underlying model
  - Free of any (possibly wrong) assumptions of parametric modeling.
  - Can be quickly applied to more data sets.
    - E.g. our older implementations required data in COCOMO format
    - W has been applied to numerous other formats

Discussion

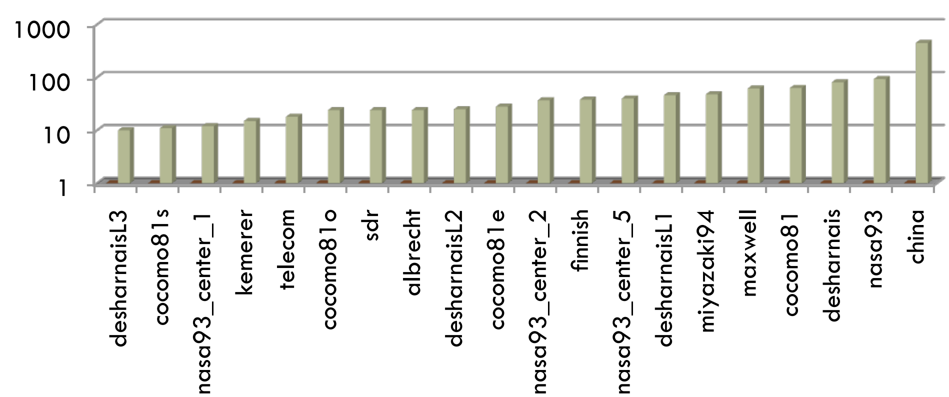
Case-based vs model-based SQP

- "W"
  - finds similar or better optimizations than more complex model-based approaches
- Simpler to code
- Faster to run
- Easier to maintain
- No model assumptions
- No restrictions arising from model ontology

Is simple sufficient?

- SE data is a shallow well?
  - Seemingly sophisticated algorithms can be no better than simpler alternatives

  See also, Lessmann et al TSE'08 , Menzies et al. ASEJ,10
- Bias in our field: always go for the more complex
  - Ant algorithms to find best splits for model trees
  - Cloud computing to try 10,000 variants of a machine learner
  - …
- Estimation data = small
  - misguided to spend day of CPU time
  - to analyze only a handful of instances



Scope of these conclusions

- Do all CBR methods are better than all model-based method?
- Nope
  - Here, just SQO
- But this result should motivate more exploration of CBR for SE data

Questions or comments?