How to Help Graduate Students Become Independent Researchers: a Challenge and Possible Solutions

Jaime Nava and Vladik Kreinovich

Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
jenava@miners.utep.edu, vladik@utep.edu

1. Graduate Students: Typical Situation

- A PhD dissertation means that a student has successfully performed his or her own novel research.
- Usually this research results in several published papers.
- Because of this research emphasis,
  - a selection of students into a PhD program is based
  - on a student's creative ability to perform independent creative scientific research.
- Some students select their own research topics.
- However, in general, selecting a doable PhD topic is a difficult task in which a PhD advisor is indispensable.
- Once a topic is selected, both advisor and student concentrate on this topic.

2. Formulation of the Problem

- Once a topic is selected, both advisor and student concentrate on this topic.
- They try to avoid distractions that can drastically slow down the student's graduation.
- As a result, many new PhDs do not get enough training in how to select doable new research problems.
- And this is what they will need to start doing once they graduate.
- *Conclusion:* we need to train students how to select doable research problems.

3. When is the Best Time for This Training?

- *We need to* train students how to select doable research problems.
- When is the *best time* for this training?
- *In the beginning*, a student is just learning the ropes, he or she is not yet ready for this training.
- *After that*, the student is concentrating on research – and additional training will be distracting.
- However, once most research results are done, come the *last stage*, of actually writing the dissertation.
- At this *last stage*, the main emphasis is
  - not on creative new results,
  - but rather on writing down in a clear way what was done.

4. Our Proposal

- At the last stage, the main emphasis is not on new results, but on writing down what was done.
- This stage is one of the most frustrating – and least creative – parts of the graduate school.
- This stage is, in our opinion, the best time to learn the art of posing doable problems.
- We propose, at this stage, the student:
  - in parallel with writing a dissertation,
  - read new papers, talk to colleagues – with the purpose of generating new problems and new ideas,
  - and then discuss the resulting problems and ideas with the advisor.
- Why? because this is exactly what the student will do when he or she graduates with a PhD.

5. What We Expect(ed) at the End of the Training and What We Achieved

- *What we expect* at the end of such training?
- That a student will learn, from the advisor, the basic skill of
  - identifying problems and ideas
  - which are doable.
- We decided to *try* this idea.
- *Result:* several interesting new results in several areas.
- In this talk, we concentrate on new results in so-called *fuzzy logic.*
- We describe one result in detail, and the second result if time allows.

6. Background. I. Why Fuzzy Logic

- In many applications, it is important to use *expert knowledge*.
- Experts often describe their knowledge in *imprecise* ("*fuzzy*") properties like "small".
- *Example* of imprecision: for a specific size, an expert may be not fully confident whether this size is small.
- To describe such properties, *fuzzy logic* was invented.
- In fuzzy logic, each statement is characterized by a *degree* of confidence.
- Usually, this degree is taken from the interval $[0, 1]$, where:
  - 0 means absolutely false and
  - 1 means absolutely true.

7. Background. II. Fuzzy Logic Operations

- *Typical situation:*
  - *we know:* the degrees $d(A)$ and $d(B)$ of expert confidence in statements A and B;
  - *we need:* to estimate the expert's degree of confidence in composite statements like $A\&B$, $A \vee B$:

$$d(A\&B) \approx f_\&(d(A), d(B));$$

$$d(A \vee B) \approx f_\vee(d(A), d(B));$$

$$d(\neg A) \approx f_\neg(d(A)).$$

- The functions providing such estimates are called *fuzzy logic operations*:
  - and-operations (a.k.a. t-norms),
  - or-operations (a.k.a. t-conorms),
  - negation operations, etc.

8. Background. II. Fuzzy Logic Operations (cont-d)

- Fuzzy logic operations must satisfy natural properties.
- Example 1:
  - *Fact:* $A\&B$ means the same as $B\&A$.
  - *Property:* the and-operation $f_\&(a, b)$ must be commutative:
$$f_\&(a, b) = f_\&(b, a).$$
- Example 2:
  - *Fact:* $A\&(B\&C)$ means the same as $(A\&B)\&C$.
  - *Property:* the and-operation $f_\&(a, b)$ must be associative:
$$f_\&(a, f_\&(b, c)) = f_\&(f_\&(a, b), c).$$
- *Known:* there exist a complete descriptions of all the operations that satisfy such properties.

9. Formulation of the Problem

- *In principle:* we can have very *complex* fuzzy logic operations.
- *In practice:* mostly simple *algebraic* operations are used:
  - linear;
  - quadratic;
  - fractional-linear; etc.
- *Foundational challenge:* how do we classify such algebraic fuzzy operations?
- *What we prove in this talk:*
  - to classify *algebraic* fuzzy logic operations,
  - we do not need to use *all* the usual properties.

10. Motivating Result: Description of All Quadratic And-Operations

- Consider quadratic functions $f_\&: [0,1]\times[0,1]\to[0,1]$:

$$f_\&(a,b) = c_0 + c_1\cdot a + c_b\cdot b + c_{aa}\cdot a^2 + c_{ab}\cdot a\cdot b + c_{bb}\cdot b^2.$$

- *Properties:*
  - the function $f_\&(a,b)$ is *monotonic* (*non-decreasing*) in each variable;
  - $f_\&$ is *conservative* in the sense that it coincides with the usual logical operation $a\&b$ for $a,b\in\{0,1\}$:

$$f_\&(0,0) = f_\&(0,1) = f_\&(1,0) = 0; \quad f_\&(1,1) = 1.$$

- *Result* (H.T. Nguyen, V. Kreinovich): the only quadratic and-operation with these properties is $f_\&(a,b) = a\cdot b$.
- *Comment:* we did not use commutativity or associativity.

11. New Result: Description of All Quadratic Or-Operations

- Consider quadratic functions $f_\vee : [0,1] \times [0,1] \to [0,1]$:

$$f_\vee(a,b) = c_0 + c_1 \cdot a + c_b \cdot b + c_{aa} \cdot a^2 + c_{ab} \cdot a \cdot b + c_{bb} \cdot b^2.$$

- *Properties:*
  - the function $f_\vee(a,b)$ is *monotonic* (*non-decreasing*) in each variable;
  - $f_\vee$ is *conservative* in the sense that it coincides with the usual logical operation $a \vee b$ for $a, b \in \{0,1\}$:

$$f_\vee(0,0) = 0, \quad f_\vee(0,1) = f_\vee(1,0) = f_\&(1,1) = 1.$$

- *Result:* the only quadratic and-operation with these properties is $f_\vee(a,b) = a + b - a \cdot b$.
- *Comment:* we did not use commutativity or associativity.

12. Negation Operations: Usual Properties

- *Main algebraic property:*
  - *Fact:* $\neg(\neg A)$ means the same as A.
  - *Property:* the negation operation $f_\neg(a)$ must satisfy the property:
    $$f_\neg(f_\neg(a)) = a.$$
- *Monotonicity:* the more we believe in A, the less we believe in $\neg A$.
- *Conclusion:* the function $f_\neg(a)$ must be non-increasing.
- *Conservative:* for $a = 0$ ("false") and for $a = 1$ ("true"), $f_\neg(a)$ must coincide with the truth value of "not a":
  $$f_\neg(0) = 1, \quad f_\neg(1) = 0.$$

13. Description of All Quadratic Negation Operations

- Consider quadratic functions $f_\neg : [0,1] \to [0,1]$:

$$f_\neg(a) = c_0 + c_1 \cdot a + c_{aa} \cdot a^2. \qquad (1)$$

- *Properties:*
  - the function $f_\neg(a)$ satisfies the property

    $$f_\neg(f_\neg(a)) = a \text{ for all } a;$$

  - $f_\neg$ is *conservative* in the sense that it coincides with the usual logical operation $\neg a$ for $a \in \{0,1\}$:

    $$f_\neg(0) = 1, \quad f_\neg(1) = 0.$$

- *Result:* the only quadratic negation operation with these properties is $f_\neg(a) = 1 - a$.
- *Comment:* we did not use monotonicity.

14. Description of All Fractional-Linear Negation Operations

- Consider fractional-linear functions $f_\neg : [0,1] \to [0,1]$:

$$f_\neg(a) = \frac{a + b \cdot x}{c + d \cdot x}.$$

- *Properties:*
  - the function $f_\neg(a)$ satisfies the property

    $$f_\neg(f_\neg(a)) = a \text{ for all } a;$$

  - $f_\neg$ is *conservative* in the sense that it coincides with the usual logical operation $\neg a$ for $a \in \{0,1\}$:

    $$f_\neg(0) = 1, \quad f_\neg(1) = 0.$$

- *Result:* the only fractional-linear negation operation with these properties is $f_\neg(a) = 1 - a$.
- *Comment:* we did not use monotonicity.

15. Acknowledgments

This work was supported in part

- by the National Science Foundation grants HRD-0734825 and DUE-0926721, and
- by Grant 1 T36 GM078000-01 from the National Institutes of Health.

16. Need for Fuzzy "Exclusive Or" Operations

- One of the main objectives of fuzzy logic is to formalize commonsense and expert reasoning.
- People use logical connectives like "and" and "or".
- Commonsense "or" can mean both "inclusive or" and "exclusive or".
- *Example:* A vending machine can produce either a coke or a diet coke, but not both.
- In mathematics and computer science, "inclusive or" is the one most frequently used as a basic operation.
- *Fact:* "Exclusive or" is also used in commonsense and expert reasoning.
- *Thus:* There is a practical need for a fuzzy version.
- *Comment:* "exclusive or" is actively used in computer design and in quantum computing algorithms

17. "Exclusive Or" Operations: Discussion

- *Fuzzy analogue* of a classical logic operation op:
  - we know the experts' degree of belief $a = d(A)$ and $b = d(B)$ in statements A and B;
  - based on a and b, we want to estimate the degree of belief in "A op B", as $f_{\text{op}}(a, b)$.
- For op $= \&$, we get an "and"-operation (t-norm).
- For op $= \vee$, we get an "or"-operation (t-conorm).
- As usual, the fuzzy "exclusive or" operation must be an extension of the corresponding crisp operation $\oplus$.
- In the traditional 2-valued logic, $0 \oplus 0 = 1 \oplus 1 = 0$ and $0 \oplus 1 = 1 \oplus 0 = 1$.
- Thus, the desired fuzzy "exclusive or" operation $f_\oplus(a, b)$ must satisfy the same properties:

$$f_\oplus(0,0) = f_\oplus(1,1) = 0; \quad f_\oplus(0,1) = f_\oplus(1,0) = 1.$$

18. Need for the Least Sensitivity: Reminder

- One of the main ways to elicit degree of certainty d is to ask to pick a value on a scale. Example:
  - on a scale of 0 to 10, an expert picks 8, so we get
  $$d = 8/10 = 0.8;$$
  - on a scale from 0 to 8, whatever we pick, we cannot get 0.8: $6/8 = 0.75 < 0.8$; $7/8 = 0.875 > 0.8$.
  - the expert will probably pick 6, with
  $$d' = 6/8 = 0.75 \approx 0.8.$$
- *It is desirable:* that the result of the fuzzy operation not change much if we slightly change the inputs:
  $$|f(a,b) - f(a',b')| \le k \cdot \max(|a-a'|, |b-b'|),$$
  with the smallest possible k.
- Such operations are called *the least sensitive* or *the most robust.*

19. For t-Norms and t-Conorms, the Least Sensitivity Requirement Leads to Reasonable Operations

- *Known results:*
  - There is only one least sensitive t-norm ("and"-operation)
  $$f_\&(a,b)=\min(a,b).$$
  - There is also only one least sensitive t-conorm ("or"-operation)
  $$f_\vee(a,b)=\max(a,b).$$
- *What we do in this presentation:* we describe the least sensitive fuzzy "exclusive or" operation.

20. Definition of a Fuzzy Exclusive-Or Operation

- **Definition:** *A function* $f : [0,1] \times [0,1] \to [0,1]$ *is called a* fuzzy "exclusive or" operation *if*

$$f(0,0) = f(1,1) = 0 \text{ and } f(0,1) = f(1,0) = 1.$$

- *Comment:* We could also require other conditions, e.g., commutativity and associativity.
- However, our main objective is to select a single operation which is the least sensitive.
- *Fact:* The weaker the condition, the larger the class of operations that satisfy these conditions.
- *Thus:* the stronger the result that our operation is the least sensitive in this class.
- *Conclusion:* We select the weakest possible condition to make our result as strong as possible.

21. Main Result

Definition:

- *Let F be a class of functions from* $[0,1]\times[0,1]$ *to* $[0,1]$.
- *We say that a function $f\in F$ is* the least sensitive *in the class F if it satisfies the following two conditions:*
  - *for some real number k, the function f satisfies the condition*

    $$|f(a,b)-f(a',b')|\le k\cdot \max(|a-a'|,|b-b'|);$$

  - *no other function $f\in F$ satisfies this condition.*

Theorem: *In the class of all fuzzy "exclusive or" operations, the following function is the least sensitive:*

$$f_\oplus(a,b)=\min(\max(a,b),\max(1-a,1-b)).$$

22. Interpretation of the Main Result

- *Reminder:* the least sensitive operation is

$$f_\oplus(a,b) = \min(\max(a,b), \max(1-a, 1-b)).$$

- *Fact:* in 2-valued logic, "exclusive or" $\oplus$ can be described in terms of the "inclusive or" operation $\vee$ as

$$a \oplus b \Leftrightarrow (a \vee b) \,\&\, \neg(a \,\&\, b).$$

- *Natural idea:*
  - replace $\vee$ with the least sensitive "or"-operation $f_\vee(a,b) = \max(a,b)$,
  - replace & with the least sensitive "and"-operation $f_\&(a,b) = \min(a,b)$, and
  - replace $\neg$ with the least sensitive negation operation $f_\neg(a) = 1-a$,
- *Result:* we get the expression given in the Theorem.

23. Fuzzy "Exclusive Or" Operations $f(a,b)$ Which Are the Least Sensitive on Average

- *Idea:* select f so that *on average*, the change in a and b leads to the smallest possible change Δc in $c = f(a,b)$.
- *Assumption:* Δa and Δb are independent random variables with 0 mean and small variance σ^2.
- *Objective:* estimate $\Delta c = f(a+\Delta a, b+\Delta b) - f(a,b)$.
- Since Δa and Δb are small, we can keep only linear terms in the Taylor series of Δc w.r.t. Δa and Δb:

$$\Delta c \approx \frac{\partial f}{\partial a} \cdot \Delta a + \frac{\partial f}{\partial b} \cdot \Delta b.$$

- Since the variables are independent with 0 mean, the mean of Δc is also 0, and variance of Δc is equal to

$$\sigma^2(a,b) = \left(\left(\frac{\partial f}{\partial a}\right)^2 + \left(\frac{\partial f}{\partial b}\right)^2\right) \cdot \sigma^2.$$

24. Fuzzy "Exclusive Or" Operations Which Are the Least Sensitive on Average (cont-d)

- *Reminder:* for each a and b, the variance $\sigma^2(a,b)$ of Δc is equal to

$$\sigma^2(a,b) = \left(\left(\frac{\partial f}{\partial a}\right)^2 + \left(\frac{\partial f}{\partial b}\right)^2\right) \cdot \sigma^2.$$

- To get the "average" variance, it is reasonable to average this value $\sigma^2(a,b)$ over all possible a and b.

- *Resulting average value:* $I \cdot \sigma^2$, where

$$I \stackrel{\text{def}}{=} \int_{a=0}^{a=1} \int_{b=0}^{b=1} \left(\left(\frac{\partial f}{\partial a}\right)^2 + \left(\frac{\partial f}{\partial b}\right)^2\right) da\, db.$$

- *We want:* the average sensitivity to be the smallest.

- *Conclusion:* we select the function $f(a,b)$ for which the integral I takes the smallest possible value.

25. New Result: Formulation

- *Reminder:* we consider *"exclusive or"* operations $f(a,b)$, i.e., functions $f:[0,1]\times[0,1]\to[0,1]$ for which:

$$f(0,b)=b,\ \ f(a,0)=a,\ \ f(1,b)=1-b,\ \text{and}\ f(a,1)=1-a.$$

- *Main result:* among all such operations, the operation which is the least sensitive on average has the form

$$f_\oplus(a,b)=a+b-2\cdot a\cdot b.$$

- *Interpretation:*
  - the classical (2-valued) "exclusive or" operation $a\oplus b$ can be represented as $(a\vee b)\,\&\,(\neg a\vee\neg b)$;
  - use the fuzzy analogues of $\&$, $\vee$, and $\neg$ which are the least sensitive on average:

$$f_\&(a,b)=\max(p+q-1,0);\ \ f_\vee(a,b)=p+q-p\cdot q;$$

$$f_\neg(a)=1-a.$$