For SSBSE 2015 Software Verification:

Epistatic Genetic Algorithm for Test Case Prioritization

Fang Yuan, Yi Bian, Zheng Li, Ruilian Zhao

Reported by:

Yi Bian

05.09.2015

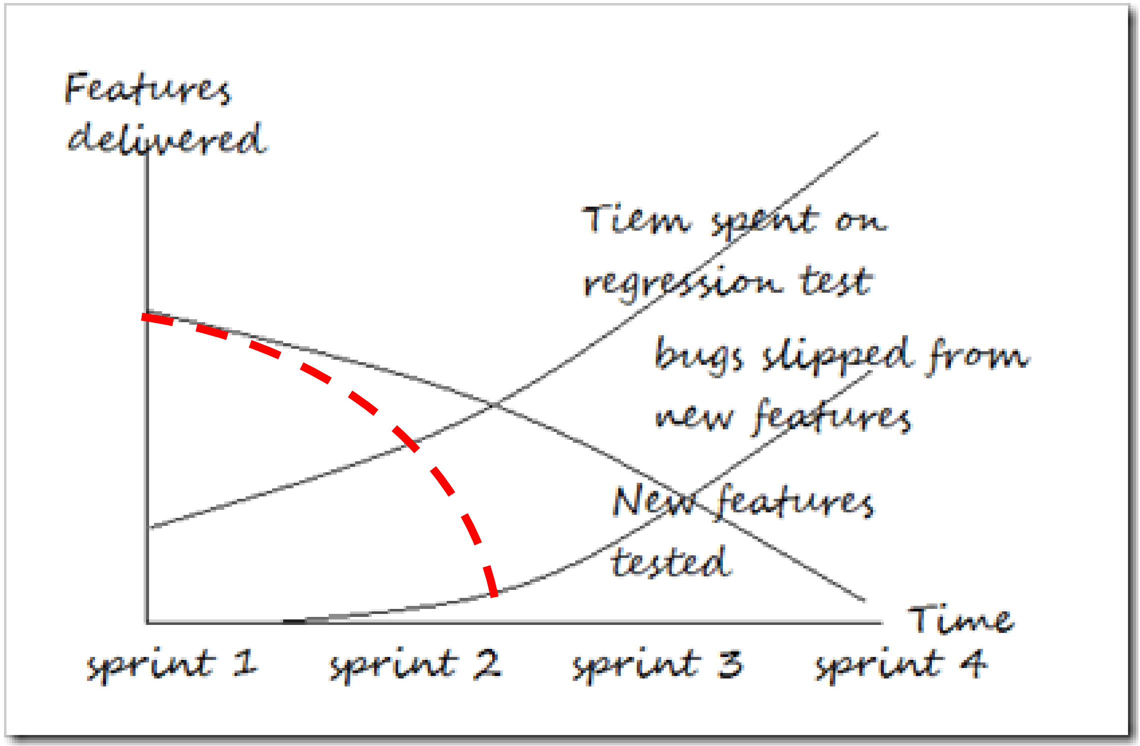
Introduction

Beijing University of Chemical Technology (BUCT)

Test Case Prioritization

In Regression Testing …

For TCP, we hope that …



*This picture is from the google.

How to Convince Developers and Management to Use Automated Test instead of Manual Test

Test Case Prioritization

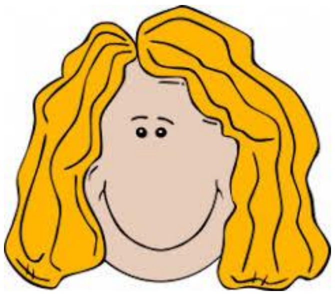
Average reversion time for some software

Table 2: Estimated commit and performance testing frequency in popular software.

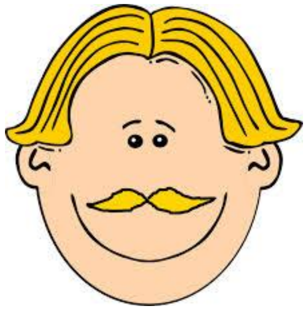
Software	Avg. Rev. per Day	Regular Perf. Testing
MySQL	~ 6	every release [49]
Chrome	~ 140	every 4 rev.
Linux	~ 140	every week [21]

*Huang P, Ma X, Shen D, et al. Performance regression testing target prioritization via performance risk analysis[C]//Proceedings of the 36th International Conference on Software Engineering. ACM, 2014: 60-71.

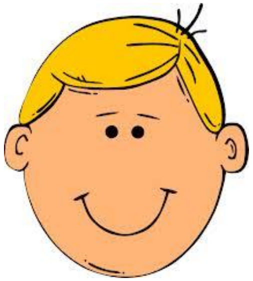
Epistasis in Biology



blond hair gene

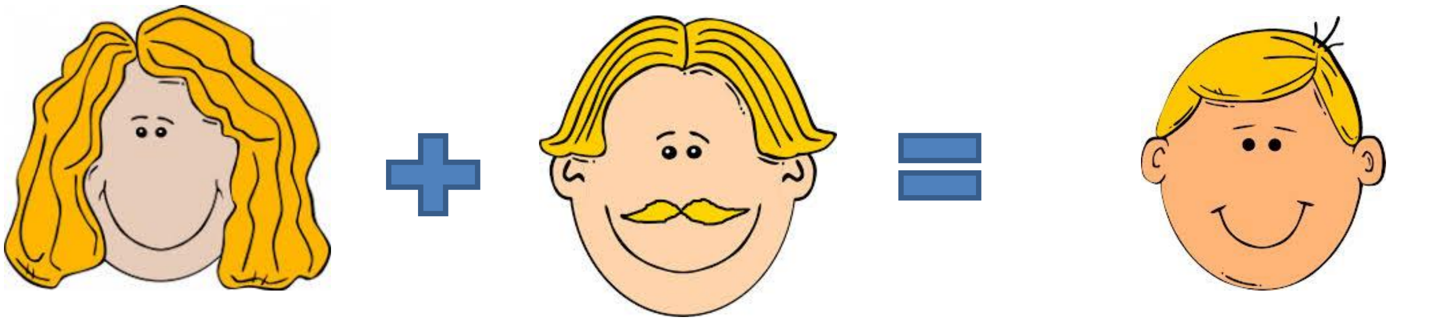


blond hair gene

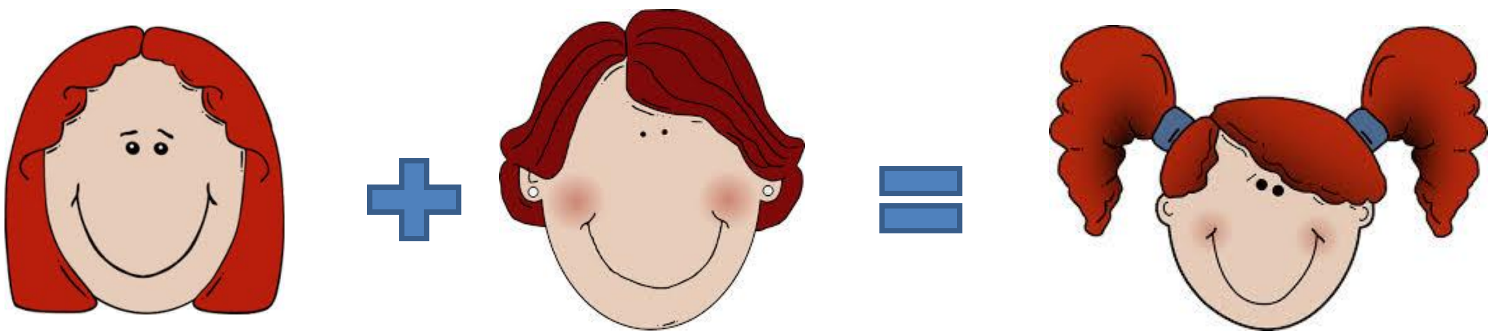


blond hair gene

Epistasis in Biology



blond hair gene + blond hair gene = blond hair gene



red hair gene + red hair gene = red hair gene

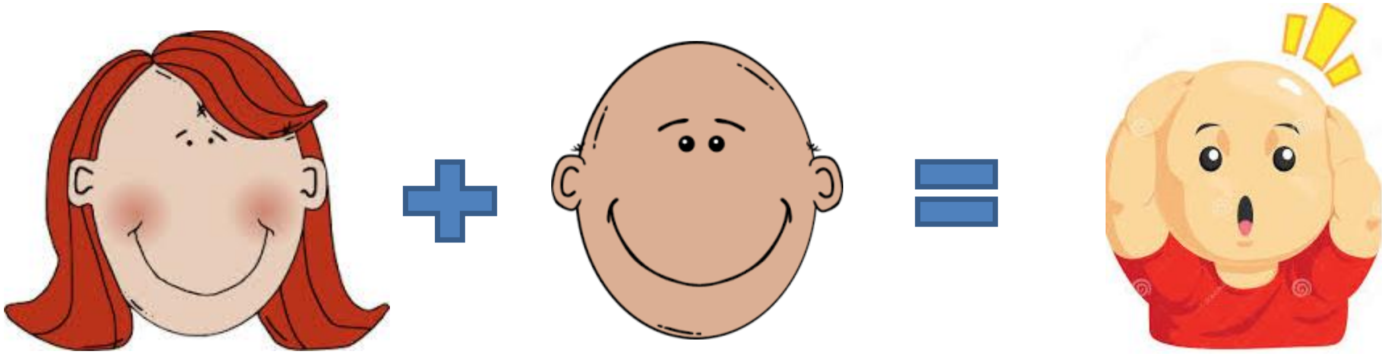
Epistasis in Biology

red hair gene

+ ? =

red hair gene

Epistasis in Biology



red hair gene
& not bald gene

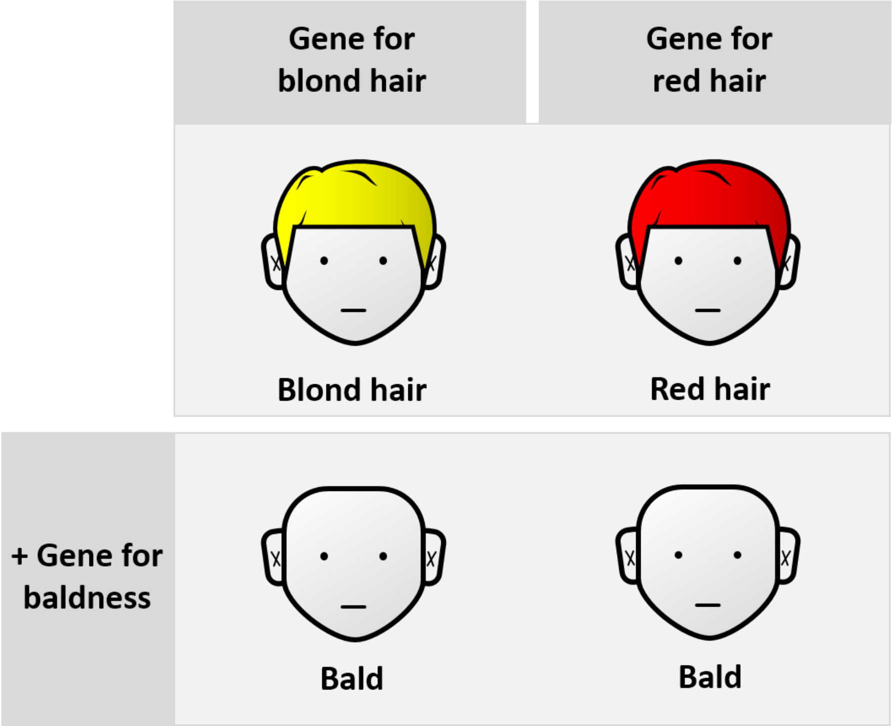
* hair gene
& bald gene

red hair gene
& bald gene

Epistasis in Biology

Wikipedia:

Epistasis is a phenomenon that consists of the effect of one gene being dependent on the presence of one or more 'modifier genes' (genetic background). Similarly, epistatic mutations have different effects in combination than individually.



The gene for total baldness is epistatic to those for blond hair or red hair. The baldness phenotype supersedes genes for hair colour and so the effects are non-additive.

Epistasis

Epistasis Variance:

Davidor Y. Epistasis variance: Suitability of a representation to genetic algorithms[J]. Complex Systems, 1990, 4(4): 369-383.

Based on binary encoding:

Ind 1:	0	0	1		0
Ind 2:	1	0	0		0
Ind 3:	1	1	1	...	1
Ind 4:	0	1	1		1
Ind 5:	0	1	1		1

Their fitness value v(s) are higher

$$A_i(a) = \frac{1}{N_i(a)} \sum_{S \in Pop_{s_i=a}} v(S)$$

$$X_i(a) = A_i(a) - \bar{V}$$

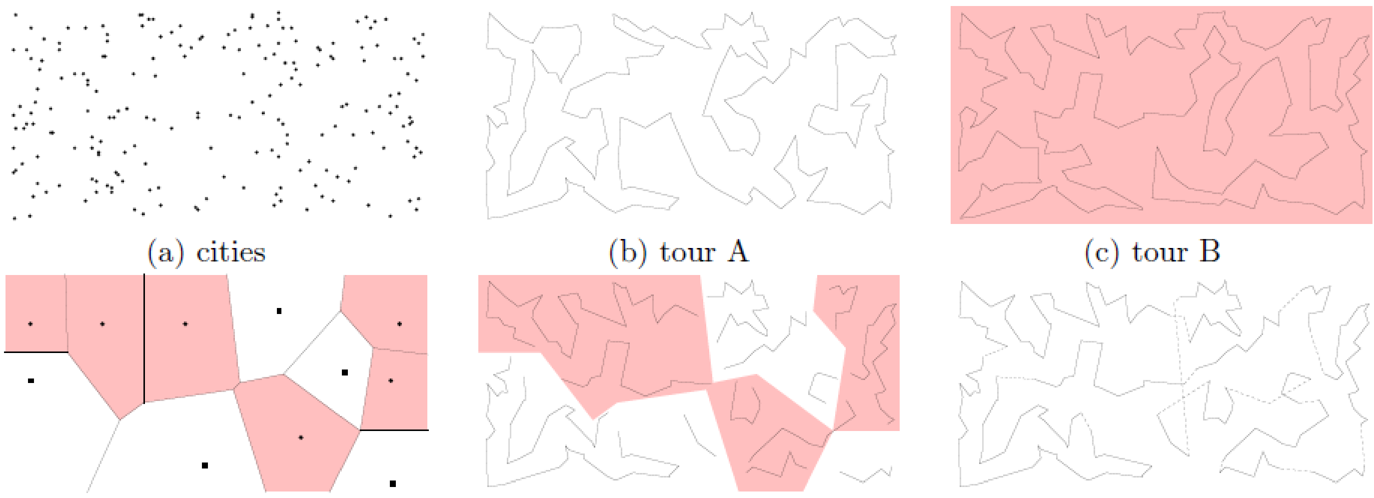
Excess genic value

Epistatic crossover

epistatic gene groups：

Seo D, Moon B R. Voronoi Quantizied Crossover For Traveling Salesman Problem[C]// GECCO. 2002: 544-552.

Voronoi quantized crossover:



(a) cities (b) tour A (c) tour B

(d) Voronoi regions (e) inherited tour segments (f) repaired tour segments

Figure 4: An example VQX for TSP (kroA200)

Epistatic crossover

epistatic gene groups：

Seo D, Moon B R. Voronoi Quantizied Crossover For Traveling Salesman Problem[C]// GECCO. 2002: 544-552.

Voronoi quantized crossover:

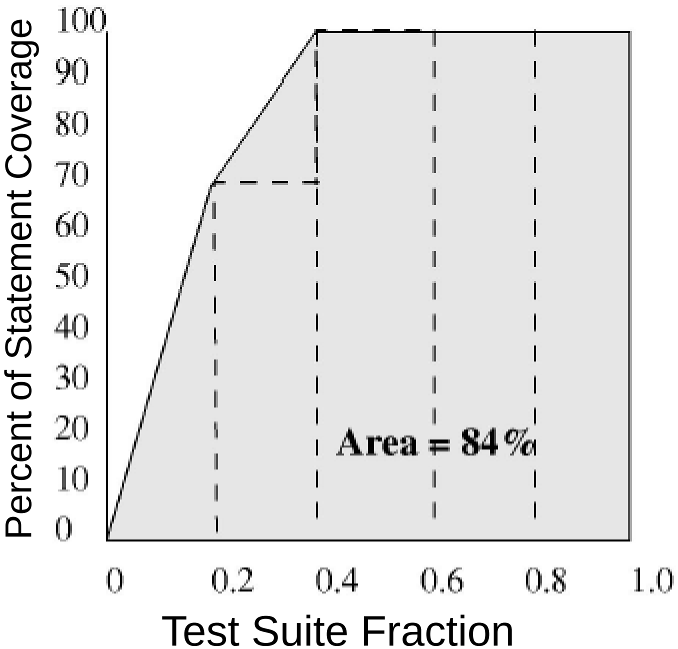
(Base on permutation encoding, similar as test case prioritization)

- The strength of the epistasis of the gene group is an inherent property of the problem.
- However, no practical method is known yet for exactly computing epistasis.

Prioritization target

In test case prioritization process

The percentage of statement coverage is gradually increasing.



Average Percentage of Statement Coverage:

$$\mathrm{APSC} = \left(1 - \frac{TS_1 + TS_1 + \cdots + TS_M}{MN} + \frac{1}{2N}\right) * 100\%$$

N is the number of test cases,
M is total number of statements,
TSi denotes the identifier of the test case that first covers the statement i in the execution sequence.

Prioritization target

An example

Code lines ***Coverage matrix:***

Code lines	TC1	TC2	TC3	TC4
0	1	0	0	1
1	0	1	0	1
2	0	0	1	1
3	0	0	0	1

Execution sequence:

Prioritization target

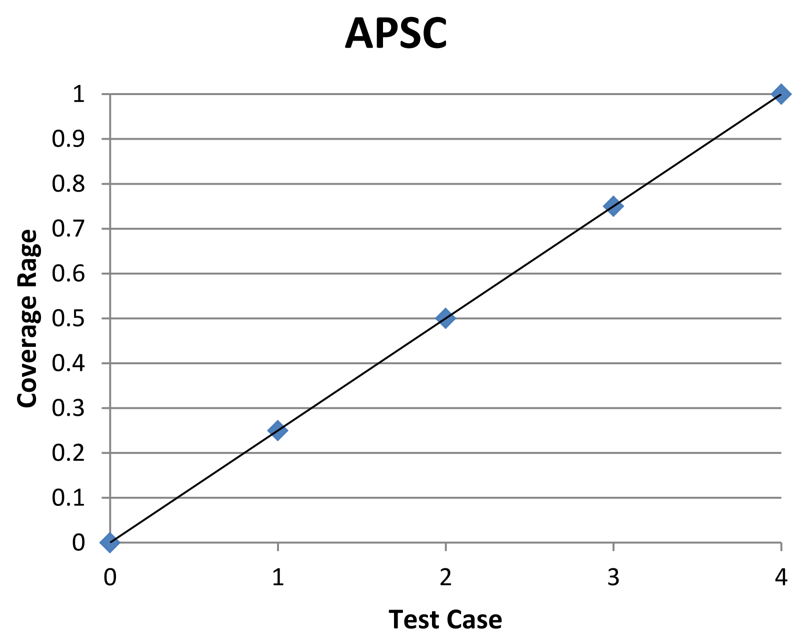
An example

Code lines ***Coverage matrix:***

Code lines	TC1	TC2	TC3	TC4
0	1	0	0	1
1	0	1	0	1
2	0	0	1	1
3	0	0	0	1

Execution sequence:

TC1 – TC2 – TC3 – TC4



$$1-\frac{1+2+3+4}{4*4}+\frac{1}{2*4}=0.5$$

Prioritization target

An example

Code lines ***Coverage matrix:***

Code lines	TC1	TC2	TC3	TC4
0	1	0	0	1
1	0	1	0	1
2	0	0	1	1
3	0	0	0	1

Execution sequence:

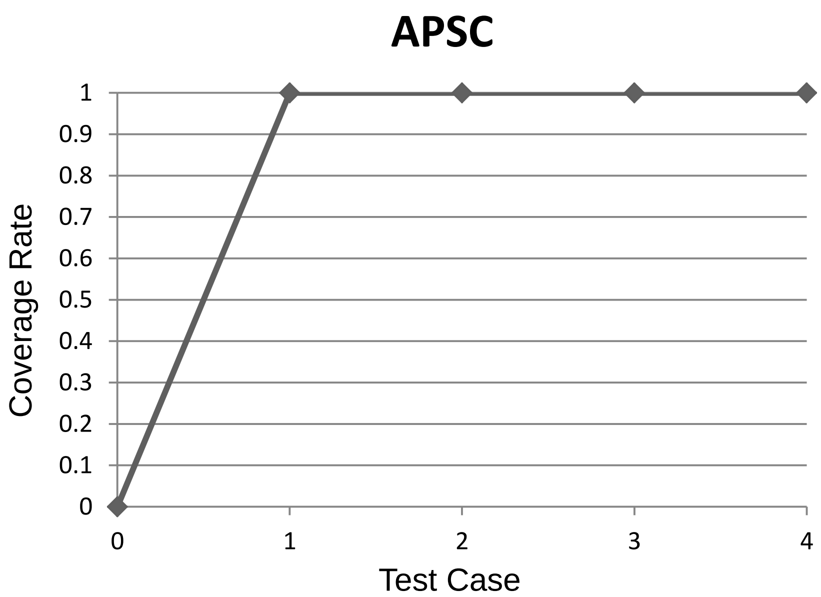
Fitness function

Prioritization target

An example

Code lines **_Coverage matrix:_**

Code lines	TC1	TC2	TC3	TC4
0	1	0	0	1
1	0	1	0	1
2	0	0	1	1
3	0	0	0	1



Execution sequence:

TC4 – TC1 – TC2 – TC3

TC3 – TC2 – TC1

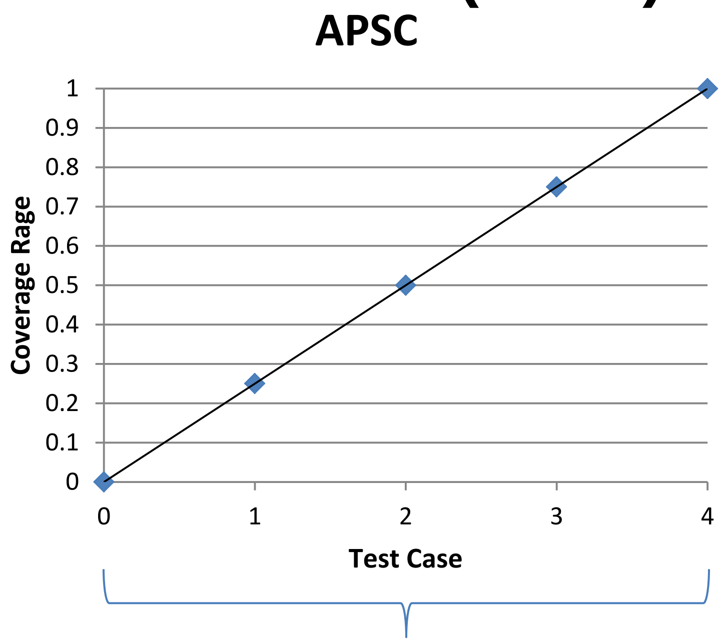
TC2 – TC1 – TC3

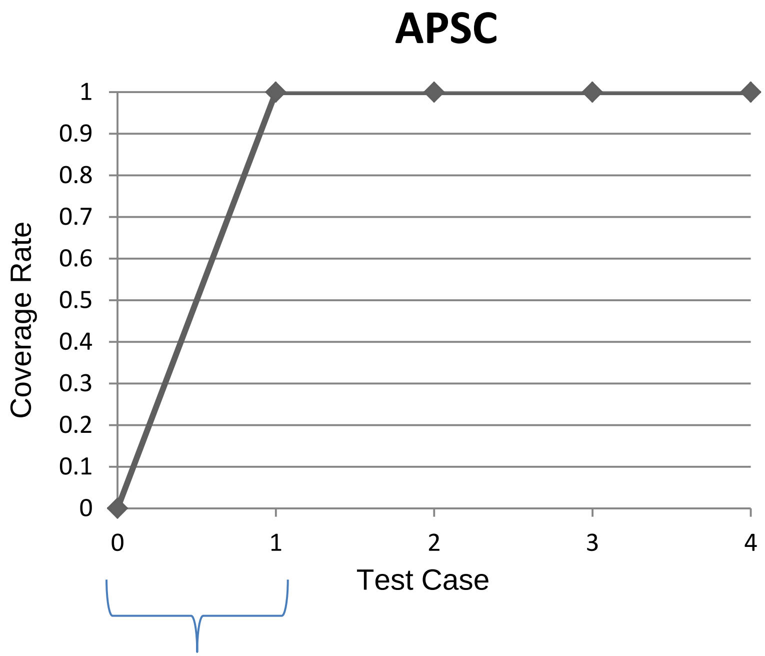
....

$$1 - \frac{1+1+1+1}{4*4} + \frac{1}{2*4} = 0.875$$

Epistatic Test case Segment (ETS)

Definition of ETS ： **(order)**





Effective part of test case sequence

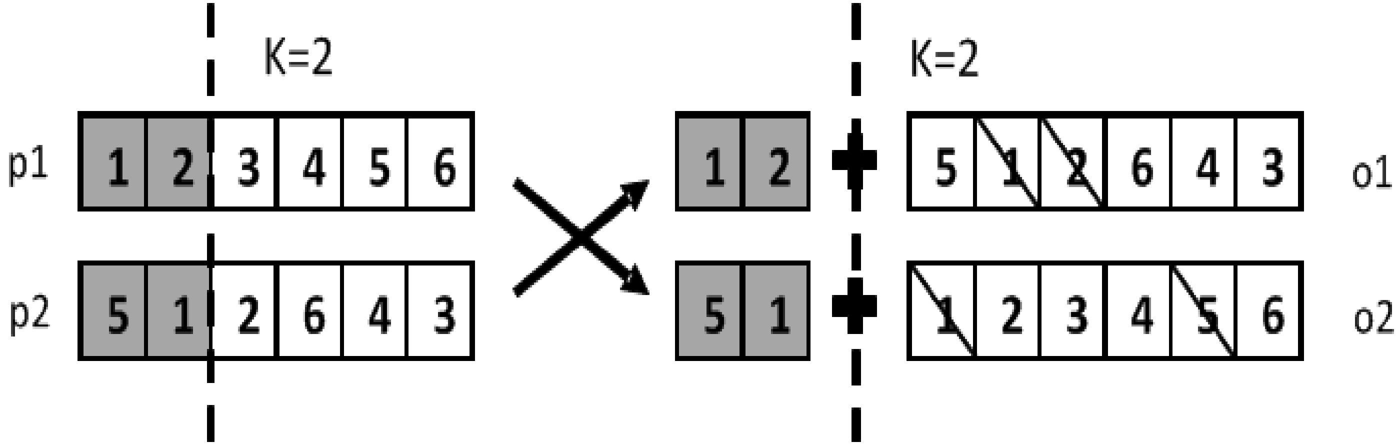
- Test case prioritization, length ETS is dynamically changed.
- In regular algorithm, crossover operation should effectively change the ETS to generate the offspring.

Epistatic in Test Case Priortization

Crossover operations

In test case prioritization ：

For **one-point** crossover



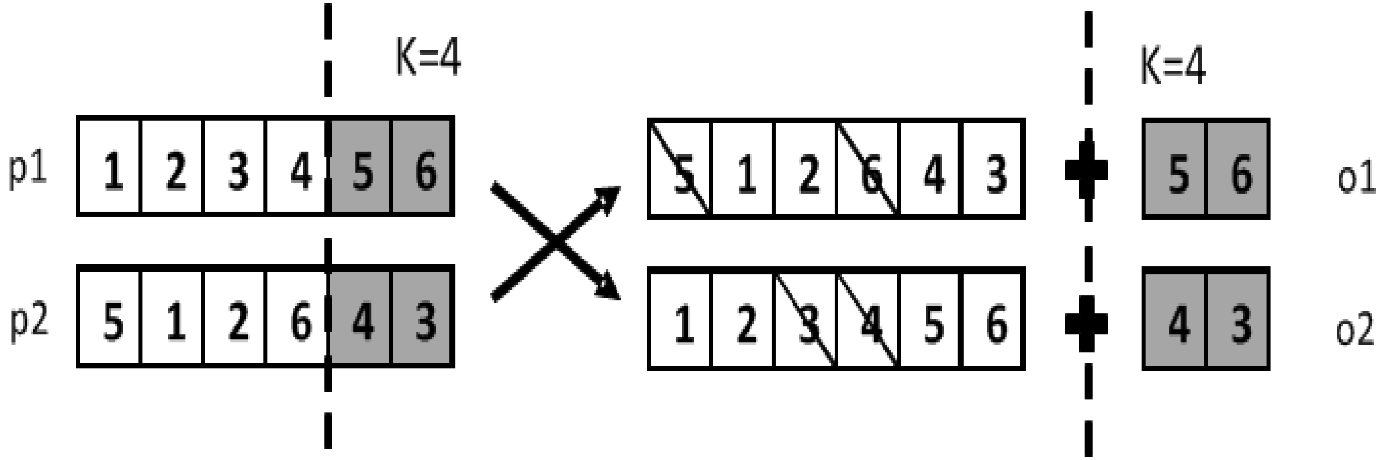
For single-point crossover, the front gen position can seldom change.

The change point may behind the ETS gene.

Crossover operations

In test case prioritization ：

For **one-point** crossover



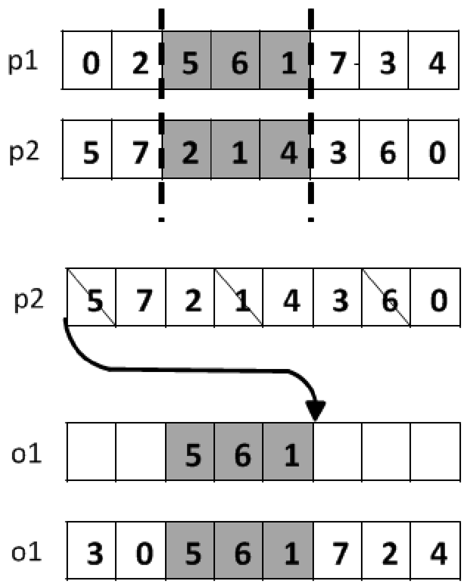
The possibility of changing ETS becomes higher.

The genes in ETS may change more frequently that for some redundancy gens move to outside of ETS.

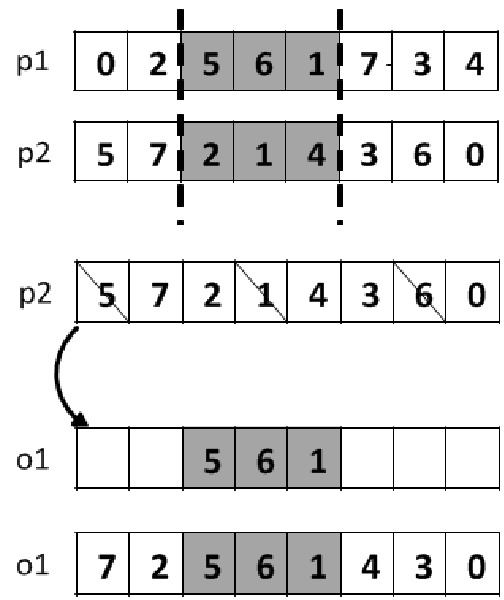
Crossover operations

In test case prioritization ：

For **order** crossover



Ord



E-Ord

To conclude which kind of crossover is more likely to inherit good genes within ETS from parent.

Crossover operations

In test case prioritization ：

Other permutation encoding crossover, about partially mapped crossover (PMX)

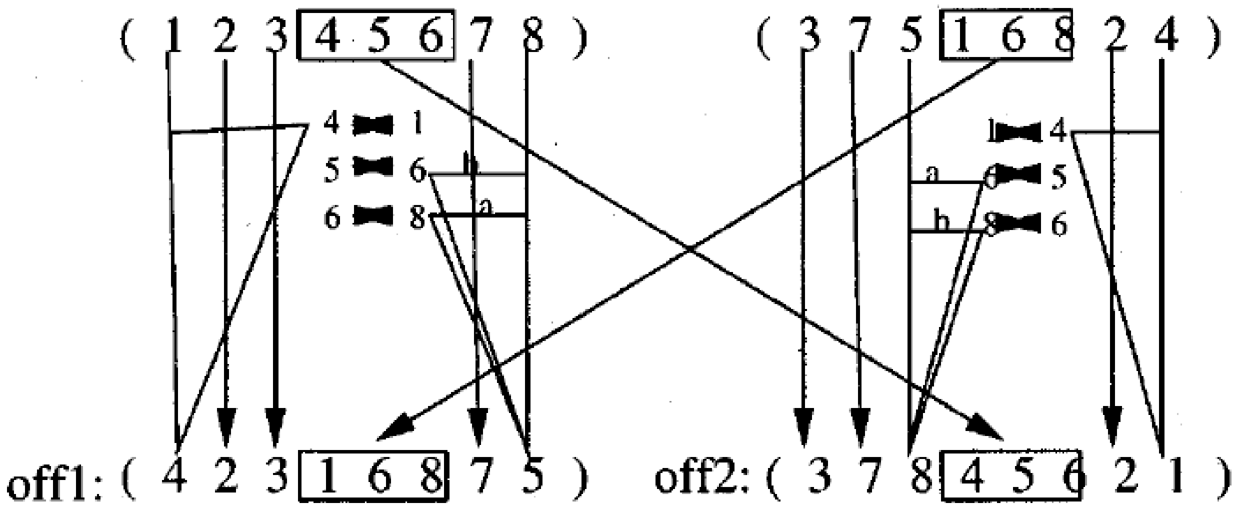


Figure 2. Partially-mapped crossover operator (PMX).

*Larrañaga P, Kuijpers C M H, Murga R H, et al. Genetic algorithms for the travelling salesman problem: A review of representations and operators[J]. Artificial Intelligence Review, 1999, 13(2): 129-170.

Epistatic in Test Case Priortization

Experiments

Subjects under test:

Subject	SLOC	test suite size		
		min	max	average
flex	3016	1047	1470	1350.17
space	3815	1208	3229	1894.29
bash	6181	764	1467	1063.17
v8	59412	2564	6159	3909.15

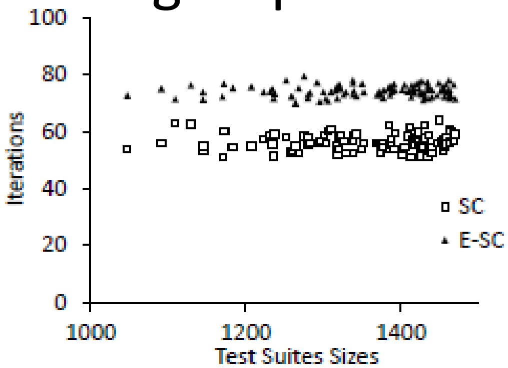
Three different group of comparison:

Experiment	Setup
1A	SC vs E-SC with termination A
1B	SC vs E-SC with termination B
2A	Ord vs E-Ord with termination A
2B	Ord vs E-Ord with termination B

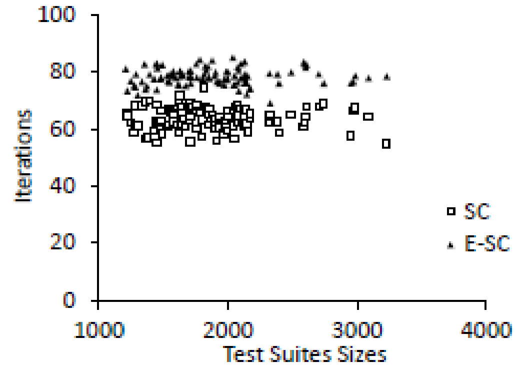
Termination condition A: APSE value is less than e-5 in continuous 10 generations.
Termination condition B: Iterate 300 generation.

Experiments

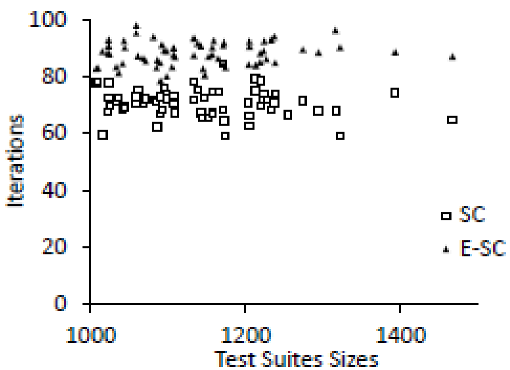
In comparison group 1:



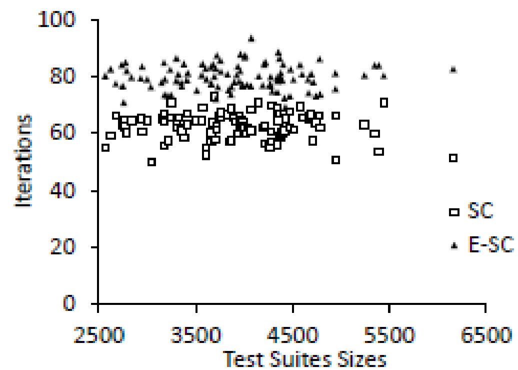
(a) flex



(b) space



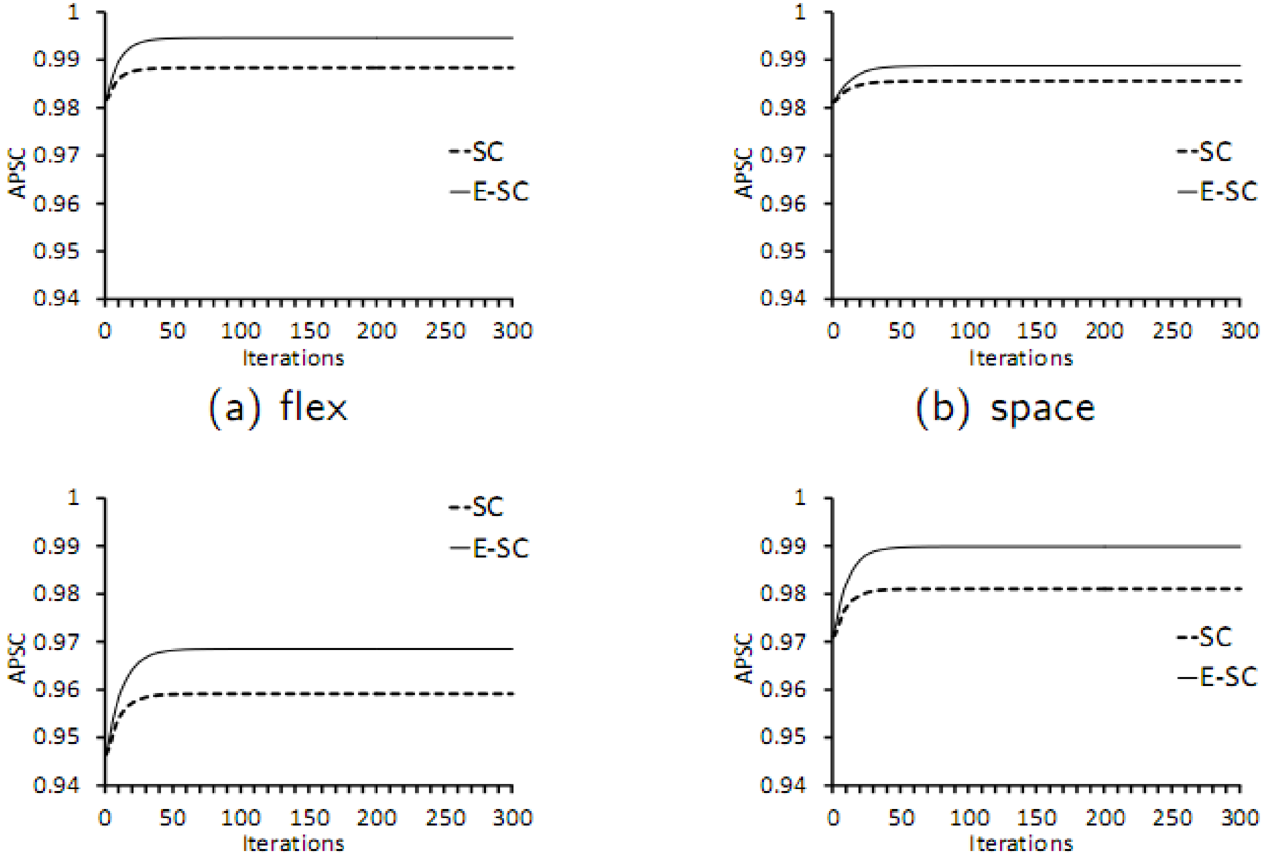
(c) bash



(d) v8

Experiments

In comparison group 1:

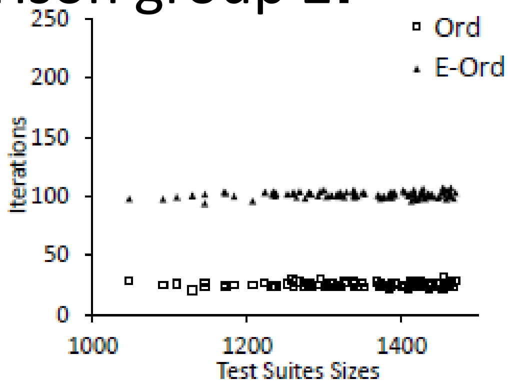


(a) flex (b) space (c) bash (d) v8

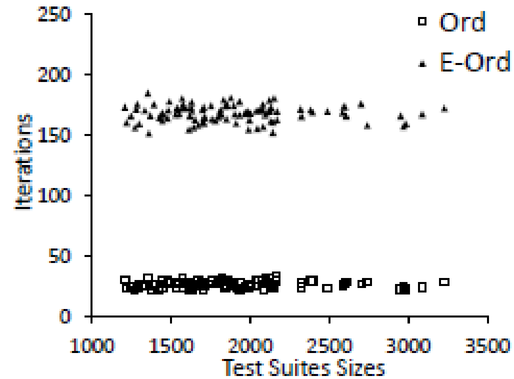
Epistatic in Test Case Priortization

Experiments

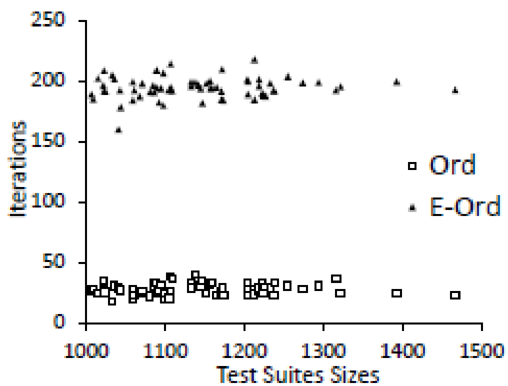
In comparison group 2:



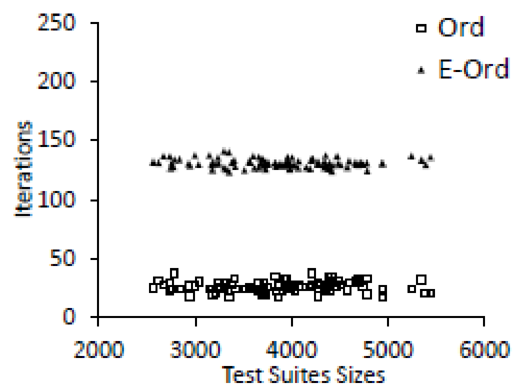
(a) flex



(b) space



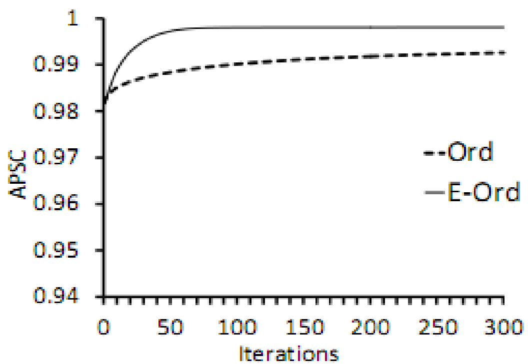
(c) bash



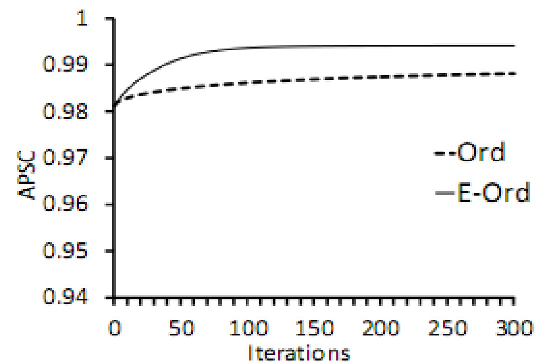
(d) v8

Experiments

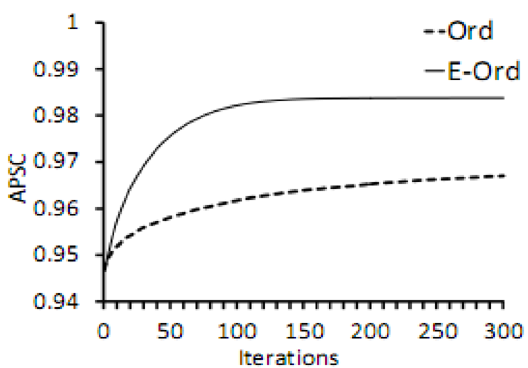
In comparison group 2:



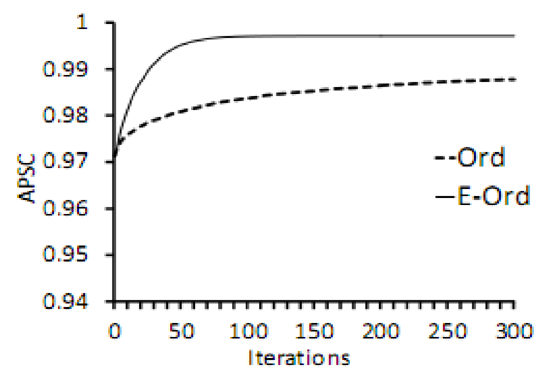
(a) flex



(b) space



(c) bash



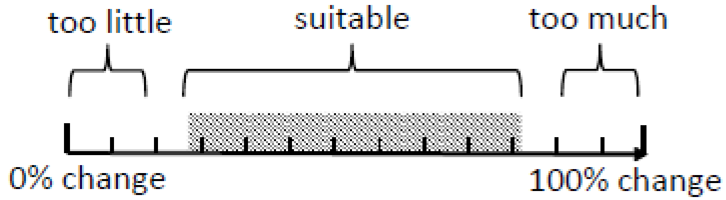
(d) v8

Experiments & Conclusion

In comparison of PMX with E-Ord:

Test Case	PMX			E-Ord		
	avg iters	avg APSC	variation	avg iters	avg APSC	variation
flex	157.04	0.9967	1.93E-06	101.08	0.9980	3.32E-08
space	187.93	0.9929	3.67E-06	167.78	0.9940	2.06E-07
bash	229.54	0.9801	7.10E-06	194.34	0.9836	1.34E-06
v8	226.88	0.9964	4.08E-06	130.64	0.9971	1.90E-07

Conclusion:

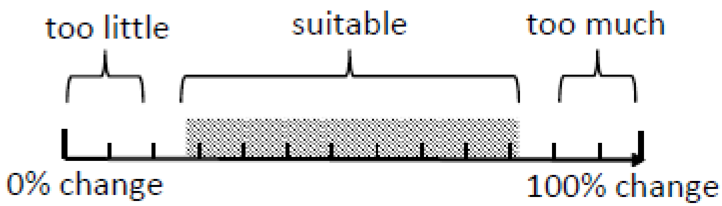


- In ETS, too large variation will cause the offspring loss of good characteristic of ETS identified by fitness function (Ord vs. E-Ord).
- But little inheritance will make GA not produce offspring from good ETSes and almost degenerate back to a random search (SC vs. E-SC).

Epistatic in Test Case Priortization

Future work

Epistatic domain is very common in SBSE area.

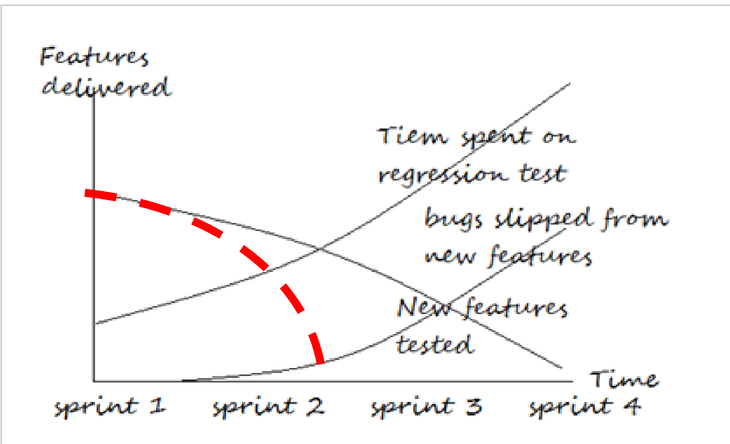


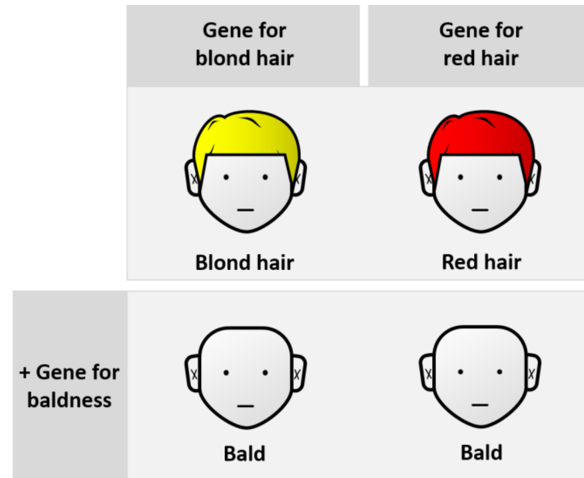
Calculate the accurate value of changing in ETS for test case prioritization.

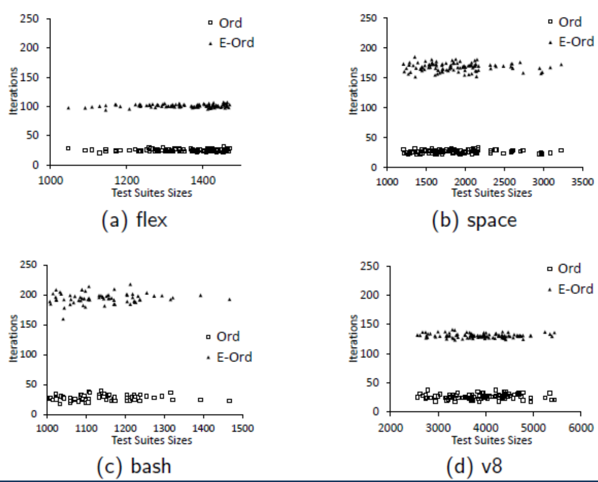
Code lines	*Coverage matrix:* tc1	tc2	tc3	tc4
0	1	0	0	1
1	0	1	0	1
2	0	0	1	1
3	0	1	0	1
Execution time:	0.1	0.2	0.1	1.0

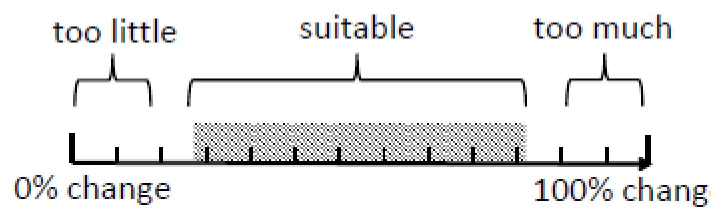
Finding out the important test cases in the opportune gen positions more quick, so need to check the change process of ETS in the prioritization process.

Summary


Features delivered
Tiem spent on regression test
bugs slipped from new features
New features tested
Time
sprint 1
sprint 2
sprint 3
sprint 4


Gene for blond hair
Gene for red hair
Blond hair
Red hair
+ Gene for baldness
Bald
Bald


Ord
E-Ord
Iterations
Test Suites Sizes
(a) flex
(b) space
(c) bash
(d) v8


too little
suitable
too much
0% change
100% change

Acknowledge

Fang Yuan

Zheng Li

Ruilian Zhao

Thanks!

Q & A