Leveraging Fuzzy System to Reduce Uncertainty of Decision Making in Software Engineering Automation

Yueke Zhang

Southwest University

Yu Huang

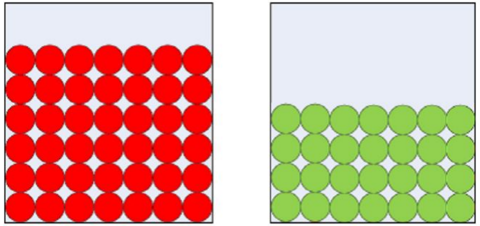
Vanderbilt University

@GI2022

Ideal Data vs. Real Life Data

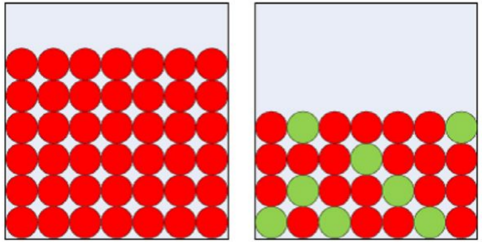
Ambiguity

ideal



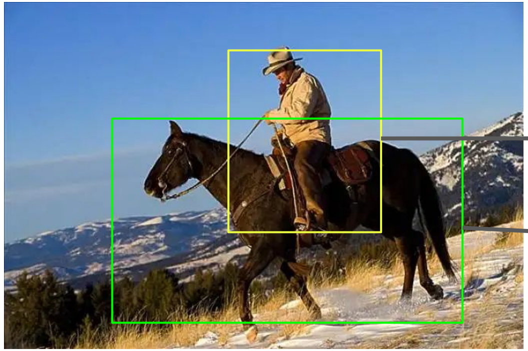
P1=0.6 P2=0.4

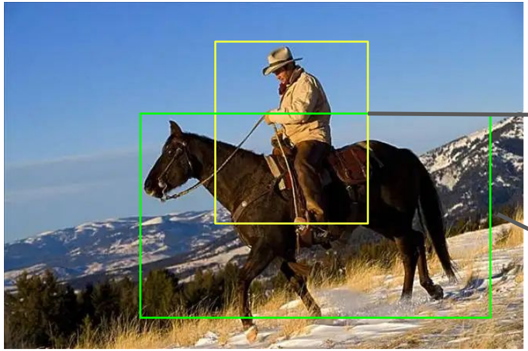
real



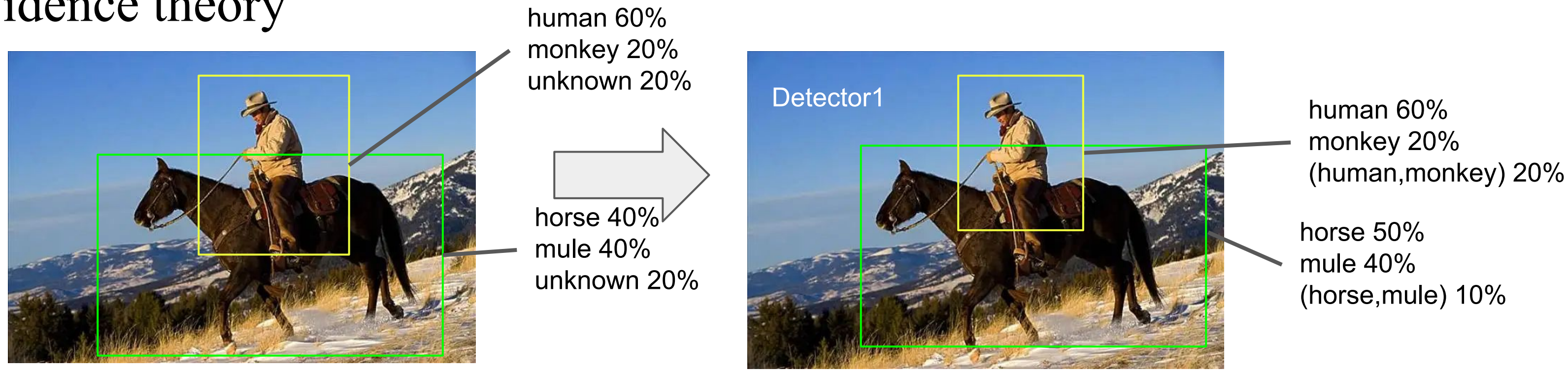
P1=0.6 P2=0.4

Partial information

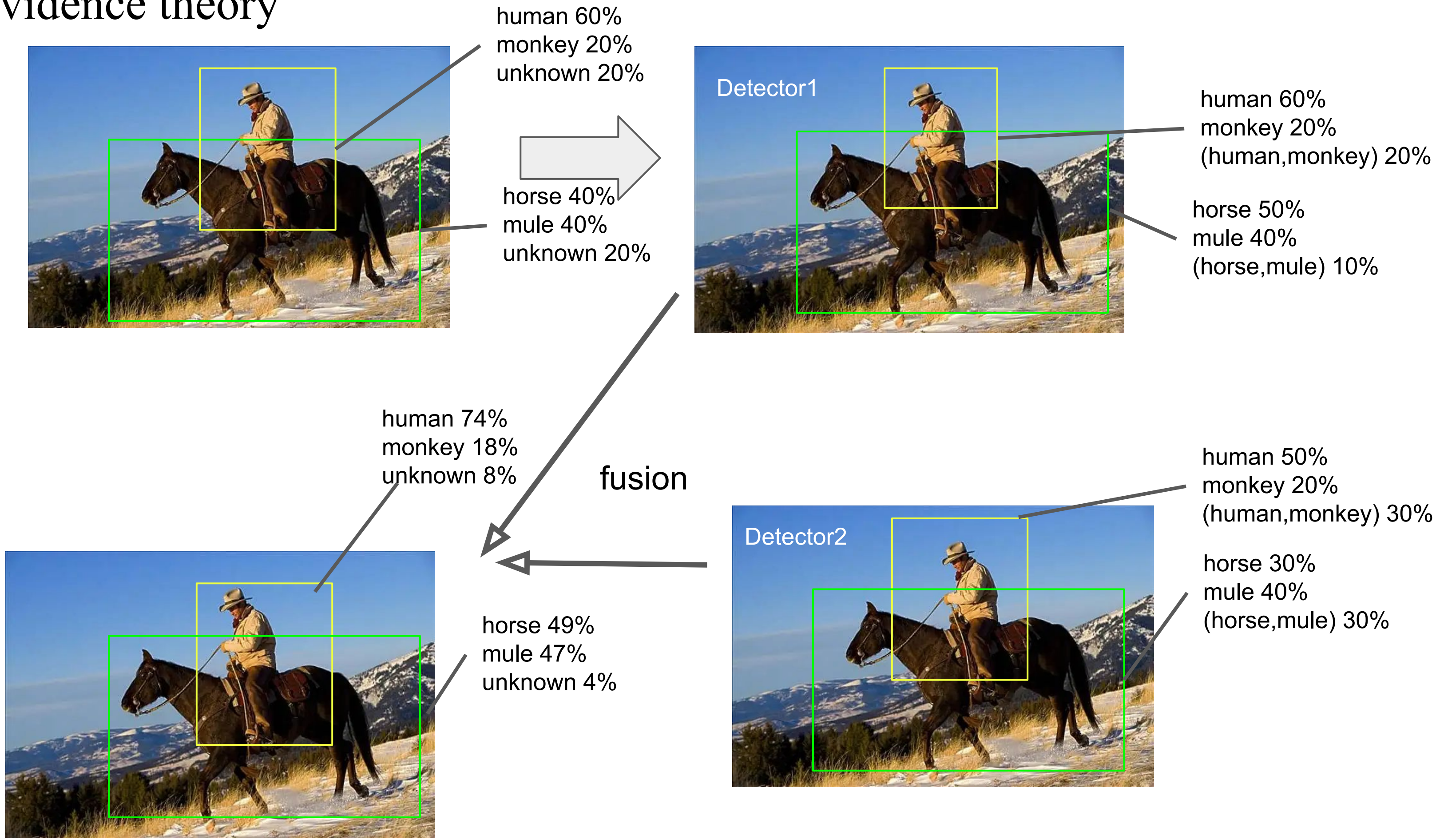




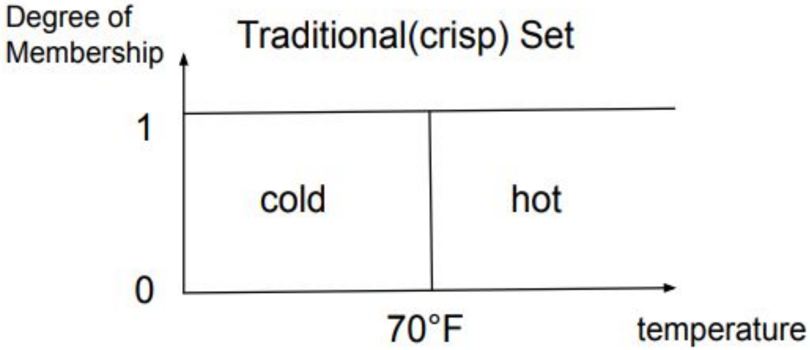
Evidence theory

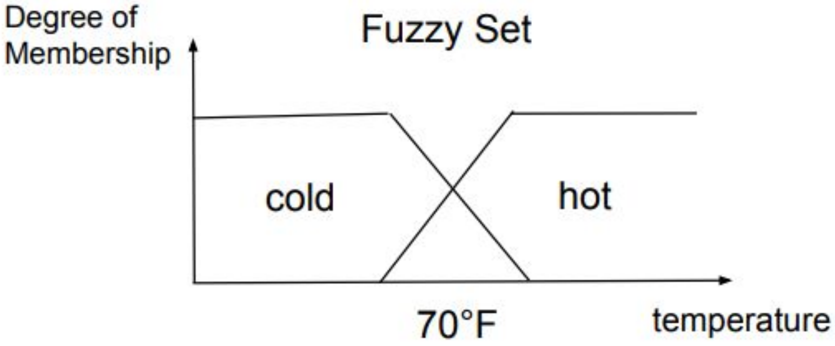


Evidence theory



The fuzzy set theory





For 70ºF

m(cold)=1 or m(hot)=1

For 70ºF

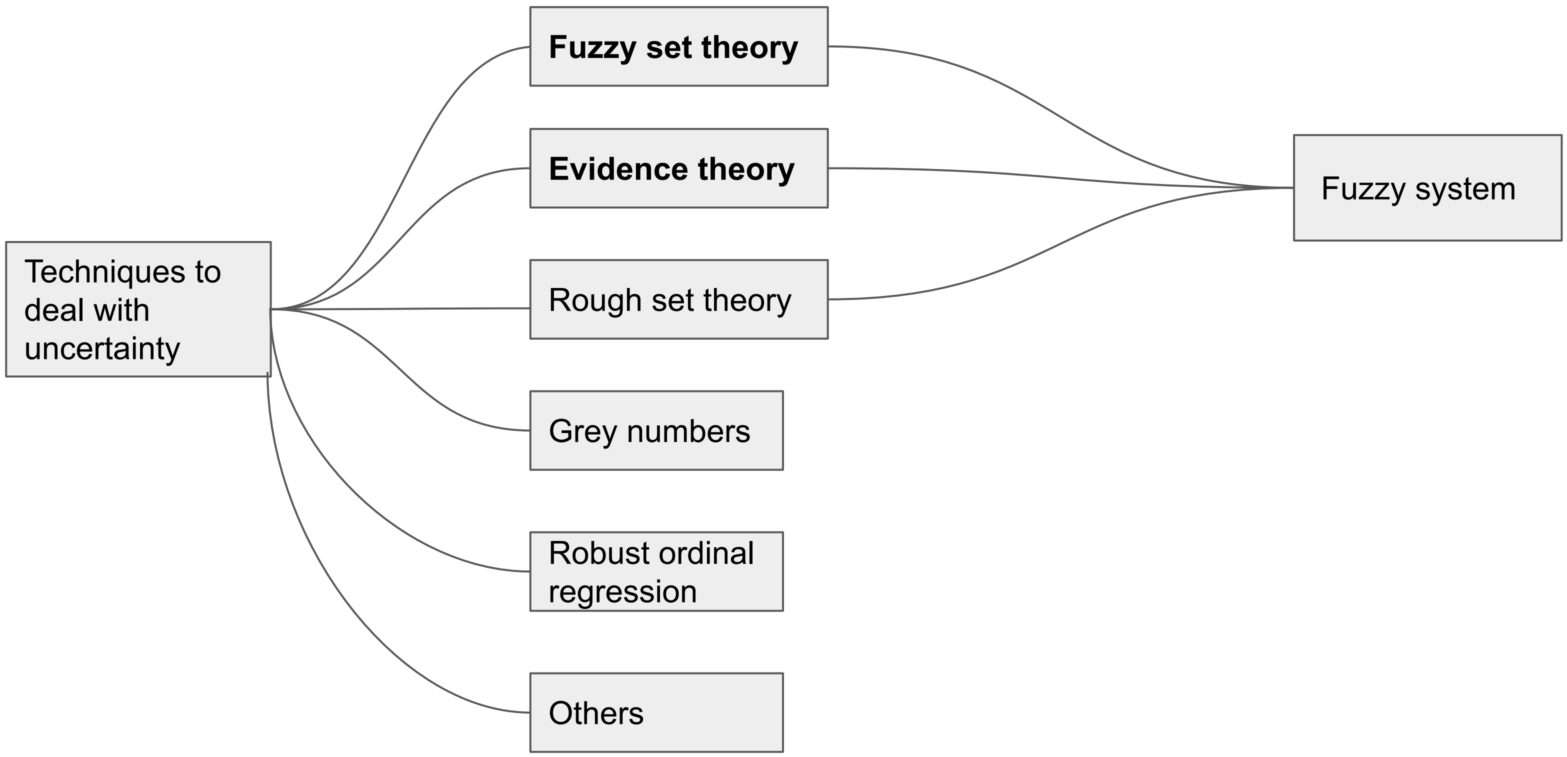
m(cold)=0.5 m(hot)=0.5

For 60ºF

m(cold)=0.4 m(hot)=0.6

m(ɑ) means membership of ɑ

What is fuzzy based uncertain decision making?



Fault localizaition and its decision-making process

Fault localization →

1.spectrum based fault localization
2.mutation based fault localization
3.statistical based fault localization
...

→ *tarantula ochiai ample zoltar….*

Fault localizaition and its decision-making process

Fault localization → 1.spectrum based fault localization → *tarantula ochiai ample zoltar….*

2.mutation based fault localization

3.statistical based fault localization

…

$$ochiai(e) = \frac{failed(e)}{totalfailed \times (failed(e)) + passed(e)}$$

$$ample(e) = \left| \frac{failed(e)}{totalfailed - failed(e)} - \frac{passed(e)}{totalpassed - passed(e)} \right|$$

$$tarantula(e) = \frac{failed(e) \times totalpassed}{failed(e) \times totalpassed + passed(e) \times totalfailed}$$

→ *A decision with three opionions*

Ochiai vs. Ample (Exam score)

	Ochiai	Ample
Math 1	0.004	0.026
Math 14	0.207	0.024
Math 19	0.013	0.016
Math 23	0.105	0.087
Math all	0.117	0.122
defects4j v.1.1.0	0.099	0.104

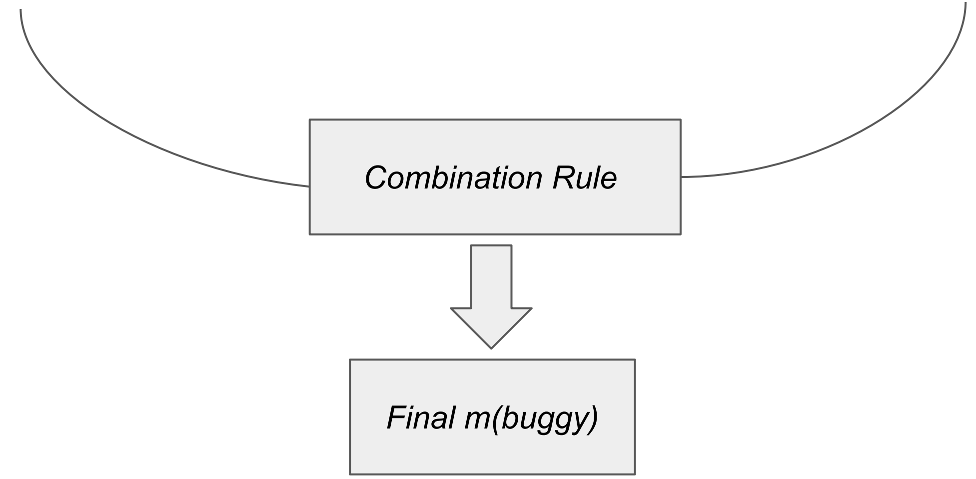
Different SBFL formulas in FL

DM1: $ochiai(e) = \frac{failed(e)}{totalfailed \times (failed(e)) + passed(e)}$

DM2: $ample(e) = \left| \frac{failed(e)}{totalfailed - failed(e)} - \frac{passed(e)}{totalpassed - passed(e)} \right|$

$$m(buggy) = norm(ochiai(e))$$
$$m(unknown) = 1 - m(buggy)$$

$$m(buggy) = norm(ample(e))$$
$$m(unknown) = 1 - m(buggy)$$



A fusion example

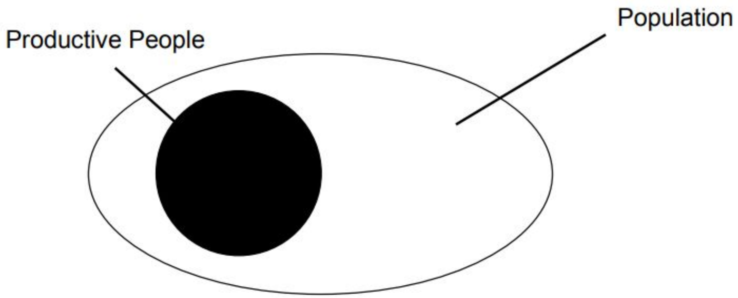
Line #	Line of Code	Ochiai	Ample	Zoltar	Fuzzy
1	`currentEvent.stepAccepted(eventT, eventY);`	0.75 (113)	0.48 (366)	0.53 (124)	0.94 (124)
2	`isLastStep = currentEvent.stop();`	0.75 (122)	0.48 (362)	0.53 (114)	0.94 (126)
3	`for (final StepHandler handler : stepHandlers) {`	0.75 (121)	0.48 (370)	0.53 (121)	0.94 (131)
4	`handler.handleStep(interpolator, isLastStep);}`	1.00 (13)	0.54 (187)	1.00 (13)	1.00 (6)
5	`if (isLastStep) {`	0.75 (95)	0.48 (390)	0.53 (95)	0.94 (105)
6	`System.arraycopy(eventY, 0, y, 0, y.length);`	0 (8283)	0.08 (7518)	0 (8283)	0.08 (7531)
7	`for (final EventState remaining : occuringEvents) {`	0 (8293)	0.08 (7519)	0 (8293)	0.08 (7529)
8	`remaining.stepAccepted(eventT, eventY); }`	0 (8269)	0.01 (7750)	0 (8269)	0.01 (7761)
	Exam Score	0.41	0.39	0.41	**0.38**

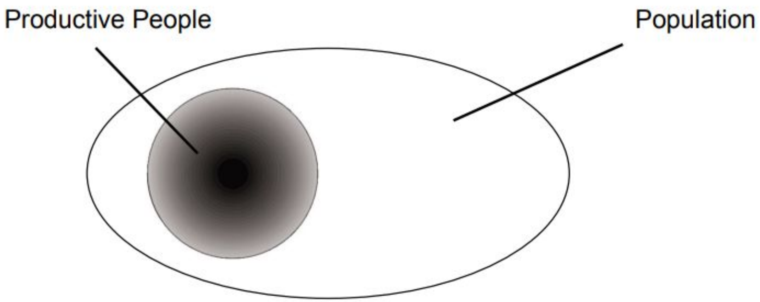
Buggyline

Exam Score

Rank in code sinppet

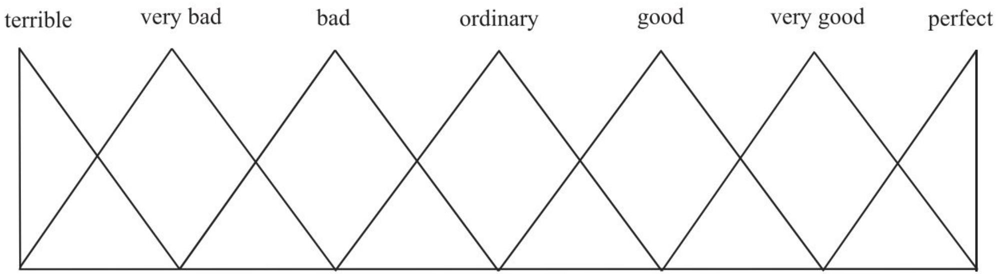
Other insights: importing fuzzy theory in human study





Traditional set

Fuzzy linguistic set



Challenge and Opportunities

- **Modeling**
  Measure uncertainty in domain knowledge

- **Computational overhead**
  The computational cost of fusion rule is negligible, but we require different information sources.

- **Evaluation**
  Golden standard
  Expert decision

Summary

We present the method of modeling uncertain decision making in SE

- How does uncertain decision making look like?
- How do we measure uncertainty in real life data?
- How do we model uncertainty in fault localization?
- Insights in other directions in SE