Human Competitiveness of Genetic Programming in Spectrum Based Fault Localization

谢晓园
武汉大学计算机学院
xxie@whu.edu.cn

缘起。。。

1. 研究背景与动机

- 现代社会，软件无处不在

1. 研究背景与动机

- 高质量软件 – 迫切要求
- 关键系统的软件缺陷将造成灾难性损失
  - 人民生命威胁，巨额财产损失，社会秩序破坏

2003.8.14 美国
美国东北部大面积停电
分布式软件缺陷

2005.11.1 日本
东京股市"停摆"
软件系统升级缺陷

2011.8.18 俄罗斯
质子火箭发射失败
推进器控制软件缺陷

2011.7.23 中国
甬温线动车事故
列控软件缺陷

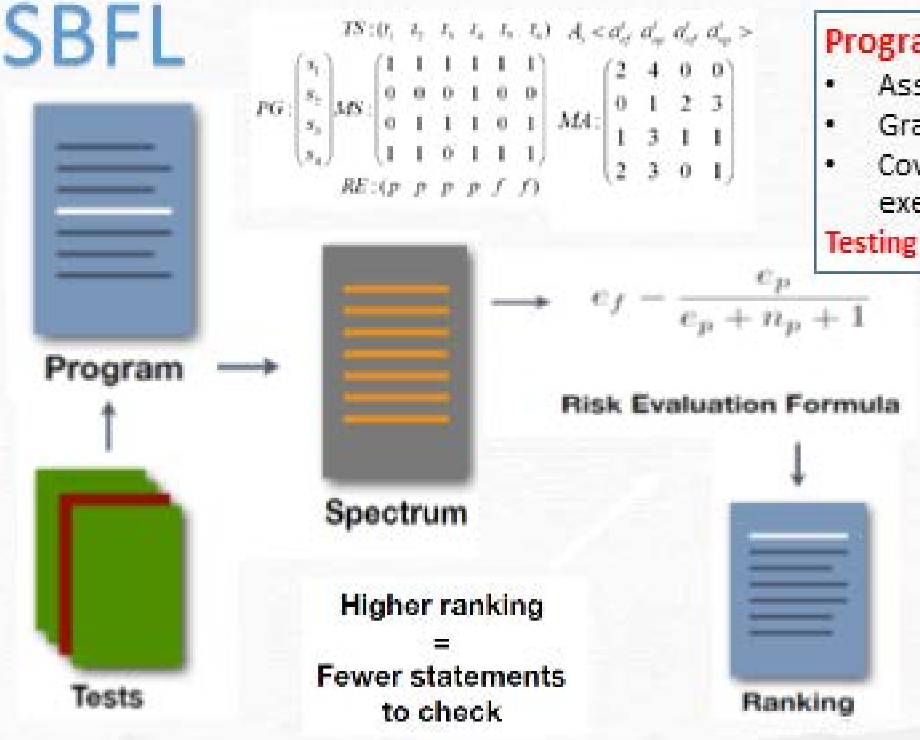
1. 研究背景与动机

- 软件缺陷的产生不可避免
  - 缺陷定位贯穿整个软件生命周期
- 常用方法 – 人工缺陷定位
  - 费时费力，效果不稳定
- 自动缺陷定位
  - 受到广泛关注
  - 存在关键问题，亟待解决

Spectrum-based Fault Localization -SBFL

1. 研究背景与动机

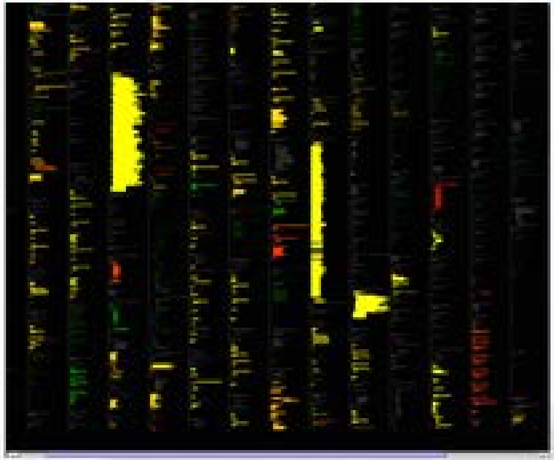
SBFL



Program spectrum: run-time profile

- Associated with each testing result
- Granularity: statements, branches, etc.
- Coverage information: binary (execution slice), execution trace, etc.

Testing result for the test cases:



1. 研究背景与动机

- Program spectrum 程序频谱: run-time profile
  - Associated with each testing result
  - Granularity: statements, branches, etc.
  - Coverage information: binary, execution trace, etc.

$$
\begin{array}{ccc}
 & & TS:(t_1 \quad t_2 \quad \ldots \quad t_m) \\
PG:\begin{pmatrix} s_1 \\ s_2 \\ \cdot \\ \cdot \\ \cdot \\ s_n \end{pmatrix} & MS: & \begin{pmatrix} 1/0 & 1/0 & \ldots & 1/0 \\ 1/0 & 1/0 & \ldots & 1/0 \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ 1/0 & 1/0 & \ldots & 1/0 \end{pmatrix} \\
 & & RE:(p/f \quad p/f \quad \ldots \quad p/f)
\end{array}
$$

1. 研究背景与动机

- 为每一个程序节点 s_i 可得 $A_i =< e^i_f, e^i_p, n^i_f, n^i_p >$

  - *'e'* : 'executed'
  - *'n'* : 'not executed'
  - *'f'* : 'fail'
  - *'p'* : 'pass'

$$TS:(t_1 \quad t_2 \quad t_3 \quad t_4 \quad t_5 \quad t_6)$$

$$PG:\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} MS:\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

$$RE:(p \quad p \quad p \quad p \quad f \quad f)$$

$$A_i < a^i_{ef} \quad a^i_{ep} \quad a^i_{nf} \quad a^i_{np} >$$

$$MA:\begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 1 & 1 \\ 2 & 3 & 0 & 1 \end{pmatrix}$$

1. 研究背景与动机

- 风险评估公式（risk evaluation formula R）
  - mapping A_i into the risk value of statement s_i
  - SBFL最核心的组成部分

$$\text{Tarantula} = \frac{\frac{e_f}{e_f+n_f}}{\frac{e_p}{e_p+n_p} + \frac{e_f}{e_f+n_f}}.$$

- 所有语句风险值降序排列，依次排查程序缺陷

1. 研究背景与动机

- SBFL示例：

Structural Elements	Test t_1	Test t_2	Test t_3	e_p	e_f	n_p	n_f	Tarantula	Rank
s_1	•			1	0	0	2	0.00	9
s_2	•			1	0	0	2	0.00	9
s_3	•			1	0	0	2	0.00	9
s_4	•			1	0	0	2	0.00	9
s_5	•			1	0	0	2	0.00	9
s_6	•		•	1	1	0	1	0.33	4
s_7 (faulty)		•	•	0	2	1	0	1.00	1
s_8	•	•		1	1	0	1	0.33	4
s_9	•	•	•	1	2	0	0	0.50	2
Result	P	F	F						

1. 研究背景与动机

- 风险评估公式的度量指标
  - Effectiveness measurement: Expense metric
    - 须被检查语句的比例
    - The lower, the better

$s_{i1}, s_{i2}, \ldots \ldots, s_{ik}\ s_f\ s_{k+2}, \ldots \ldots\ s_n$

$Expense = (k+1)/n$

1. 研究背景与动机

SBFL重要研究内容：设计高效的风险评估公式

$$
\begin{array}{ccc}
 & & TS:(t_1 \quad t_2 \quad \cdots \quad t_n) \\
PG: \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_m \end{pmatrix} & MS: & \begin{pmatrix} 1/0 & 1/0 & \cdots & 1/0 \\ 1/0 & 1/0 & \cdots & 1/0 \\ & \vdots & & \\ 1/0 & 1/0 & \cdots & 1/0 \end{pmatrix} \\
 & & RE:(p/f \quad p/f \quad \cdots \quad p/f)
\end{array}
$$

R：决定定位效果的关键

2. 人工风险评估公式的设计

- 第一个公式：Tarantula in 2002

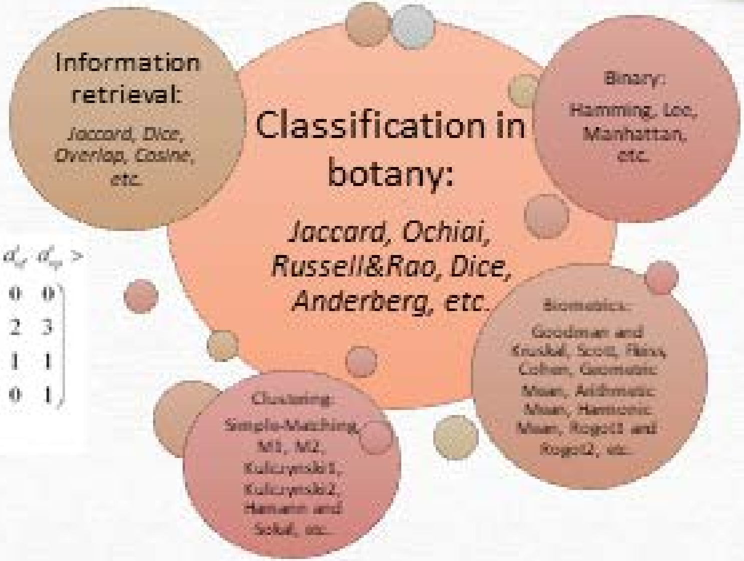
$$\text{Tarantula} = \frac{\frac{e_f}{e_f+n_f}}{\frac{e_p}{e_p+n_p} + \frac{e_f}{e_f+n_f}}.$$

JONES, J. A.,HARROLD, M. J., AND STASKO, J. 2002. Visualization of test information to assist fault localization. In *Proceedings of the 24th International Conference on Software Engineering*. 467–477.

2. 人工风险评估公式的设计

- 手工设计思想: based on various intuitions

$$
\begin{array}{c}
TS:(t_1\ \ t_2\ \ t_3\ \ t_4\ \ t_5\ \ t_6) \\
PG:\begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} MS:\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 & 1 \end{pmatrix} \\
RE:(p\ \ p\ \ p\ \ p\ \ f\ \ f)
\end{array}
\qquad
\begin{array}{c}
A_i < d^i_{ef}\ d^i_{ep}\ d^i_{nf}\ d^i_{np} > \\
MA:\begin{pmatrix} 2 & 4 & 0 & 0 \\ 0 & 1 & 2 & 3 \\ 1 & 3 & 1 & 1 \\ 2 & 3 & 0 & 1 \end{pmatrix}
\end{array}
$$



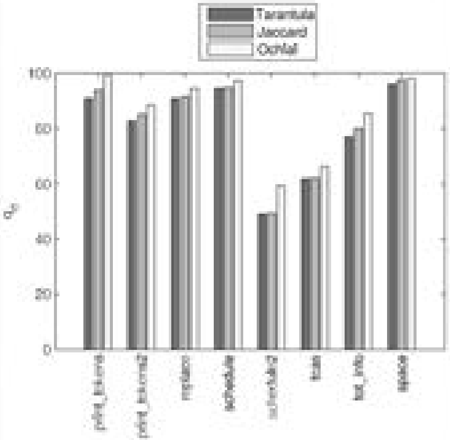
群雄割据。。。

Name	Formula	Name	Formula
O	-1 if $a_{nf} > 0$, otherwise a_{np}	O^p	$a_{ef} - \frac{a_{ep}}{totP+1}$
Jaccard	$\frac{a_{ef}}{totF + a_{ep}}$	Ochiai	$\frac{a_{ef}}{\sqrt{totF(a_{ef}+a_{ep})}}$
Tarantula	$\frac{\frac{a_{ef}}{totF}}{\frac{a_{ef}}{totF}+\frac{a_{ep}}{totP}}$		
Zoltar	$\frac{a_{ef}}{totF + a_{ep} + \frac{10000 a_{nf} a_{ep}}{a_{ef}}}$		
Ample	$\left\lvert\frac{a_{ef}}{totF} - \frac{a_{ep}}{totP}\right\rvert$	Ample2	$\frac{a_{ef}}{totF} - \frac{a_{ep}}{totP}$
Wong1	a_{ef}	Wong2	$a_{ef} - a_{ep}$
Wong3	$a_{ef} - h$, where $h = \begin{cases} a_{ep} & \text{if } a_{ep} \le 2 \\ 2 + 0.1(a_{ep} - 2) & \text{if } 2 < a_{ep} \le 10 \\ 2.8 + 0.001(a_{ep} - 10) & \text{if } a_{ep} > 10 \end{cases}$		
Wong3'	$a_{ef} - h$, where $h = \begin{cases} -1000 & \text{if } a_{ep} + a_{ef} = 0 \\ Wong3 & \text{otherwise} \end{cases}$		
CBI Inc	$\frac{a_{ef}}{a_{ef}+a_{ep}} - \frac{totF}{T}$	CBI Log	$\frac{2}{\frac{1}{CBIInc} + \frac{\log totF}{\log a_{ef}}}$
CBI Sqrt	$\frac{2}{\frac{1}{CBIInc} + \frac{\sqrt{totF}}{\sqrt{a_{ef}}}}$	M1	$\frac{a_{ef}+a_{np}}{a_{nf}+a_{ep}}$
M2	$\frac{a_{ef}}{a_{ef}+a_{np}+2(a_{nf}+a_{ep})}$	M3	$\frac{2(a_{ef}+a_{np})}{T}$
Sørensen-Dice	$\frac{2a_{ef}}{2a_{ef}+a_{nf}+a_{ep}}$		
Kulczynski1	$\frac{a_{ef}}{a_{nf}+a_{ep}}$		
Kulczynski2	$\frac{1}{2}\left(\frac{a_{ef}}{totF}+\frac{a_{ef}}{a_{ef}+a_{ep}}\right)$		
Russell and Rao	$\frac{a_{ef}}{T}$	Lee	$a_{ef} + a_{np}$
Rogers & Tanimoto	$\frac{a_{ef}+a_{np}}{a_{ef}+a_{np}+2(a_{nf}+a_{ep})}$	Goodman	$\frac{2a_{ef}-a_{nf}-a_{ep}}{2a_{ef}+a_{nf}+a_{ep}}$
Simple Matching	$\frac{a_{ef}+a_{np}}{T}$	Hamann	$\frac{a_{ef}+a_{np}-a_{nf}-a_{ep}}{T}$
Hamming	$a_{ef} + a_{np}$	Euclid	$\sqrt{a_{ef} + a_{np}}$
Ochiai2	$\frac{a_{ef}a_{np}}{\sqrt{(a_{ef}+a_{ep})(a_{np}+a_{nf})(totF)(totP)}}$		
Ochiai3	$\frac{a_{ef}^2}{totF(a_{ef}+a_{ep})}$		
Piatetsky-Shapiro	$a_{ef} + a_{ef}^2 + totF - a_{ep}a_{nf} - a_{np}a_{ef}$		
Collective Strength	$1 - \frac{a_{ef}+a_{np}}{(a_{ef}+a_{ep})(totF)+(a_{nf}+a_{np})(totP)} \cdot \frac{1-(a_{ef}+a_{ep})(totF)-(a_{nf}+a_{np})(totP)}{1-a_{ef}-a_{np}}$		
Geometric Mean	$\frac{a_{ef}a_{np}-a_{nf}a_{ep}}{\sqrt{(a_{ef}+a_{ep})(a_{np}+a_{nf})(totF)(totP)}}$		
Harmonic Mean	$\frac{(a_{ef}a_{np}-a_{nf}a_{ep})((a_{ef}+a_{ep})(a_{np}+a_{nf})+(totF)(totP))}{(a_{ef}+a_{ep})(a_{np}+a_{nf})(totF)(totP)}$		

Name	Formula	Name	Formula
Arithmetic Mean	$\frac{2a_{ef}a_{np}-2a_{nf}a_{ep}}{(a_{ef}+a_{ep})(a_{np}+a_{nf})+(totF)(totP)}$		
Cohen	$\frac{2a_{ef}a_{np}-2a_{nf}a_{ep}}{(a_{ef}+a_{ep})(totP)+(a_{nf}+a_{np})(totF)}$		
Scott	$\frac{4a_{ef}a_{np}-4a_{nf}a_{ep}-(a_{nf}-a_{ep})^2}{(2a_{ef}+a_{nf}+a_{ep})(2a_{np}+a_{nf}+a_{ep})}$		
Fleiss	$\frac{4a_{ef}a_{np}-4a_{nf}a_{ep}-(a_{nf}-a_{ep})^2}{(2a_{ef}+a_{nf}+a_{ep})+(2a_{np}+a_{nf}+a_{ep})}$		
Rogot1	$\frac{1}{2}\left(\frac{a_{ef}}{2a_{ef}+a_{nf}+a_{ep}}+\frac{a_{np}}{2a_{np}+a_{nf}+a_{ep}}\right)$		
Rogot2	$\frac{1}{4}\left(\frac{a_{ef}}{a_{ef}+a_{ep}}+\frac{a_{ef}}{totF}+\frac{a_{np}}{totP}+\frac{a_{np}}{a_{np}+a_{nf}}\right)$		
Binary	0 if $a_{nf} > 0$, otherwise 1	Gower1	$\frac{a_{ef}-(a_{nf}+a_{ep})+a_{np}}{T}$
Gower2	$\frac{a_{ef}+a_{np}}{a_{ef}+a_{np}+0.5(a_{nf}+a_{ep})}$	Gower3	$\frac{a_{ef}a_{np}-a_{nf}a_{ep}}{a_{ef}a_{np}+a_{nf}a_{ep}}$
Anderberg	$\frac{a_{ef}}{a_{ef}+2(a_{nf}+a_{ep})}$		
AddedValue	$\frac{a_{ef}}{\max(a_{ef}+a_{ep}, totF)}$	Interest	$\frac{a_{ef}}{(a_{ef}+a_{ep})totF}$
Confidence	$\max\left(\frac{a_{ef}}{a_{ef}+a_{ep}}, \frac{a_{ef}}{totF}\right)$		
Certainty	$\max\left(\frac{a_{ef}}{a_{ef}+a_{ep}} - (a_{ef}+a_{ep}), 1-(a_{ef}+a_{ep})\right)$		
Sneath & Sokal 1	$\frac{2(a_{ef}+a_{np})}{T+a_{ef}+a_{np}}$		
Sneath & Sokal 2	$\frac{a_{ef}}{a_{ef}+2(a_{nf}+a_{ep})}$		
Phi	$\frac{a_{ef}a_{np}-a_{nf}a_{ep}}{\sqrt{(totF)(a_{ef}+a_{ep})(a_{nf}+a_{np})(totP)}}$		
Kappa	$1 - \frac{a_{ef}+a_{np}-(a_{ef}+a_{ep})(totF)-(a_{nf}+a_{np})(totP)}{1-(a_{ef}+a_{ep})(totF)-(a_{nf}+a_{np})(totP)}$		
Conviction	$\max\left(\frac{(a_{ef}+a_{ep})(totP)}{a_{ep}}, \frac{(a_{nf}+a_{np})(totF)}{a_{nf}}\right)$		
Mountford	$\frac{2a_{ef}}{2(a_{ef}+a_{ep})(totF)-(2a_{ef}+a_{ep}+a_{nf})a_{ef}}$		
Klosgen	$\sqrt{a_{ef}} \cdot \max\left(\frac{a_{ef}}{a_{ef}+a_{ep}} - totF, \frac{a_{ef}}{totF} - (a_{ep}+a_{ef})\right)$		
YuleQ	$\frac{a_{ef}a_{np}-a_{nf}a_{ep}}{a_{ef}a_{np}+a_{nf}a_{ep}}$	YuleY	$\frac{\sqrt{a_{ef}a_{np}}-\sqrt{a_{nf}a_{ep}}}{\sqrt{a_{ef}a_{np}}+\sqrt{a_{nf}a_{ep}}}$
YuleV	$\frac{a_{ef}a_{np}-(a_{ef}+a_{ep})(totF)}{a_{ef}a_{np}+(a_{ef}+a_{ep})(totF)}$		
Correlation	$\frac{\lvert a_{ef}a_{np}-(a_{ef}+a_{ep})(totF)\rvert - \frac{T}{2}}{\sqrt{(a_{ef}+a_{ep})(totF)\,a_{np}(totF)\,a_{ef}(totP)\,a_{nf}}}$		
Manhattan	$1 - \frac{a_{ep}+a_{nf}}{T}$	Braun	$\frac{a_{ef}}{a_{ef}+a_{nf}}$
Baroni	$\frac{\sqrt{a_{ef}a_{np}}+a_{ef}}{\sqrt{a_{ef}a_{np}}+totF+a_{ep}}$	Coef	$\frac{a_{ef}}{a_{ef}+a_{ep}}$
Levandowsky	$\frac{a_{ef}}{totF+a_{ep}}$	Watson	$1 - \frac{a_{nf}+a_{ep}}{2a_{ef}+a_{nf}+a_{ep}}$
JacCube	$\frac{a_{ef}}{\sqrt[3]{totF+a_{ep}}}$	NFD	$a_{ef} + a_{np}$
SokalDist	$\sqrt{\frac{a_{nf}+a_{ep}}{T}}$	Overlap	$\frac{a_{ef}}{\min(a_{ef}, a_{nf}, a_{ep})}$
CorRatio	$\frac{a_{ef}^2}{totF(a_{ef}+a_{ep})}$	Forbes	$\frac{Ta_{ef}}{totF(a_{ef}+a_{ep})}$
Fager	$\frac{a_{ef}}{\sqrt{totF(a_{ef}+a_{ep})}} - \frac{1}{2\sqrt{a_{ef}+a_{ep}}}$		

3. 人工公式的比较

- 实验比较



Metric	Tcas	Sch	Sch2	Pt	Pt2	Tot	Rep
O^p	9.90	3.88	23.21	3.40	0.78	3.11	4.68
O	9.90	7.65	23.21	3.40	0.78	3.11	4.68
Wong3'	10.11	3.88	17.35	3.40	0.82	3.11	3.95
Wong3	17.78	17.71	29.10	3.40	0.82	9.27	11.29
Wong4	10.03	4.74	20.77	3.55	0.78	3.18	4.68
Zoltar	9.90	3.88	20.77	3.40	0.78	3.13	3.92
M2	10.27	1.57	23.17	4.31	2.00	4.59	4.49
Kulczynski2	9.94	3.88	20.80	3.40	1.08	3.27	4.28
Overlap	14.47	11.36	18.51	7.15	10.65	6.63	9.11
Ochiai	10.66	1.63	23.56	6.22	3.70	5.52	4.90
Amean	10.82	5.14	23.89	7.31	5.09	9.37	5.24
Jaccard	10.77	1.68	26.04	8.37	5.55	6.53	6.25
Tarantula	10.80	1.77	26.04	8.93	5.70	7.09	6.45
Russell	14.47	11.36	18.51	7.15	10.65	6.60	9.10
Binary	14.47	15.15	18.51	7.15	10.65	6.60	9.10
Ample	12.83	12.62	27.50	8.32	5.47	15.09	6.92
Ample2	10.91	4.98	28.42	7.71	4.94	9.38	5.72
Pearson	10.92	4.95	25.13	7.84	4.92	9.31	5.05
McCon	9.94	10.94	20.80	3.40	1.08	3.27	4.28
CBI Log	11.47	7.20	26.04	8.88	5.70	10.26	6.25
JacCube	10.35	2.08	22.91	4.57	2.00	4.53	4.58
Rogot2	10.91	10.83	27.76	7.74	4.92	14.37	5.80

Table X. Description of Siemens Test Suite and Space

Program	Versions	LOC	Test Cases	Description
tcas	41	173	1608	altitude separation
schedule	9	410	2650	priority scheduler
schedule2	10	307	2710	priority scheduler
print_tokens	7	563	4130	lexical analyser
print_tokens2	10	508	4115	lexical analyser
tot_info	23	406	1052	information measure
replace	32	563	5542	pattern recognition
Space	32	9059	13585	array definition language

ABREU, R.et al. 2009. A practical evaluation of spectrum-based fault localization. *J. Syst. Softw. 82*, 11, 1780–1792.

3. 人工公式的比较

- 基于ITE2的模型分析

```
if (t1())
    s1();          /* S1 */
else
    s2();          /* S2 */
if (t2())
    x = True;      /* S3 */
else
    x = t3();      /* S4 - BUG */
```

Fig. 1. Program segment If-Then-Else-2 (ITE2).

Optimal against ITE2

$$O(a_{np}, a_{nf}, a_{ep}, a_{ef}) = \begin{cases} -1 & \text{if } a_{nf} > 0 \\ a_{np} & \text{otherwise} \end{cases}$$

$$O^{p}(a_{np}, a_{nf}, a_{ep}, a_{ef}) = a_{ef} - \frac{a_{ep}}{P+1}$$

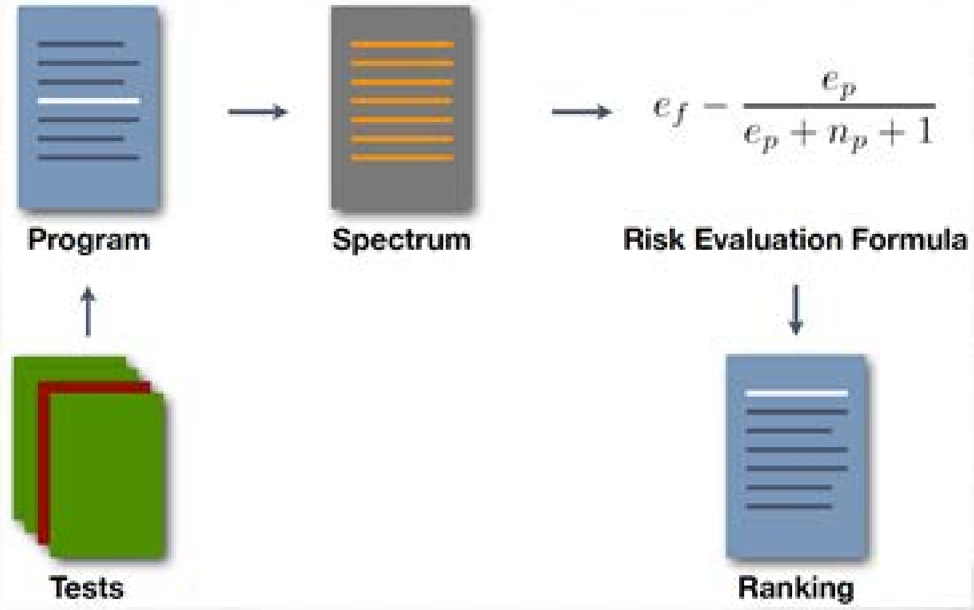
NAISH, Lee et al., A model for spectra-based software diagnosis. *ACM Trans. Softw. Engin. Methodol. 20*, 3, 11:1–11:32.

Table IV. Influence of Test Suite Size and Metric on Total Score (%) for $ITE2_8$

Num. of tests	2	5	10	20	50	100	500	1000
O, O^p	60.00	72.59	81.10	87.98	94.13	96.81	99.31	99.65
Wong3′	60.00	67.04	76.51	86.87	94.06	96.81	99.31	99.65
Wong3	56.67	63.15	75.59	86.79	94.06	96.81	99.31	99.65
Zoltar	60.00	71.48	79.97	87.53	94.06	96.80	99.31	99.65
M2	60.00	72.59	80.98	87.34	92.41	94.33	95.87	96.05
Kulczynski2	60.00	71.48	79.72	86.30	91.78	93.96	95.78	95.99
Russell etc	53.33	61.11	69.57	78.79	88.89	93.80	98.64	99.31
Overlap	48.89	54.88	66.57	77.91	88.79	93.79	98.64	99.31
Ochiai	60.00	71.48	79.12	84.90	89.51	91.28	92.75	92.93
Rogot2	56.67	67.78	77.10	83.50	88.42	90.28	91.81	92.00
HMean	48.33	67.13	77.06	83.50	88.42	90.28	91.81	92.00
GMean	48.33	67.13	76.92	83.33	88.23	90.07	91.59	91.78
AMean	48.33	67.13	76.75	83.18	88.01	89.86	91.39	91.58
Ample2	56.67	67.78	76.45	82.79	87.84	89.73	91.28	91.48
Jaccard etc	60.00	71.48	78.22	83.12	87.02	88.55	89.85	90.02
Ochiai2	48.33	67.13	75.13	80.93	85.24	86.89	88.30	88.47
Cohen	48.33	67.13	75.16	80.63	84.75	86.35	87.70	87.87
Tarantula etc	55.56	62.10	69.68	75.68	80.40	82.21	83.73	83.92
Fleiss	56.67	65.93	72.54	76.70	80.10	81.48	82.70	82.86
Scott etc	56.67	66.67	72.43	76.46	79.66	80.95	82.12	82.27
CBI Log	33.33	50.86	63.69	73.13	78.70	80.30	82.35	82.89
CBI Sqrt	28.89	46.73	60.71	70.99	77.16	78.69	79.63	79.73
Rogers etc	56.67	63.15	67.60	71.02	73.93	75.15	76.26	76.41
Ample	36.67	34.54	38.28	41.40	43.92	44.87	45.64	45.74

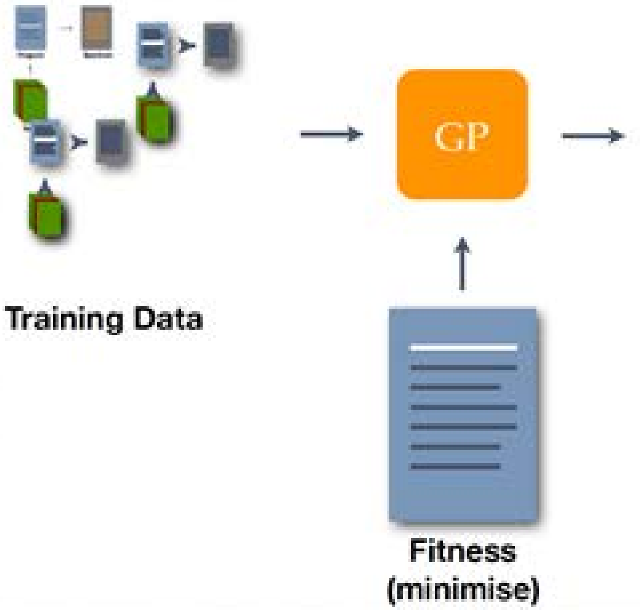
4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL



4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL



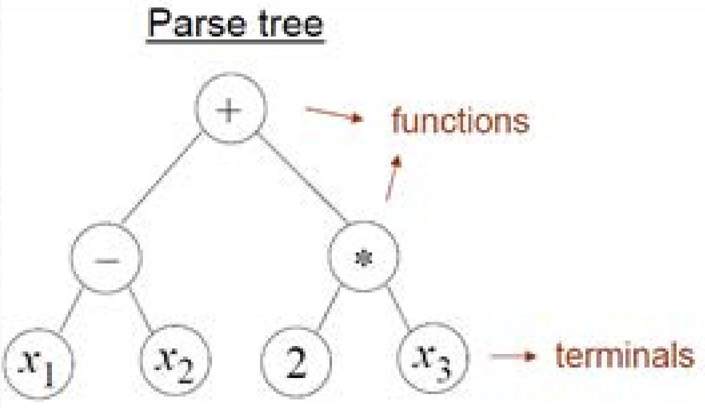
$$e_f^2(2e_p + 2e_f + 3n_p)$$

$$e_f^2(e_f^2 + \sqrt{n_p})$$

$$\ldots$$

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Tree-based Representation
  - For terminal symbols, we use $< e_f^i, e_p^i, n_f^i, n_p^i >$, as well as constants.



4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - GP operators

Operator Node	Definition
gp_add(a, b)	a + b
gp_sub(a, b)	a - b
gp_mul(a, b)	ab
gp_div(a, b)	1 if $b = 0$, $\frac{a}{b}$ otherwise
gp_sqrt(a)	$\sqrt{\lvert a \rvert}$

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Fitness function
    - To avoid over-fitting, we consider the average E over n bugs $B = \{b_1, b_2, \ldots, b_n\}$ from n programs $P = \{p_1, p_2, \ldots, p_n\}$

$$E(\tau, p, b) = \frac{\text{Ranking of } b \text{ according to } \tau}{\text{Number of statements in } p} * 100$$

$$\text{fitness}(\tau, B, P) = \frac{1}{n}\sum_{i=1}^{n} E(\tau, p_i, b_i) \text{ (to be minimised)}$$

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Configuration
    - Population size: 40;
    - Initialization of maximum tree depth: 4;
    - Stopping criterion: a fixed run of 100 generations;
    - A rank selection operator;
    - A single point crossover operator with the rate of 1.0;
    - A subtree replacement mutation operator with the rate of 0.08.

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Data for evolution

Subject	Number of Tests	Lines of Code	Executable Lines of Code	Number of Faults
flex	567	12,407–14,244	3,393–3,965	47
grep	199	12,653–13,363	3,078–3,314	11
gzip	214	6,576– 7,996	1,705–1,993	18
sed	360	8,082–11,990	1,923–2,172	16

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Protocol
    - The Genetic Programming algorithm was repeated 30 times to cater for its stochastic nature
    - Each individual run of the GP uses a random sample of 20 faults out of 92 to evolve a risk evaluation formula
    - The remaining 72 faults are reserved for evaluation purposes.

4. 首个自动化的风险公式生成方法

- Genetic Programming for SBFL
  - Result

ID	Refined Formula	ID	Refined Formula
GP01	$e_f(n_p + e_f(1 + \sqrt{e_f}))$	GP16	$\sqrt{e_f^{\frac{3}{2}} + n_p}$
GP02	$2(e_f + \sqrt{n_p}) + \sqrt{e_p}$	GP17	$\frac{2e_f + n_f}{e_f - n_p} + \frac{n_p}{\sqrt{e_f}} - e_f - e_f^2$
GP03	$\sqrt{\lvert e_f^2 - \sqrt{e_p}\rvert}$	GP18	$e_f^3 + 2n_p$
GP04	$\sqrt{\lvert \frac{n_p}{e_p - n_p} - e_f\rvert}$	GP19	$e_f\sqrt{\lvert e_p - e_f + n_f - n_p\rvert}$
GP05	$\frac{(e_f + n_p)\sqrt{e_f}}{(e_f + e_p)(n_p n_f + \sqrt{e_p})(e_p + n_p)\sqrt{\lvert e_p - n_p\rvert}}$	**GP20**	$2(e_f + \frac{n_p}{e_p + n_p})$
GP06	$e_f n_p$	GP21	$\sqrt{e_f + \sqrt{e_f + n_p}}$
GP07	$2e_f(1 + e_f + \frac{1}{2n_p}) + (1 + \sqrt{2})\sqrt{n_p}$	GP22	$e_f^2 + e_f + \sqrt{n_p}$
GP08	$e_f^2(2e_p + 2e_f + 3n_p)$	GP23	$\sqrt{e_f}(e_f^2 + \frac{n_p}{e_f} + \sqrt{n_p} + n_f + n_p)$
GP09	$\frac{e_f\sqrt{n_p}}{n_p + n_p} + n_p + e_f + e_f^3$	GP24	$e_f + \sqrt{n_p}$
GP10	$\sqrt{\lvert e_f - \frac{1}{n_p}\rvert}$	GP25	$e_f^2 + \sqrt{n_p} + \frac{\sqrt{e_f}}{\sqrt{\lvert e_p - n_p\rvert}} + \frac{n_p}{(e_f - n_p)}$
GP11	$e_f^2(e_f^2 + \sqrt{n_p})$	**GP26**	$2e_f^2 + \sqrt{n_p}$
GP12	$\sqrt{e_p + e_f + n_p - \sqrt{e_p}}$	GP27	$\frac{n_p\sqrt{(n_p n_f - e_f)}}{e_f + n_p n_f}$
GP13	$e_f(1 + \frac{1}{2e_p + e_f})$	GP28	$e_f(e_f + \sqrt{n_p} + 1)$
GP14	$e_f + \sqrt{n_p}$	GP29	$e_f(2e_f^2 + e_f + n_p) + \frac{(e_f - n_p)\sqrt{n_p e_f}}{e_p - n_p}$
GP15	$e_f + \sqrt{n_f + \sqrt{n_p}}$	GP30	$\sqrt{\lvert e_f - \frac{n_f - n_p}{e_f + n_f}\rvert}$

华山论剑。。。

5.1 实验方法比较

- Mean Expense

ID	GP	Op1	Op2	Ochiai	AMPLE	Jacc'd	Tarant.	Wong1	Wong2	Wong3
GP01	5.73	9.20	5.30	32.66	10.96	6.10	15.06	22.24	17.10	6.63
GP02	12.04	9.67	5.72	32.60	11.91	6.63	14.92	23.45	19.49	8.92
GP03	14.46	11.35	6.11	29.99	12.18	6.99	15.68	23.55	18.55	8.85
GP04	7.80	9.70	4.46	30.98	8.83	5.03	13.88	22.62	14.64	6.33
GP05	9.35	11.04	5.80	29.95	10.63	6.42	14.46	23.15	18.54	8.53
GP06	12.15	11.11	5.87	28.02	12.51	6.79	15.35	23.12	16.70	7.01
GP07	8.93	11.18	5.94	29.53	12.19	6.85	14.81	23.88	19.74	8.68
GP08	**6.32**	**10.23**	**6.34**	**30.91**	**11.67**	**7.04**	**16.21**	**23.54**	**19.94**	**9.05**
GP09	9.66	10.58	5.33	31.56	11.40	6.17	14.06	22.58	18.31	8.20
GP10	**6.31**	**11.55**	**6.31**	**29.83**	**12.51**	**7.16**	**15.79**	**22.99**	**19.74**	**8.56**
GP11	**5.83**	**11.07**	**5.83**	**33.52**	**12.12**	**6.69**	**16.77**	**22.05**	**18.16**	**6.96**
GP12	12.09	8.84	6.23	32.15	11.65	7.02	16.65	22.91	19.42	9.09
GP13	**5.11**	**9.05**	**5.11**	**31.67**	**10.27**	**5.90**	**15.92**	**22.03**	**17.00**	**6.69**
GP14	9.91	8.52	5.91	31.69	11.10	6.55	15.88	23.15	18.10	8.65
GP15	5.62	9.54	5.59	33.02	10.23	6.19	15.16	23.85	17.17	8.44
GP16	6.79	8.32	5.71	30.52	10.74	6.41	14.60	23.06	18.36	8.42
GP17	7.67	11.46	6.22	33.62	12.06	6.98	16.85	22.44	17.94	8.59
GP18	9.42	10.78	5.54	34.17	11.46	6.33	15.45	22.17	17.46	8.14
GP19	6.42	9.01	5.11	31.28	10.18	5.78	15.03	22.84	15.26	7.79
GP20	**5.69**	**10.93**	**5.69**	**29.34**	**10.88**	**6.38**	**15.23**	**23.41**	**19.30**	**8.42**
GP21	10.17	10.13	6.24	29.82	10.86	6.89	15.70	23.01	19.85	9.43
GP22	7.58	8.50	5.91	28.06	10.46	6.60	13.67	23.25	18.60	8.63
GP23	6.14	10.76	5.52	30.86	10.57	6.16	14.69	21.77	16.90	7.25
GP24	9.18	10.15	6.21	28.74	12.53	7.10	15.76	23.41	20.16	8.35
GP25	9.34	10.19	6.29	32.56	12.36	7.18	17.59	22.63	20.19	9.48
GP26	**6.38**	**11.62**	**6.38**	**32.83**	**12.27**	**7.25**	**18.28**	**23.77**	**16.18**	**7.69**
GP27	9.75	8.53	5.89	33.28	12.01	6.85	16.42	22.99	19.23	7.81
GP28	5.56	9.18	5.25	30.02	11.18	6.15	13.52	22.86	17.17	6.85
GP29	7.16	10.12	6.17	34.17	12.83	7.14	17.00	22.94	20.18	8.88
GP30	10.68	9.10	5.14	30.02	10.17	5.78	14.49	22.79	17.09	8.34

5.1 实验方法比较

- Vargha & Delaney's *A*-test

ID	Op1		Op2		AMPLE		Jaccard		Wong3	
	A	Count	*A*	Count	*A*	Count	*A*	Count	*A*	Count
GP01	0.51	3/63/6	0.50	2/64/6	0.53	25/46/1	0.51	22/47/3	0.50	7/60/5
GP02	0.38	9/16/47	0.35	8/16/48	0.39	22/8/42	0.36	19/10/43	0.39	13/15/44
GP03	0.45	4/52/16	0.42	0/56/16	0.45	21/33/18	0.42	20/33/19	0.44	5/54/13
GP04	0.37	11/9/52	0.34	7/9/56	0.37	16/9/47	0.34	10/9/53	0.37	9/9/54
GP05	0.49	6/53/13	0.47	4/53/15	0.49	19/42/11	0.47	15/41/16	0.50	10/51/11
GP06	0.49	4/48/20	0.47	3/48/21	0.50	6/56/10	0.47	5/48/19	0.48	6/46/20
GP07	0.46	6/38/28	0.44	2/42/28	0.47	19/30/23	0.44	14/31/27	0.46	7/38/27
GP08	**0.51**	**3/59/10**	**0.50**	**3/59/10**	**0.54**	**25/47/0**	**0.51**	**26/46/0**	**0.52**	**9/54/9**
GP09	0.50	6/51/15	0.48	2/55/15	0.50	17/43/12	0.48	17/42/13	0.50	4/53/15
GP10	**0.52**	**4/67/1**	**0.50**	**0/71/1**	**0.53**	**23/45/4**	**0.50**	**24/44/4**	**0.51**	**8/63/1**
GP11	**0.52**	**4/68/0**	**0.50**	**0/72/0**	**0.53**	**24/45/3**	**0.50**	**23/46/3**	**0.52**	**5/67/0**
GP12	0.48	2/53/17	0.47	2/53/17	0.50	19/46/7	0.48	19/45/8	0.49	2/55/15
GP13	**0.51**	**3/69/0**	**0.50**	**0/72/0**	**0.52**	**23/47/2**	**0.50**	**22/48/2**	**0.50**	**6/66/0**
GP14	0.50	2/59/11	0.49	2/59/11	0.52	20/49/3	0.49	18/49/5	0.50	5/56/11
GP15	0.51	3/63/6	0.50	3/63/6	0.51	21/48/3	0.50	21/48/3	0.52	10/56/6
GP16	0.50	2/58/12	0.49	2/58/12	0.53	22/47/3	0.50	17/50/5	0.52	10/53/9
GP17	0.48	5/50/17	0.45	1/53/18	0.49	22/33/17	0.46	18/35/19	0.48	8/49/15
GP18	0.50	4/61/7	0.48	0/65/7	0.50	21/42/9	0.48	20/43/9	0.50	2/64/6
GP19	0.50	4/49/19	0.49	3/49/20	0.52	20/46/6	0.50	16/46/10	0.51	8/49/15
GP20	**0.52**	**4/68/0**	**0.50**	**0/72/0**	**0.52**	**23/46/3**	**0.50**	**23/46/3**	**0.53**	**9/63/0**
GP21	0.50	3/61/8	0.49	3/61/8	0.51	22/46/4	0.49	20/46/6	0.51	9/55/8
GP22	0.50	2/67/3	0.49	0/69/3	0.52	22/47/3	0.50	20/49/3	0.52	5/65/2
GP23	0.52	4/63/5	0.50	0/67/5	0.52	23/45/4	0.50	19/47/6	0.52	5/64/3
GP24	0.51	3/56/13	0.50	3/56/13	0.52	20/50/2	0.50	19/49/4	0.51	6/54/12
GP25	0.48	11/46/15	0.47	8/47/17	0.50	17/37/18	0.48	18/36/18	0.50	12/43/17
GP26	**0.52**	**4/68/0**	**0.50**	**0/72/0**	**0.52**	**23/46/3**	**0.50**	**22/47/3**	**0.51**	**5/67/0**
GP27	0.51	2/58/12	0.50	2/58/12	0.52	21/51/0	0.50	11/51/10	0.51	6/54/12
GP28	0.52	3/60/9	0.51	3/60/9	0.53	22/50/0	0.51	21/49/2	0.52	8/57/7
GP29	0.51	6/45/21	0.49	5/45/22	0.52	19/41/12	0.50	18/39/15	0.52	11/42/19
GP30	0.50	3/60/9	0.49	1/62/9	0.50	18/46/8	0.49	17/46/9	0.51	4/59/9

5.1 实验方法比较

比较结果几乎完全基于实验观测

实验噪音

实验设置不完全相同

数据选取是否没有bias

实验规模是否足够大

没有统一确凿的评判方法

群龙无首
没有实质性进展

试图统一实验设置，尽可能减少风险，实验对象改用抽象程序

5.3 有限集合下的最优理论框架

- Definitions： 环境无关的任意两个公式之间的关系
  - Formula R_1 is better than formula R_2 ($R_1 \rightarrow R_2$)
    - $R_1 \rightarrow R_2$ if for any program, faulty statement, test suite and consistent tie-breaking scheme, $E_1 \leq E_2$
  - Formula R_1 is equivalent to formula R_2 ($R_1 \leftrightarrow R_2$)
    - $R_1 \leftrightarrow R_2$ if for any program, faulty statement, test suite and consistent tie-breaking scheme, $E_1 = E_2$

5.3 有限集合下的最优理论框架

- 集合的划分
  - Given a test suite and a formula R, a program PG=$\{s_1, s_2, \ldots, s_n\}$ can be divided into three subsets
    - S_B^R: higher risk values than the fault statement s_f, ranked before s_f
    - S_F^R: same risk values as the fault statement s_f, equally ranked with s_f
    - S_A^R: lower risk values than the fault statement s_f, ranked after s_f

$$S_B^R = \{s_i \in S \mid R(s_i) > R(s_f), 1 \le i \le n\}$$
$$S_F^R = \{s_i \in S \mid R(s_i) = R(s_f), 1 \le i \le n\}$$
$$S_A^R = \{s_i \in S \mid R(s_i) < R(s_f), 1 \le i \le n\}$$

只与R的定义有关

包含关系 ---- 真正决定定位效果的关键因素

5.3 有限集合下的最优理论框架

- 为了公平比较公式，提出Consistent tie-breaking scheme
  - Given any two sets of statements S_1, S_2
  - A scheme is consistent if it gives same relative order for statements in $S_1 \cap S_2$

5.3 有限集合下的最优理论框架

- Theorem 1：给定两个公式，如果其对应的程序语句子集存在关系 $S_B^{R_1} = S_B^{R_2}$，$S_F^{R_1} = S_F^{R_2}$，　$S_A^{R_1} = S_A^{R_2}$，则 $R_1 \boxed{\leftrightarrow} R_2$ ，也即，$R_2 \rightarrow R_1$ 且 $R_1 \rightarrow R_2$

5.3 有限集合下的最优理论框架

- Theorem 2：给定两个公式，如果其对应的程序语句子集存在关系$S_B^{R_1} \subseteq S_B^{R_2}$，$S_A^{R_2} \subseteq S_A^{R_1}$，则$R_1 \rightarrow R_2$
- Theorem 3：给定两个公式，如果其对应的程序语句子集存在关系$S_B^{R_1} \subsetneq S_B^{R_2}$，$S_A^{R_2} \subsetneq S_A^{R_1}$，则$R_1 \rightarrow R_2$ ，且R_2 --\> R_1, 即R_1严格优于R_2

5.2 局部最优理论框架

- Theorem 4：给定两个公式，如果其对应的程序语句子集存在反例使得相对应的子集间无确定包含关系，则R_1 --\> R_2，且R_2 --\> R_1

```
start
if (S1)
   S2;
else
{
   if (S3)
      S4;
   else
   {
      if (S5)
      {
         S6;
         S8;
         S9;
      }
      else
         S7;
      S10;
   }
}
exit
```

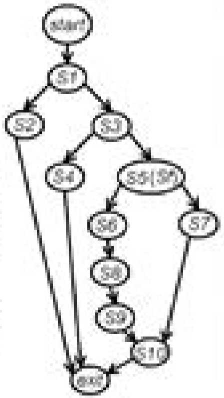


Fig. 6. Sample program PG_2.

Table III. A_i for PG_2 with TS_3

Statement	$A_i = <a^i_{ef}, a^i_{ep}, a^i_{nf}, a^i_{np}>$
s_1	<40, 160, 0, 0>
s_2	<0, 70, 40, 90>
s_3	<40, 90, 0, 70>
s_4	<0, 30, 40, 130>
s_5	<40, 60, 0, 100>
s_6	<40, 30, 0, 130>
s_7	<0, 30, 40, 130>
s_8	<40, 30, 0, 130>
s_9	<40, 30, 0, 130>
s_{10}	<40, 60, 0, 100>

5.3 有限集合下的最优理论框架

- GP2, GP3, GP13, GP19在与35个人工公式对比中被证明是maximal

6组等价公式

Equivalence Group	Formula	Expression
ER_1	Naish1 [37]	$\begin{cases} -1 & \text{if } e_f < F \\ n_p & \text{if } e_f = F \end{cases}$
	Naish2 [37]	$e_f - \frac{e_p}{e_p+n_p+1}$
	GP13 [55]	$e_f(1 + \frac{1}{2e_p+e_f})$
ER_5	Wong1 [51]	e_f
	Russel & Rao [46]	$\frac{e_f}{e_f+n_f+e_p+n_p}$
	Binary [37]	$\begin{cases} 0 & \text{if } e_f < F \\ 1 & \text{if } e_f = F \end{cases}$
GP02 [55]		$2(e_f + \sqrt{n_p}) + \sqrt{e_p}$
GP03 [55]		$\sqrt{\lvert e_f^2 - \sqrt{e_p} \rvert}$
GP19 [55]		$e_f\sqrt{\lvert e_p - e_f + n_f - n_p \rvert}$

证实或推翻过去20年的研究成果

5.3 有限集合下的最优理论框架

Proposition 1: *GP13 ↔ Op1 and GP13 ↔ Op2.*

$$X^{Op}=\{s_i|e_f^i=F \ \ and \ \ e_p^f>e_p^i, 1\le i\le n\}$$
$$Y^{Op}=\{s_i|e_f^i=F \ \ and \ \ e_p^f=e_p^i, 1\le i\le n\}$$
$$Z^{Op}=S\backslash X^{Op}\backslash Y^{Op}$$

$$S_B^{GP13}=\{s_i|e_f^i(1+\frac{1}{2e_p^i+e_f^i})>F(1+\frac{1}{2e_p^f+F}), 1\le i\le n\}$$
$$S_F^{GP13}=\{s_i|e_f^i(1+\frac{1}{2e_p^i+e_f^i})=F(1+\frac{1}{2e_p^f+F}), 1\le i\le n\}$$

Proof. Refer to Lemma 3 and Lemma 4, we have $S_B^{N1}=S_B^{N2}=S_B^{GP13}$, $S_F^{N1}=S_F^{N2}=S_F^{GP13}$ and $S_A^{N1}=S_A^{N2}=S_A^{GP13}$, respectively. After Theorem 1, GP13 $\leftrightarrow$ Naish1 and GP13 $\leftrightarrow$ Naish2.

5.3 有限集合下的最优理论框架

Proposition 1: *GP13* $\leftrightarrow$ *Op1 and GP13* $\leftrightarrow$ *Op2.*

1. To prove that $S_B^{GP13} = X^{Op}$.

   (1) To prove $X^{Op} \subseteq S_B^{GP13}$.

   For any $s_i \in X^{Op}$, we have $F(1+\frac{1}{2e_p^i+F}) > F(1+\frac{1}{2e_p^f+F})$ because $e_p^f > e_p^i$ and $F>0$. Since $e_f^i = F$, we have $e_f^i(1+\frac{1}{2e_p^i+e_f^i}) > F(1+\frac{1}{2e_p^f+F})$, which implies $s_i \in S_B^{GP13}$. Thus, we have proved $X^{Op} \subseteq S_B^{GP13}$.

   (2) To prove $S_B^{GP13} \subseteq X^{Op}$.

   For any $s_i \in S_B^{GP13}$, we have $e_f^i(1+\frac{1}{2e_p^i+e_f^i}) > F(1+\frac{1}{2e_p^f+F})$. Let us consider the following two exhaustive cases.

5.3 有限集合下的最优理论框架

Propo

2. To prove that $S_F^{GP13} = Y^{Op}$.

(1) To prove $Y^{Op} \subseteq S_F^{GP13}$.

For any $s_i \in Y^{Op}$, we have $e_f^i(1+\frac{1}{2e_p^i+e_f^i})=F(1+\frac{1}{2e_p^f+F})$ because $e_f^i=F$ and $e_p^f=e_p^i$. After the definition of S_F^{GP13}, $s_i \in S_F^{GP13}$. Thus, we have proved $Y^{Op} \subseteq S_F^{GP13}$.

(2) To prove $S_F^{GP13} \subseteq Y^{Op}$.

For any $s_i \in S_F^{GP13}$, we have $e_f^i(1+\frac{1}{2e_p^i+e_f^i})=F(1+\frac{1}{2e_p^f+F})$. Let us consider the following two exhaustive cases.

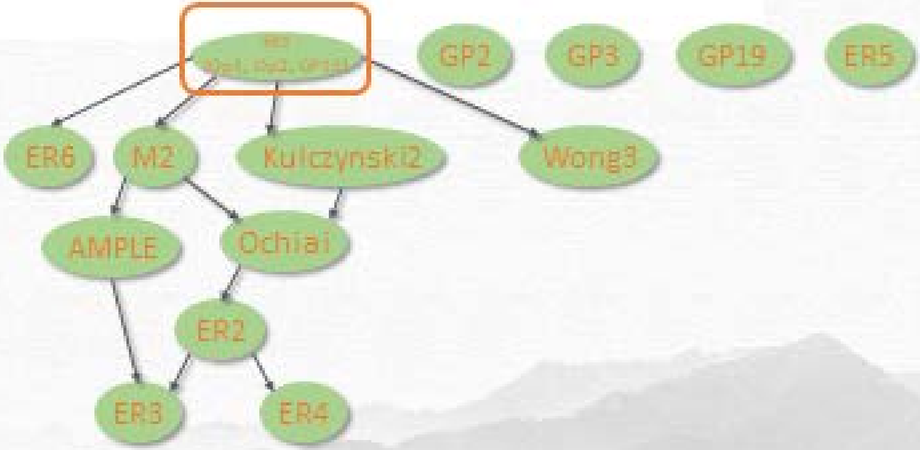
- Case (i) $e_f^i<F$. First, consider the sub-case that $e_f^i=0$. Then we have $e_f^i(1+\frac{1}{2e_p^i+e_f^i})=0$. It follows from the definition of S_F^{GP13} that $0=F(1+\frac{1}{2e_p^f+F})$, which is however contradictory to $F>0$ and $e_p^f \geq 0$. Thus, it is impossible to have $e_f^i=0$. Now, consider the sub-case that $0<e_f^i<F$. Similar to the above proof of $S_B^{GP13} \subseteq X^{Op}$, we can prove that $(\frac{1}{1+2\frac{e_p^i}{e_f^i}}-\frac{1}{1+2\frac{e_p^f}{F}})<(F-e_f^i)$, which is however contradictory to $e_f^i(1+\frac{1}{2e_p^i+e_f^i})=F(1+\frac{1}{2e_p^f+F})$. Therefore, it is impossible to have $0<e_f^i<F$. Therefore, we have proved that if $s_i \in S_F^{GP13}$, then we cannot have $e_f^i<F$.
- Case (ii) $e_f^i=F$. Assume further $e_p^i \neq e_p^f$. Obviously, we have $F(1+\frac{1}{2e_p^i+F}) \neq F(1+\frac{1}{2e_p^f+F})$, which can be re-written as $e_f^i(1+\frac{1}{2e_p^i+e_f^i}) \neq F(1+\frac{1}{2e_p^f+F})$. However, this is contradictory to $e_f^i(1+\frac{1}{2e_p^i+e_f^i}) = F(1+\frac{1}{2e_p^f+F})$. Thus, the only possible case is $e_p^f=e_p^i$.

We have proved that if $s_i \in S_F^{GP13}$, then $e_f^i=F$ and $e_p^f=e_p^i$, which imply $s_i \in Y^{Op}$. Therefore, $S_F^{GP13} \subseteq Y^{Op}$.

5.3 有限集合下的最优理论框架

Proposition 1: *GP13 ↔ Op1 and GP13 ↔ Op2.*

Proof. Refer to Lemma 3 and Lemma 4, we have $S_B^{N1} = S_B^{N2} = S_B^{GP13}$, $S_F^{N1} = S_F^{N2} = S_F^{GP13}$ and S_A^{N1}=S_A^{N2} =S_A^{GP13}, respectively. After Theorem 1, GP13 ↔ Naish1 and GP13 ↔ Naish2.



5.3 有限集合下的最优理论框架

Proposition 2. *GP02, GP03, GP19, ER1 and ER5 are distinct maximal formulas (or groups of equivalent formulas.*

- With *TS*12 ER1′ → GP02 does not hold; with *TS*21 GP02 → ER1′ does not hold.
- With *TS*12 ER5 → GP02 does not hold; with *TS*21 GP02 → ER5 does not hold
- With *TS*11 ER1′ → GP03 does not hold; with *TS*12 GP03 → ER1′ does not hold.
- With *TS*11 ER5 → GP03 does not hold; with *TS*12 GP03 → ER5 does not hold.
- With *TS*11 ER1′ → GP19 does not hold; with *TS*12 GP19 → ER1′ does not hold.
- With *TS*11 ER5 → GP19 does not hold; with *TS*12 GP19 → ER5 does not hold.
- With *TS*11 GP02 → GP03 does not hold; with *TS*12 GP03 → GP02 does not hold.
- With *TS*11 GP02 → GP19 does not hold; with *TS*12 GP19 → GP02 does not hold.
- With *TS*21 GP03 → GP19 does not hold; with *TS*22 GP19 → GP03 does not hold.

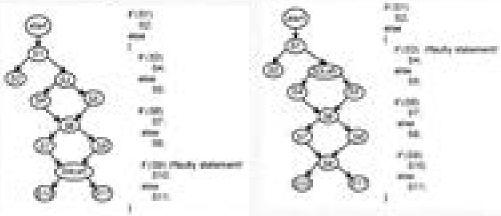
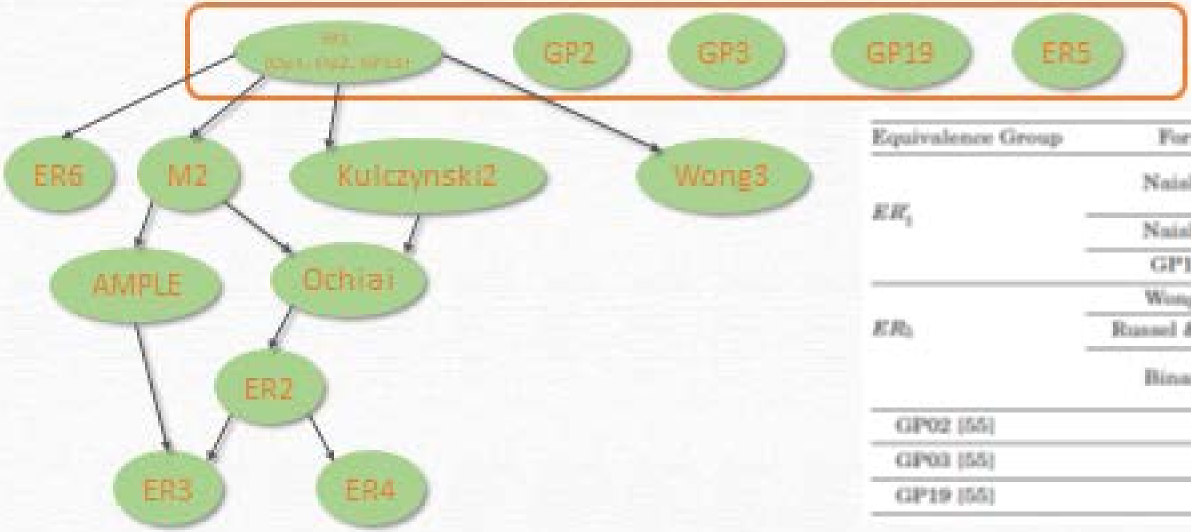


Table 2. ε for PG_1 and PG_2 with different test suites

Statement	$\varepsilon=\langle e_f^s, e_p^s, e_f^n, e_p^n\rangle$			
	$TS1_1$	$TS1_2$	$TS2_1$	$TS2_2$
s_1	$\langle 1,6,0,0\rangle$	$\langle 1,8,0,0\rangle$	$\langle 2,15,0,0\rangle$	$\langle 10,15,0,0\rangle$
s_2	$\langle 0,1,1,5\rangle$	$\langle 0,6,1,2\rangle$	$\langle 0,1,2,14\rangle$	$\langle 0,1,10,14\rangle$
s_3	$\langle 1,5,0,1\rangle$	$\langle 1,2,0,6\rangle$	$\langle 2,14,0,1\rangle$	$\langle 10,14,0,1\rangle$
s_4	$\langle 1,4,0,2\rangle$	$\langle 1,1,0,7\rangle$	$\langle 1,7,1,8\rangle$	$\langle 9,0,1,13\rangle$
s_5	$\langle 0,1,1,5\rangle$	$\langle 0,1,1,7\rangle$	$\langle 1,7,1,8\rangle$	$\langle 1,14,9,1\rangle$
s_6	$\langle 1,5,0,1\rangle$	$\langle 1,2,0,6\rangle$	$\langle 2,14,0,1\rangle$	$\langle 10,14,0,1\rangle$
s_7	$\langle 1,4,0,2\rangle$	$\langle 1,1,0,7\rangle$	$\langle 1,8,1,7\rangle$	$\langle 5,6,5,9\rangle$
s_8	$\langle 0,1,1,5\rangle$	$\langle 0,1,1,7\rangle$	$\langle 1,6,1,9\rangle$	$\langle 5,8,5,7\rangle$
s_9	$\langle 1,5,0,1\rangle$	$\langle 1,2,0,6\rangle$	$\langle 2,14,0,1\rangle$	$\langle 10,14,0,1\rangle$
s_{10}	$\langle 1,4,0,2\rangle$	$\langle 1,1,0,7\rangle$	$\langle 1,9,1,6\rangle$	$\langle 1,12,9,1\rangle$
s_{11}	$\langle 0,1,1,5\rangle$	$\langle 0,1,1,7\rangle$	$\langle 1,5,1,10\rangle$	$\langle 9,2,1,13\rangle$

5.3 有限集合下的最优理论框架

- GP2, GP3, GP13 GP19在与35个人工公式对比中被证明是maximal



Equivalence Group	Formula	Expression
ER_1	Naish1 [37]	$\begin{cases} -1 & \text{if } e_f < F \\ n_p & \text{if } e_f = F \end{cases}$
	Naish2 [37]	$e_f - \frac{e_p}{e_p + n_p + 1}$
	GP13 [55]	$e_f(1 + \frac{1}{2e_p + e_f})$
ER_5	Wong1 [51]	e_f
	Russel & Rao [46]	$\frac{e_f}{e_f + n_f + e_p + n_p}$
	Binary [37]	$\begin{cases} 0 & \text{if } e_f < F \\ 1 & \text{if } e_f = F \end{cases}$
GP02 [55]		$2(e_f + \sqrt{n_p}) + \sqrt{e_p}$
GP03 [55]		$\sqrt{\lvert e_f^2 - \sqrt{e_p} \rvert}$
GP19 [55]		$e_f \sqrt{\lvert e_p - e_f + n_f - n_p \rvert}$

5.3 无限集合下的最优理论框架

- 定义： Limited maximality --- A risk evaluation formula R_1 from a subset of formulas, $\mathbb{S} \subset \mathcal{F}$, is said to be a *maximal* formula of $\mathbb{S}$ if for any element $R_2 \in \mathbb{S}$, $R_2 \rightarrow R_1$ implies $R_2 \leftrightarrow R_1$.
- 定义： Maximality --- A risk evaluation formula R_1 from a subset of formulas, $\mathbb{S}$ is said to be a *maximal* formula in $\mathcal{F}$, if for any distinct formula $R_2 \in \mathcal{F}$ such that $R_2 \rightarrow R_1$ implies $R_2 \leftrightarrow R_1$

有限集 → 无限集

How to generalized the analysis?

5.3 无限集合下的最优理论框架

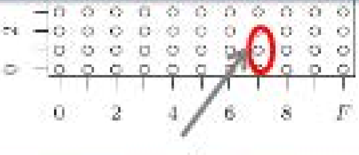
- 频谱坐标系：

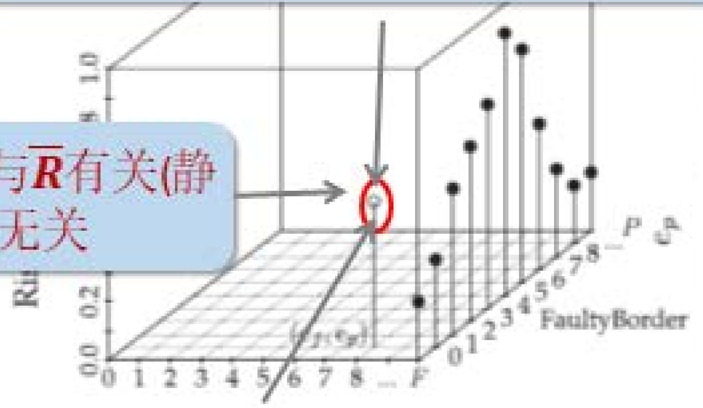
$$\overline{R}(e_{if}, e_{ip}) = R(e_{if}; e_{ip}; n_{if}; n_{ip}).$$

$$\overline{\mathcal{F}} = \{\overline{R} \mid \overline{R}: I_f \times I_p \rightarrow Real\}$$

$\overline{R}$ 将平面上每个<e_{if}, e_{ip}>点映射到三维空间Risk Value轴的相应高度

Risk Value轴的每个点的取值只与$\overline{R}$有关(静态信息)，与动态测试执行情况无关





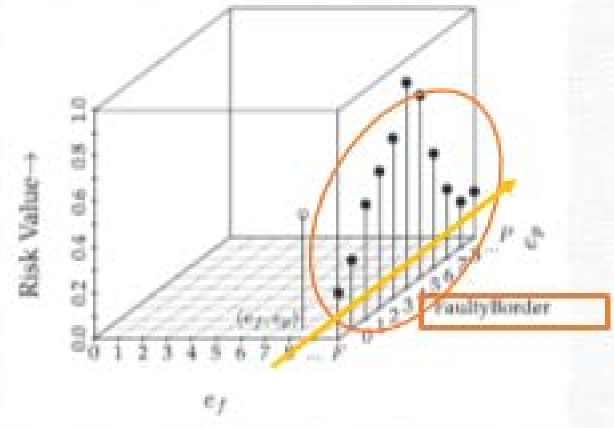
给定P&TS(动态执行信息)，每个点上将会附结>=0个程序语句, 这些语句所获得的risk值就是其所附结点的risk值；不同执行信息决定了每个点上所附结语句的密度，不同的密度会导致当前P&TS情况下sf的ranking不同

can be interpreted as a 2D plane with discrete coordinates depicted above.

la $\overline{R}$: SBFL formular assign ... be mapped to the spectral ... btain a 3D space, in which ... res. The points whose e_f values are equal to F form the faulty border: the faulty statement is guaranteed to be mapped on the border.

5.3 无限集合下的最优理论框架

- 频谱坐标系中的Faulty Border： we call the sequential points <(F, 0), (F, 1), ..., (F, e_p), ..., (F, P)> 为Faulty Border，记为*E*



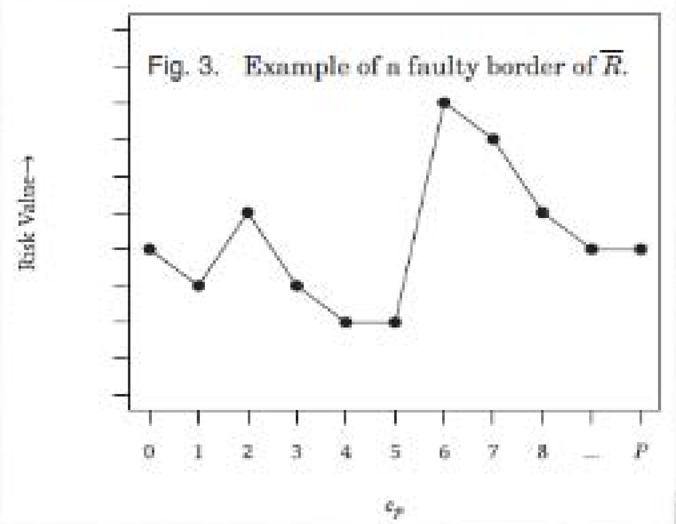
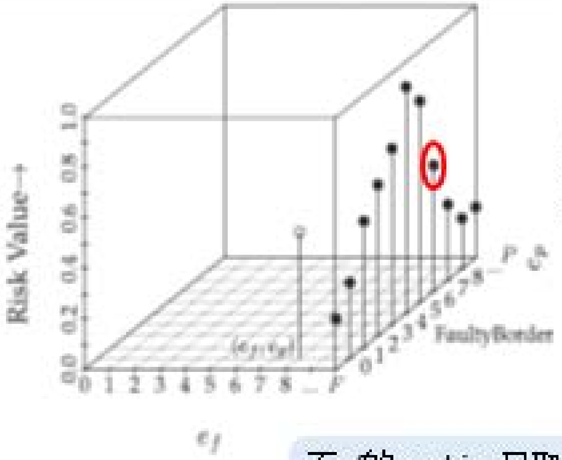


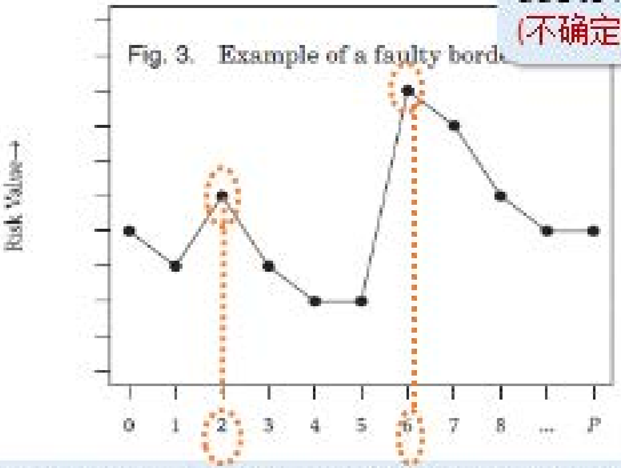
Fig. 3. Example of a faulty border of $\overline{R}$.

5.3 无限集合下的最优理论框架

- Faulty Border作用：single fault s_f 一定附结于Faulty Border的某个点上，因为e_f=F

附在哪个点取决于s_f的ep值 -never know due to inevitable coincidental correctness (不确定性1)



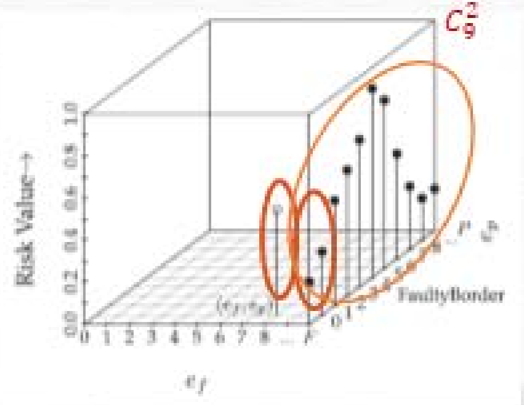


而s_f的ranking又取决于R对于s_f附结点所赋予的risk值在所有faulty border上点risk值的排名 (不确定性2)

5.3 无限集合下的最优理论框架

- 解决办法：构建辅助概念
  - U_R denotes the set of points outside E that have risk higher than or equal to those of some points on E, for formula R.
  - P_R denotes the set of relations of risk values between any two distinct points (F, e_{pi}) and (F, e_{pj}) $\{\langle i, j, op\rangle\}$

Monotonicity with e_p



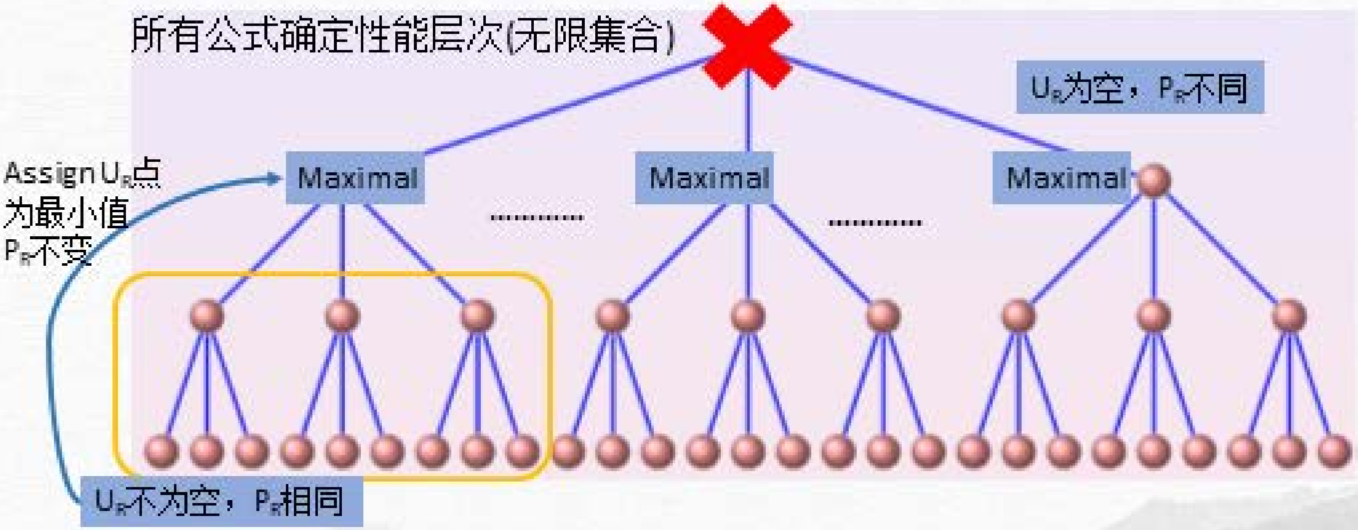
5.3 无限集合下的最优理论框架

- Theorem 1: Any formula $R \in F$, R is maximal, iff $\boldsymbol{U_R} = \emptyset$
  - $U_R = \emptyset$是 maximality 的充要条件!
  - 性质: If $U_{R1} = U_{R2} = \emptyset$, and $P_{R1} = P_{R2,}$ then, R_2 $\boxed{\leftrightarrow}$ R_1
  - 性质: If $U_{R1} = U_{R2} = \emptyset$, and $P_{R1} \neq P_{R2,}$ then, R_2 <-\> R_1

- Theorem 2: There is no formula that is greatest against the set of all formulas F

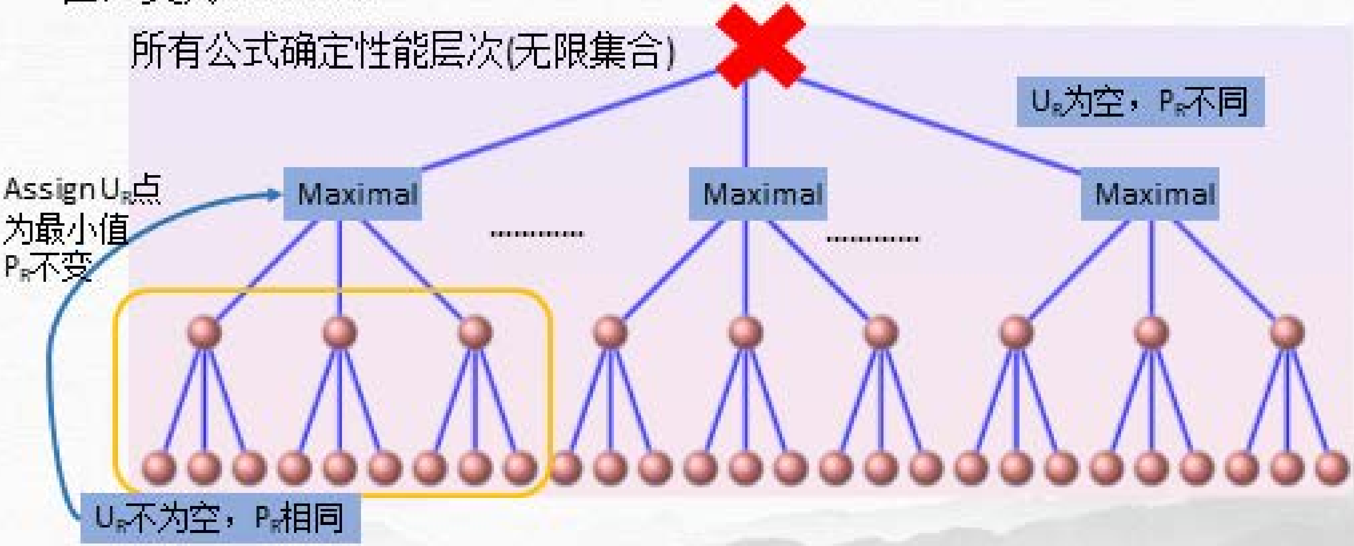
5.3 无限集合下的最优理论框架

- 全局视图：

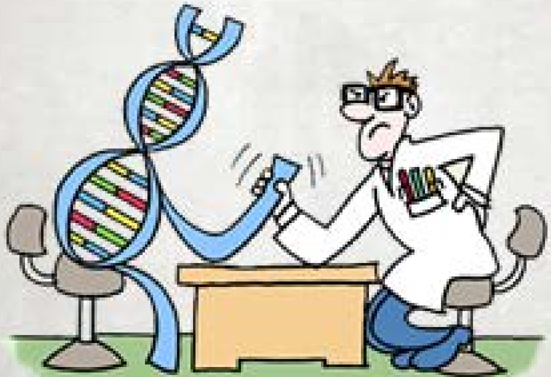


5.3 无限集合下的最优理论框架

- GP13是 无限集合下的maximal
- GP2、GP3、GP19不是无限集合下的maximal，但可以通过修改U_R中点的risk值，变为maximal



第肆回合



5.4 实证分析

- GP13是“practical greatest”formula

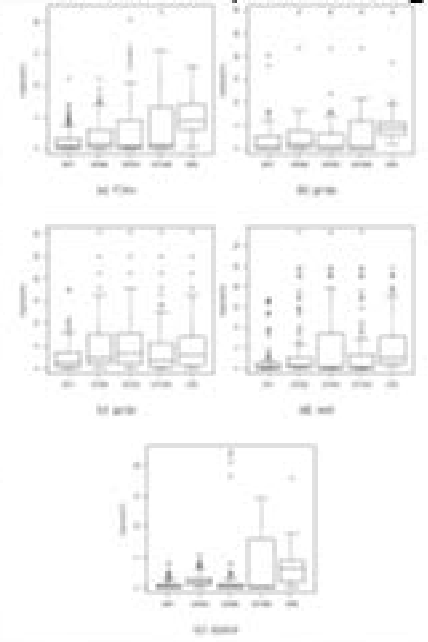


Table VI. The p-values from Shapiro-Wilk Normality Test on Observed Expense Values

Subject	ER_1	$GP2^M$	$GP3^M$	$GP19^M$	ER_5
flex	< 1e-4	< 1e-4	< 1e-4	< 1e-4	[illegible]
grep	< 1e-4	< 1e-4	< 1e-4	< 1e-4	< 1e-4
gzip	< 1e-4	< 1e-4	< 1e-4	< 1e-4	< 1e-4
sed	< 1e-4	< 1e-4	< 1e-4	< 1e-4	< 1e-4
space	< 1e-4	< 1e-4	< 1e-4	< 1e-4	< 1e-4

- flex: $ER_1 \geq (GP_5^M \geq GP_{13}^M \geq GP_2^M) \geq ER_5$.
- grep: $ER_1 \geq (GP_5^M \geq GP_{13}^M \geq GP_2^M) \geq ER_5$.
- gzip: $ER_1 \geq (GP_{13}^M \geq GP_5^M) \geq (ER_5 \geq GP_2^M)$
- sed: $ER_1 \geq (GP_{13}^M \geq GP_5^M \geq GP_2^M) \geq ER_5$.
- space: $(ER_1 \geq GP_2^M) \geq (GP_{13}^M \geq GP_5^M) \geq ER_5$.

Table VII. Interpretation of the Hypotheses in the Context of SBFL

Hypotheses	H_0	H_1
Acceptance Condition:	p-value $\geq$ 0.05	p-value < 0.05
2-tailed	$E(A) \simeq E(B)$: A and B DO have similar performance	$E(A) \not\simeq E(B)$: A and B DO NOT have similar performance
1-tailed (lower)	$E(A) \geq E(B)$: A DOES NOT tend to be worse than B	$E(A) < E(B)$: A DOES tend to be worse than B
1-tailed (upper)	$E(A) \leq E(B)$: A DOES NOT tend to be better than B	$E(A) > E(B)$: A DOES tend to be better than B

尘埃落定。。。

6. 小结

~2012

人工设计评估公式

2012

GP-evolved formulas出现，并在实验中胜出

2013

风险评估公式理论分析框架提出，GP-evolved公式被证明在现有人工公式中为理论maximal

2017~2019

理论分析框架被扩展至无限集合，GP-evolved 公式被证明为general maximal，并为practical greatest

6. 小结

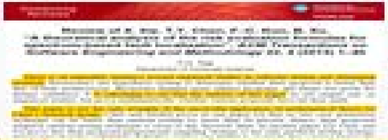
- 相关论文
  - J.Tu , **X.Xie (*)**, T.Y.Chen , B.Xu , On the analysis of spectrum based fault localization using hitting sets, **Journal of Systems and Software,** 2019, Vol.147, pp. 106-123.
  - S. Yoo (*), **X. Xie (*),** F-C. Kuo, T. Y. Chen, M. Harman, Human Competitiveness of Genetic Programming in Spectrum-Based Fault Localisation: Theoretical and Empirical Analysis, **ACM Transactions on Software Engineering and Methodology**, 2017, 26(1), pp. 4:1-4:30.
  - **X.Xie**, Z.Liu, S. Song, Z. Chen, J. Xuan and B.W. Xu, Revisit of Automatic Debugging via Human Focus-tracking, International Conference on Software Engineering **(ICSE)** 2016.
  - T. Y. Chen, **X. Xie (*),** F-C. Kuo, B.W.Xu, A Revisit of a Theoretical Analysis on Spectrum-Based Fault Localization, Annual International Computers, Software & Applications Conference **(COMPSAC)**, 2015
  - **X. Xie (*),** T. Y. Chen, F.-C. Kuo, B. W. Xu, A Theoretical Analysis of The Risk Evaluation Formulas for Spectrum-Based Fault Localization, **ACM Transaction on Software Engineering and Methodology,** 2013, Vol. 22, No. 4, pp. 31:1-31:40.
  - **X. Xie (*)**, F.-C. Kuo, T.Y. Chen, S. Yoo and M. Harman, Provably Optimal and Human-Competitive Results in SBSE for Spectrum Based Fault Localisation, Symposium on Search Based Software Engineering **(SSBSE)**, 2013, pp. 224-238.

6. 小结

“Human Competitiveness of Genetic Programming in Spectrum Based Fault Localisation: Theoretical and Empirical Analysis.”

14th Annual ACM SIGEVO “HUMIES” Awards (银奖) 2017, Berlin, Germany

6. 小结



"There is an unhealthy tendency toward empirical studies in software testing and debugging research... ...It is refreshing to see that the authors of this paper ... ...example of the successful application of mathematical theory ... ..."

ACM Computing Reviews, 29(12): “令人耳目一新的成果，阻止了该领域先前错误的研究方向，是软件工程理论研究的成功案例”

George K. Baah: “非常高兴你们团队对Harrold教授的猜想给出了证明，阻止了未来所有新公式的提出”

"I am happy that your team has been able to make this contribution. Mary Jean, Andy Podgurski and I suspected something like this based on published empirical results and I am happy your team has been able to prove it. I hope it puts a stop to all the fault localization formulas that are being created or imported from the statistics literature."

自动缺陷定位的奠基人

(1947-2013)

ACM/IEEE Fellow
ACM TOSEM 编委
ACM SIGSOFT 执行主席
北美计算研究协会CRA-W主席
2004年美国杰出科学家总统奖

Mary Jean Harrold教授 Georgia Tech

谢谢