From Interval Computations to Processing Fuzzy Data: Main Ideas

Slides for Interval Computations class

Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA

1. Classification of Practical Problems

1) We want to *learn* what is happening in the world.

- In particular, we want to know the numerical values of different quantities:
  - distances,
  - masses,
  - charges,
  - coordinates, etc.

2) Based on these values, we would like to *predict* how the state of the world will change over time.

3) Finally, we would like to find out what *changes* we need to make to get the desired results.

- A real-life problem often involves solving subproblems of all three types.

2. Classification of Practical Problems (cont-d)

- Learning the current state of the world and predicting the future are usually classified as *science*.
- The tasks of finding the appropriate change are usually classified as *engineering*.
- Measuring the flow of Rio Grande and predicting how this flow will change over time are problems of science.
- Finding the best way to change this flow (e.g., by building a levee) is a problem of engineering.

3. First Class of Practical Problems: Learning the State of the World

- Let us start with the first class of practical problems: the problem of learning the state of the world.
- In particular, that we want to know the numerical values of different quantities y that characterize this state.
- Some quantities y we can simply directly measure, e.g.:
  - when we want to know the current state of a patient in a hospital,
  - we can measure the patient's body temperature, blood pressure, weight, and many other important characteristics.
- In some situations, we do not even need to measure:
  - we can simply ask an expert,
  - and the expert will provide us with an (approximate) value $\widetilde{y}$ of the quantity y.

4. Sometimes, We Cannot Directly Measure

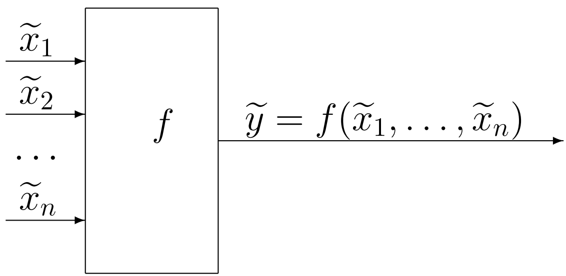
- However, many other quantities of interest are difficult or even important to measure or estimate directly.
- Examples of such quantities include the amount of oil in a given well or a distance to a star.
- Since we cannot directly measure these quantities, the only way to get information about them is:
  - to measure (or ask an expert to estimate) some other easier-to-measure quantities $x_1, \ldots, x_n$,
  - then to estimate y based on the measured values $\widetilde{x}_i$ of these auxiliary quantities x_i.
- For example, to estimate the amount of oil in a given well, we perform *seismic* experiments:
  - we set up small explosions at some locations and
  - measure the resulting seismic waves at different distances from the location of the explosion.

5. Sometimes, We Can't Measure (cont-d)

- To find the distance to a faraway star, we measure:
  - the direction to the star from different locations on Earth (and/or in different seasons) and
  - the coordinates of (and the distances between) the locations of the corresponding telescopes.
- To estimate the value of the desired quantity y, we must know the relation:
  - between y and
  - the easier-to-measure (or easier-to-estimate) quantities $x_1, \ldots, x_n$.
- Specifically, we want to use the estimates of x_i to come up with an estimate for y.

6. Need for an Algorithm

- Thus, the relation between y and x_i must be given in the form of an *algorithm* $f(x_1, \ldots, x_n)$ which:
  - transforms the values of x_i
  - into an estimate for y.
- Once we know this algorithm f and the measured values $\widetilde{x}_i$ of the auxiliary quantities, we can estimate y:



Home Page
Title Page
Page 7 of 103
Go Back
Full Screen
Close
Quit

7. Algorithms Can Have Different Complexity

- In different practical situations, we have algorithms f of different complexity.
- To find the distance to star, we can usually have an explicit trigonometric formula.
- In this case, f is a simple formula.
- To find the amount of oil, we must numerically solve a complex partial differential equation.
- In this case, f is a complex iterative algorithm for solving this equation.
- When the values x_i are obtained by measurement, this 2-stage process does involve measurement.
- To distinguish it from *direct* measurements, this 2-stage process is called an *indirect* measurement.

8. Possible Exam Questions

- What is indirect measurement?
- Why do we need indirect measurements? Provide examples.

9. Second Class of Practical Problems: Predicting the Future State of the World

- Once
  - we know the values of the quantities $y_1, \ldots, y_m$ characterizing the current state of the world,
  - we can start predicting the future state of the world,
  - i.e., the future values of these quantities.
- To be able to predict the future value z of each of these quantities, we must know:
  - how exactly this value z
  - depends on the current values $y_1, \ldots, y_m$.
- Specifically, we want to use the known estimates $\widetilde{y}_i$ for y_i to come up with an estimate for z.

10. Need for an Algorithm

- Thus, the relation between z and y_i must be given in the form of an *algorithm* $g(y_1, \ldots, y_m)$ which:
  - transforms the values of y_i
  - into an estimate for z.
- Once:
  - we know this algorithm g and the estimates $\widetilde{y}_i$ for the current values of the quantities,
  - we can estimate z as $\widetilde{z} = g(\widetilde{y}_1, \ldots, \widetilde{y}_n)$.
- The corresponding algorithm g can be very complicated and time-consuming.
- This is, e.g., how weather is predicted now:
  - weather prediction requires so many computations
  - that it can only be performed on fast supercomputers.

11. The General Notion of Data and Knowledge Processing

- So far, we have analyzed two different classes of practical problems:
  - the problem of *learning* the current state of the world (i.e., the problem of indirect measurement),
  - and the problem of *predicting* the future state of the world.
- From the *practical* viewpoint, these two problems are drastically different.
- However, as we have seen, from the *computational* viewpoint, these two problems are very similar.

12. Data and Knowledge Processing (cont-d)

- In both problems:
  - we start with the estimates $\widetilde{x}_1, \ldots, \widetilde{x}_n$ for the quantities $x_1, \ldots, x_n$, and then
  - we apply the known algorithm f to these estimates,
  - this results in an estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ for the desired quantity y.
- When the inputs come from measurements, they constitute *data*.
- The computational part of the corresponding procedure is called *data processing*.
- When the inputs come from experts, it constitutes *knowledge*.
- The computational part of the corresponding procedure is called *knowledge processing*.

13. Possible Exam Questions

- What is data processing? Give examples.
- What is knowledge processing? Give examples.
- Why do we need data processing and knowledge processing?
- What is the difference between data processing and knowledge processing?

14. Third Class of Practical Problems: How to Change the World

- Once:
  - we know the current state of the world and
  - we know how to predict the consequences of different decisions (designs, etc.),
  - it is desirable to find the decision (design, etc.) which guarantees the given results.
- Depending on what we want from this design:
  - we can subdivide all the problems from this class
  - into two subclasses.
- In both subclasses, the design must satisfy some constraints.
- Thus, we are interested in finding a design that satisfies all these constraints.

15. Two Subclasses

- In some practical situations, satisfaction of all these constraints is all we want.
- In general, there may be several possible designs which satisfy given constraints.
- In the problems from the first subclass, we do not have any preferences for one of these designs.
- Any one of them will suffice.
- Such problems are called the problems of *constraint satisfaction*.

16. Two Subclasses (cont-d)

- In other practical situations, we do have a clear preference between different designs x.
- This preference is usually described in terms of an *objective function* $F(x)$:
  - more preferable designs x
  - correspond to larger values of $F(x)$.
- In such situation:
  - among all the designs which satisfy given constraints,
  - we would like to find a design x for which the value $F(x)$ of the given objective function is the largest.
- Such problems are called *optimization problems*.

17. Possible Exam Questions

- What is constraint satisfaction? Why do we need it? Provide examples.
- What is optimization? Why do we need it? Provide examples.
- What is the difference between constraint satisfaction and optimization?

18. Need to Process Fuzzy Uncertainty

- In many practical situations, we only have expert estimates for the inputs x_i.
- Sometimes, experts provide guaranteed bounds on the values of x_i.
- Sometimes, they even provide the probabilities of different values within these bounds.
- However, such cases are rare.
- Usually, the experts' opinions are described by imprecise ("fuzzy") words from natural language.
- For example, an expert can say that:
  - the value x_i of the i-th quantity is approximately equal to 1.0,
  - with an accuracy most probably of about 0.1.

19. Need to Process Fuzzy Uncertainty (cont-d)

- Based on such "fuzzy" information, what can we say about $y = f(x_1, \ldots, x_n)$?
- The need to process such "fuzzy" information was first emphasized in the early 1960s by Lotfi Zadeh.
- He designed a special technique for such processing called *fuzzy logic*.

20. Possible Exam Question

- What is fuzzy processing? Why is it called "fuzzy"?
- Why do we need fuzzy processing? Give an example.
- Who was the first to develop fuzzy techniques?

21. Processing Fuzzy Uncertainty: Main Idea

- A value y is a reasonable value of the desired quantity if $y = f(x_1, \ldots, x_n)$ for some reasonable values x_i.
- In other words: if for some values $x_1, \ldots, x_n$, x_1 is reasonable, x_2 is reasonable, …, and $y = f(x_1 \ldots, x_n)$.
- Thus, to describe to what extent different values of y are reasonable, we must be able:
  - to describe to what extent (to what degree) different values of x_i are reasonable, and
  - to combine these degrees into the desired degree of belief in reasonability of y.

22. Degrees of Belief

- Let us first introduce the basic concept of degrees of belief.
- For example, we would like to estimate to what extent the value $x_i = 0.89$ is consistent with the statement:
  - *the value x_i of the i-th quantity is approximately equal to 1.0,*
  - *with an accuracy most probably about 0.1.*
- In the absence of uncertainty, every statement is either true or false.
- In the computer, "true" is usually represented as 1, and "false" as 0.

23. Degrees of Belief (cont-d)

- It is thus reasonable to use numbers between 0 and 1 to represent levels of confidence intermediate between:
  - the absolute confidence that a given statement is true and
  - the absolute confidence that a given statement is false.
- How do we determine this degree of confidence?
- For example, we can:
  - ask several (N) experts whether $x_i = 0.89$ is consistent with the above statement, and
  - if M of them reply "yes", take the ratio M/N as the desired degree of confidence.

24. Degrees of Belief (cont-d)

- If we do not have access to numerous experts, we can simply:
  - ask the only available expert to describe his or her degree of confidence
  - by marking a number on a scale from 0 to N (e.g., on a scale from 0 to 5).
- If an expert marks his or her degree as M, we take the ratio M/N as the desired degree of confidence.

25. Possible Exam Questions

- Why degree of belief 1 usually corresponds to absolute confidence, and 0 to absolute lack of confidence?
- Describe two ways to elicit degrees of belief from experts.
- Suppose that we are trying to formalize a medical rule about small tumors.
- Out of 20 experts, 16 think that 0.3 mm is small.
- What will be the corresponding degree of belief that 0.3 mm is small?
- An expert marks her degree of confidence that 0.3 mm is small as 3.5 on a 0 to 5 scale.
- What is the resulting degree of belief?

26. Membership Functions

- To formally describe the original expert's statement S about x_i, we need to know:
  - for every real number x_i,
  - the degree $\mu_S(x_i)$ to which this real number is consistent with this statement S.
- In practice, we can only ask finitely many questions.
- So, no matter how many questions we ask, we can only find $\mu_S(x_i)$ for finitely many values x_i.
- To estimate the values $\mu_S(x_i)$ for all other real numbers x_i, we must use interpolation and extrapolation.
- Usually, a piece-wise linear interpolation is used.
- Sometimes, a more sophisticated procedure is applied: e.g., a piecewise quadratic interpolation.

27. Membership Functions (cont-d)

- The function $\mu_S(x_i)$ which is obtained by this interpolation is called a *membership function*.
- This function describes:
  - for every real number x_i,
  - the degree $\mu_S(x_i)$ to which this real number is consistent with this statement S.

28. Possible Exam Questions

- What is a membership function?
- How can we determine the membership function?
- We know that $\mu(0) = 1$ and $\mu(2) = 0$. Use linear interpolation to find $\mu(x)$ for $x \in [0, 2]$.
- *Reminder:* if we know that $f(x_1) = y_1$ and $f(x_2) = y_2$, then as an estimate for $f(x)$, we take

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1} \cdot (x - x_1).$$

- In our case, $x_1 = 0$, $x_2 = 2$, $y_1 = 1$, $y_2 = 0$, so

$$y = 1 + \frac{0 - 1}{2 - 0} \cdot (x - 0) = 1 - x/2.$$

- *Now students have to work:* we know that $\mu(-2) = 0$, $\mu(0) = 1$, use linear interpolation to find $\mu(-1.6)$.

29. Need for "And"- and "Or"-Operations: t-Norms and t-Conorms

- As we have mentioned earlier:
  - we are not directly interested
  - in the degree to which a given real number x_i is consistent with the expert's knowledge S_i about the i-th input.
- We are mainly interested in the degree to which:
  - the value x_1 is consistent with the knowledge about the first input *and*
  - the value x_2 is consistent with the knowledge about the second input *and*
  - … *and*
  - the value x_n is consistent with the knowledge about the n-th input.

30. "And"- and "Or"-Operations (cont-d)

- In principle, we can determine the degree of belief in such a composite statement by asking an expert:
  - for each possible combination of values $x_1, x_2, \ldots, x_n$,
  - what is the degree to which this combination is consistent with all the available expert knowledge.
- However, as we have mentioned earlier:
  - even for a single input,
  - we cannot realistically elicit degrees of confidence about too many values.
- If we consider N possible values of each input, then:
  - we would need to elicit the expert's degree of confidence
  - about $N^n \gg N$ possible combinations.

31. “And”- and “Or”-Operations (cont-d)

- Eliciting N^n values is even less realistic than eliciting N values.
- We cannot directly elicit the expert's degree of confidence in all composite statements.
- So, a natural idea is:
  - to estimate the degree of confidence in the composite statement
  - based on the degrees of confidence in individual statements
  - such as “x_i is consistent with the expert's knowledge S_i about the i-th input.”
- How can we come up with such an estimate?

32. "And"- and "Or"-Operations (cont-d)

- Let us reformulate this estimation problem:
  - we know the expert's degree of confidence in statements A_1, A_2, ..., A_n, and
  - we want to estimate the expert's degree of confidence in a composite statement

$$A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n,$$

  - i.e., "A_1 and A_2 and ... and A_n".
- Here, e.g., $A_1 \,\&\, A_2 \,\&\, A_3$ can be represented as $(A_1 \,\&\, A_2) \,\&\, A_3$.
- So, it is sufficient to solve this estimation problem for the case of two statements.

33. "And"- and "Or"-Operations (cont-d)

- Once we have a solution for this particular case, we will then be able to solve the general problem as well:
  - first, we apply the two-statement solution to the degrees of certainty in A_1 and A_2,
  - we thus get an estimate for the expert's degree of certainty in $A_1 \,\&\, A_2$;
  - then, we apply the same solution to the degrees of certainty in $A_1 \,\&\, A_2$ and A_3,
  - we thus get an estimate for the expert's degree of certainty in $A_1 \,\&\, A_2 \,\&\, A_3$;
  - after that, we apply the same solution to the degrees of certainty in $A_1 \,\&\, A_2 \,\&\, A_3$ and A_4,
  - we thus get an estimate for the expert's degree of certainty in $A_1 \,\&\, A_2 \,\&\, A_3 \,\&\, A_4$;
  - etc.

34. "And"- and "Or"-Operations (cont-d)

- Eventually, we will get the degree of confidence in the desired composite statement $A_1 \,\&\, A_2 \,\&\, \ldots \,\&\, A_n$.
- Thus, we need a procedure that would transform:
  - the degree of belief d_1 in a statement A_1 and
  - the degree of belief d_2 in a statement A_2
  - into a (reasonable) estimate for a degree of belief in a composite statement $A_1 \,\&\, A_2$.
- Let us denote the estimate corresponding to given values d_1 and d_2 by $f_\&(d_1, d_2)$.
- The procedure $f_\&$ is called an *"and"-operation*, or, for historical reasons, a *t-norm*.
- This procedure maps:
  - degrees of belief d_1 and d_2 in statements A_1 and A_2
  - into a degree of belief $d = f_\&(d_1, d_2)$ in $A_1 \,\&\, A_2$.

35. "Or"-Operations

- Similarly:
  - to estimate the degree of belief in a composite statement $A_1 \vee A_2$ ("A_1 or A_2"),
  - we need a procedure $f_\vee$ that maps:
    * degrees of belief d_1 and d_2 in statements A_1, A_2
    * into a degree of belief $d = f_\vee(d_1, d_2)$ in $A_1 \vee A_2$.
- Such a procedure is called an *"or"-operation*.
- In logic:
  - "or" is a kind of dual to "and", and
  - "or"-operation can be viewed as a dual to an "and"-operation (t-norm).
- Because of this duality, an "or"-operation is also called a *t-conorm*.

36. Possible Exam Questions

- What is an "and"-operation?
- Why do we need "and"-operations? Why cannot we explicitly elicit all the degrees from the experts?
- What is an "or"-operation?
- Why do we need "or"-operations?

37. Properties of "And"- and "Or"-Operations

- From the intended meaning of the "and"- and "or"-operations, we can deduce reasonable properties.
- For example, intuitively, "A_1 and A_2" means the same as "A_2 and A_1".
- Thus, it is reasonable to require that:
  - our estimate $f_\&(d_1, d_2)$ for the degree of confidence in "A_1 and A_2"
  - should be the same our estimate $f_\&(d_2, d_1)$ for the degree of confidence in "A_2 and A_1".
- In other words, we must have $f_\&(d_1, d_2) = f_\&(d_2, d_1)$ for all possible values of d_1 and d_2.
- In mathematical terms, this means that the function $f_\&$ must be *commutative*.

38. Associativity

- Similarly, "(A_1 and A_2) and A_3" means the same as "A_1 and (A_2 and A_3)".
- Indeed, both mean the same as "A_1 and A_2 and A_3".
- For each "and"-operation $f_\&$, the expression "(A_1 and A_2) and A_3" means that we:
  - first estimate the degree of belief in "A_1 and A_2" as $f_\&(d_1, d_2)$, and
  - then estimate the degree of belief in "(A_1 and A_2) and A_3" as $f_\&(f_\&(d_1, d_2), d_3)$.
- Similarly, the expression "A_1 and (A_2 and A_3)" means that we:
  - first estimate the degree of belief in "A_2 and A_3" as $f_\&(d_2, d_3)$, and
  - then estimate the degree of belief in "A_1 and (A_2 and A_3)" as $f_\&(d_1, f_\&(d_2, d_3))$.

39. Associativity and Other Properties

- Since the expressions are equivalent, it is reasonable to require that these estimates coincide.
- So, we must have $f_\&(f_\&(d_1, d_2), d_3) = f_\&(d_1, f_\&(d_2, d_3))$ for all possible values of d_1, d_2, and d_3.
- In mathematical terms, this means that the function $f_\&$ must be *associative*.
- There are several other reasonable properties of "and"-operations.
- For example, "A_1 and A_2" implies A_1.
- So, our degree of belief in the composite statement "A_1 and A_2" cannot exceed our degree of belief in A_1.

40. Other Properties (cont-d)

- Thus, it is reasonable to require that:
  - the estimate $f_\&(d_1, d_2)$ for this degree of belief
  - should also not exceed our degree of belief d_1 in the statement A_1.
- In other words, we should have $f_\&(d_1, d_2) \le d_1$ for all possible values of d_1 and d_2.
- If A_1 is absolutely true (i.e., $d_1 = 1$), then intuitively:
  - the composite statement "A_1 and A_2"
  - has exactly the same truth value as A_2.
- Thus, it is reasonable to require that $f_\&(1, d_2) = d_2$ for all possible values of d_2.

41. Other Properties (cont-d)

- On the other hand:
  - if A_1 is absolutely false (i.e., $d_1 = 0$),
  - then the composite statement "A_1 and A_2" should also be absolutely false,
  - no matter how much we may believe in A_2.
- Thus, it is reasonable to require that $f_\&(0, d_2) = 0$ for all possible values of d_2.
- Finally:
  - if, due to a new evidence, our degree of belief in one of the statements A_1 and A_2 increases,
  - the resulting degree of belief in "A_1 and A_2" will either increase or stay the same,
  - but it cannot decrease.

42. Other Properties (cont-d)

- Thus, it is it is reasonable to require that the operation $f_\&$ be *monotonic* in the sense that:
  - if $d_1 \le d'_1$ and $d_2 \le d'_2$,
  - then $f_\&(d_1, d_2) \le f_\&(d'_1, d'_2)$.
- All these properties are indeed required of an "and"-operation (t-norm).
- Similarly, it is reasonable to require that an "or"-operation (t-conorm) $f_\vee$ should be:
  - commutative,
  - associative,
  - monotonic, and
  - satisfy the conditions $d_1 \le f_\vee(d_1, d_2)$, $f_\vee(1, d_2) = 1$, and $f_\vee(0, d_2) = d_2$ for all d_1 and d_2.

43. Possible Exam Questions

- Why should an "and"-operation be commutative? Explain.
- Why should an "and"-operation be associative? Explain.
- Why should an "or"-operation be commutative? Explain.
- Why should an "or"-operation be associative? Explain.

44. Simplest "And"- and "Or"-Operations: Derivation

- There exist many different "and"- and "or"-operations which satisfy the above properties.
- In some applications such as fuzzy control, it is crucial to select appropriate operations:
  - we can use the additional degrees of freedom
  - to tune the resulting control and
  - thus make it an even better fit for the corresponding objective function.
- However, in knowledge processing:
  - when we are very uncertain about the inputs,
  - it is probably more reasonable to select the simplest "and"- and "or"-operations
  - which are consistent with the expert knowledge.

45. Simplest "And"-Operation (cont-d)

- To select such operations, it makes sense to consider yet another property of "and" and "or" – that:
  - for every statement A,
  - "A and A" means the same as simply A.
- Thus, it is reasonable to require that:
  - for every statement A with a degree of confidence d,
  - our estimate $f_\&(d, d)$ of the expert's degree of confidence in "A and A"
  - should be the same as the original degree of confidence d in the original statement A.
- Thus, it is reasonable to require that $f_\&(d, d) = d$ for all possible values of d.
- In mathematical terms, this means that the function $f_\&$ must be *idempotent*.

46. Simplest "Or"-Operation

- Similarly, "A or A" means the same as simply A.
- So, it is reasonable to require that $f_\vee(d, d) = d$ for all possible values of d.
- So, the function $f_\vee$ must also be *idempotent*.

47. Uniqueness of Idempotent "And"-Operation

- Let us first show that the only idempotent "and"-operation is $f_\&(d_1, d_2) = \min(d_1, d_2)$.
- Without loss of generality, let us assume that $d_1 \le d_2$.
- In this case, the desired equality takes the form

$$f_\&(d_1, d_2) = d_1.$$

- Since the operation $f_\&$ is idempotent, we have

$$f_\&(d_1, d_1) = d_1.$$

- Due to $d_1 \le d_2$, monotonicity implies that $f_\&(d_1, d_1) \le f_\&(d_1, d_2)$, hence $d_1 \le f_\&(d_1, d_2)$.
- On the other hand, for an "and"-operation, we always have $f_\&(d_1, d_2) \le d_1$.
- So, we can conclude that $f_\&(d_1, d_2) = d_1$, i.e., indeed, $f_\&(d_1, d_2) = \min(d_1, d_2)$.

48. Uniqueness of Idempotent "Or"-Operation

- Let us now prove that the only idempotent "or"-operation is $f_\vee(d_1, d_2) = \max(d_1, d_2)$.
- Without loss of generality, let us again assume that
$$d_1 \le d_2.$$
- In this case, the desired equality takes the form
$$f_\vee(d_1, d_2) = d_2.$$
- Since the operation $f_\vee$ is idempotent, we have
$$f_\vee(d_2, d_2) = d_2.$$
- Due to $d_1 \le d_2$, monotonicity implies that $f_\vee(d_1, d_2) \le f_\vee(d_2, d_2)$, hence $f_\vee(d_1, d_2) \le d_2$.
- On the other hand, for an "or"-operation, we always have $d_2 \le f_\vee(d_1, d_2)$, so $f_\vee(d_1, d_2) = d_2$, i.e.,
$$f_\vee(d_1, d_2) = \max(d_1, d_2).$$

49. Simplest "And"- and "Or"–Operations: General

- $f_\&(d_1, d_2) = \min(d_1, d_2)$ and $f_\vee(d_1, d_2) = \max(d_1, d_2)$ were actually the first designed by L. Zadeh.
- They are still actively used in various applications of fuzzy techniques.
- If the expert's degree of confidence in A_1 is 0.6 and in A_2 is 0.6, then:
  - what is the expert's degree of confidence in $A \,\&\, B$?
  - what is the expert's degree of confidence in $A \vee B$?

50. Possible Exam Questions

- Why should an "and"-operation be idempotent? Explain.
- Why should an "or"-operation be idempotent? Explain.
- What is the only idempotent "and"-operation? Who invented it? Is it used in practice?
- What is the only idempotent "or"-operation? Who invented it? Is it used in practice?

51. Fuzzy Data Processing

- Let us apply the above simple operations to knowledge processing, i.e., to processing fuzzy uncertainty.
- In this situation, we know an algorithm $y = f(x_1, \ldots, x_n)$.
- This algorithm relates:
  - the value of the desired difficult-to-estimate quantity y
  - with the values of easier-to-estimate auxiliary quantities $x_1, \ldots, x_n$.
- We also have expert knowledge about each of the quantities x_i.
- For each i, this knowledge is described in terms of the corresponding membership function $\mu_i(x_i)$.

52. Fuzzy Data Processing (cont-d)

- For each i and for each value x_i, the value $\mu_i(x_i)$ is:
  - the degree of confidence
  - that this value is indeed a possible value of the i-th quantity.
- Based on this information, we want to find the membership function $\mu(y)$ which describes:
  - for each real number y,
  - the degree of confidence that this number is a possible value of the desired quantity.

53. Fuzzy Data Processing (cont-d)

- As we have mentioned earlier, y is a possible value of the desired quantity if:
  - for some values $x_1, \ldots, x_n$,
  - x_1 is a possible value of the first input quantity, and
  - x_2 is a possible value of the 2nd input quantity,
  - $\ldots$, and
  - we have $y = f(x_1 \ldots, x_n)$.
- We know:
  - that the degree of confidence that x_1 is a possible value of the first input quantity is equal to $\mu_1(x_1)$,
  - that the degree of confidence that x_2 is a possible value of the 2nd input quantity is equal to $\mu_2(x_2)$,
  - etc.

54. Fuzzy Data Processing (cont-d)

- The degree of confidence $d(y, x_1, \ldots, x_n)$ in an equality $y = f(x_1 \ldots, x_n)$ is, of course, equal:
  - to 1 if this equality holds, and
  - to 0 if this equality does not hold.
- We have already agreed to represent "and" as min.
- Thus, for each combination of values $x_1, \ldots, x_n$,
  - the degree of confidence in a composite statement

    "x_1 is a possible value of the first input quantity, and x_2 is a possible value of the second input quantity, ..., and $y = f(x_1 \ldots, x_n)$"

  - is equal to

$$\mu(y) = \min(\mu_1(x_1), \mu_2(x_2), \ldots, d(y, x_1, \ldots, x_n)).$$

55. Fuzzy Data Processing (cont-d)

- We got the expression

$$\mu(y) = \min(\mu_1(x_1), \mu_2(x_2), \ldots, d(y, x_1, \ldots, x_n)).$$

- We can simplify this expression if we consider two possible cases:
  - when the equality $y = f(x_1 \ldots, x_n)$ holds, and
  - when this equality does not hold.
- When the equality $y = f(x_1 \ldots, x_n)$ holds, we get $d(y, x_1, \ldots, x_n) = 1$.
- Thus, the above degree of confidence is equal to

$$\min(\mu_1(x_1), \mu_2(x_2), \ldots, \mu_n(x_n)).$$

- When the equality $y = f(x_1 \ldots, x_n)$ does not hold, we get $d(y, x_1, \ldots, x_n) = 0$.

56. Fuzzy Data Processing (cont-d)

- Thus, the above degree of confidence is equal to 0.
- We want to combine these degrees of belief into a single degree of confidence that

  "for some values $x_1, \ldots, x_n$, x_1 is a possible value of the first input quantity, and x_2 is a possible value of the first input quantity, ..., and $y = f(x_1 \ldots, x_n)$".

- The words "for some values $x_1, \ldots, x_n$" means that the following composite property holds:
  - either for one combination of real numbers $x_1, \ldots, x_n$,
  - or for another combination
  - until we exhaust all (infinitely many) such combinations.

57. Fuzzy Data Processing (cont-d)

- We have already agreed to represent "or" as max.
- Thus, the desired degree of confidence $\mu(y)$ is equal:
  - to the maximum of the degrees
  - corresponding to different combinations $x_1, \ldots, x_n$.
- Since we have infinitely many possible combinations, the maximum is not necessarily attained.
- So we should, in general, consider supremum instead of maximum:

$$\mu(y) = \sup \min(\mu_1(x_1), \mu_2(x_2), \ldots, d(y, x_1, \ldots, x_n)).$$

- Here, the supremum is taken over all possible combinations.
- We know that the maximized degree is non-zero only when $y = f(x_1 \ldots, x_n)$.

58. Zadeh's Extension Principle

- So, it is sufficient to only take supremum over such combinations.
- For such combinations, we can omit the term $d(y, x_1, \ldots, x_n)$ in the maximized expression.
- So, we arrive at the following formula for $\mu(y)$:

$$\sup\{\min(\mu_1(x_1), \ldots, \mu_n(x_n)) : y = f(x_1, \ldots, x_n)\}.$$

- This formula describes a reasonable way to extend an arbitrary data processing algorithm $f(x_1, \ldots, x_n)$:
  - from real-valued inputs
  - to a more general case of fuzzy inputs.
- It was first proposed by L. Zadeh and is thus called *Zadeh's extension principle.*
- This is the main formula that describes knowledge processing under fuzzy uncertainty.

59. Fuzzy Data Processing Can Be Reduced to Interval Computations

- We will show that from the computational viewpoint:
  - the application of this formula
  - can be reduced to interval computations.
- This reduction is how knowledge processing under fuzzy uncertainty is usually done.

60. Possible Exam Questions

- Describe the main problem of fuzzy data processing in precise terms.
- Formulate Zadeh's extension principle.
- *For extra credit:* explain how Zadeh's extension principle is derived.

61. An Alternative Set Representation of a Membership Function: Alpha-Cuts

- In some situations, an expert knows exactly which values of x_i are possible and which are not.
- In this situation, the expert's knowledge can be naturally represented by the *set* of all possible values.
- In general, the expert's knowledge is fuzzy:
  - we may still have some values about which the expert 100% believes that they are possible, and
  - we may still have some values about which the expert 100% believes that they are impossible, but
  - in general, the expert is not 100% confident about which values of x_i are possible and which are not.

62. Alpha-Cuts (cont-d)

- For example, a geophysicist may be confident:
  - that the density x_i of some mineral can take on values ranging from 3.4 to 3.7 g/cm^3, and
  - she may know that values smaller than 3.0 or larger than 4.0 are absolutely impossible, but
  - she is not sure whether values from 3.0 to 3.4 or from 3.7 to 4.0 are indeed realistically possible.
- As we have mentioned, the ultimate purpose of the measurements and estimates is to make decisions.

63. Alpha-Cuts (cont-d)

- In the geophysical example, we have measured the density at a certain depth, and we need to decide:
  - whether it is possible that we have the desired mineral
  - in this case we should undertake more measurements, or
  - whether it is not possible that we have the desired mineral
  - in this case we should not waste our resources on this region and move to more promising regions.
- In practice, decisions are made under uncertainty:
  - if we only have a fuzzy expert description of possible values, as a membership function $\mu_S(x_i)$,
  - which values x_i should we then classify as possible ones and which as impossible?

64. Alpha-Cuts (cont-d)

- Under uncertainty, a reasonable idea is to select a threshold $\alpha \in (0, 1]$.
- In this case,
  - all the values x_i for which the expert's degree of confidence is strong enough ($\mu_S(x_i) \geq \alpha$)
  - are classified as possible.
- Similarly:
  - all the values x_i for which the expert's degree of confidence is not sufficiently strong ($\mu_S(x_i) < \alpha$)
  - are classified as impossible.
- The resulting set of possible elements is called the α-*cut* of the membership function $\mu_S(x_i)$:

$$\mathbf{x}_i(\alpha) \stackrel{\text{def}}{=} \{x_i : \mu_S(x_i) \geq \alpha\}$$

65. How to Select α

- The choice of a threshold α depends on the practical problem.
- For example:
  - if we are looking for a potentially very valuable mineral deposit,
  - then it makes sense to continue prospecting even when our degree of confidence is not very high.
- In this case, it makes sense to select a reasonably small threshold α.
- On the other hand:
  - if the potential benefit is not high and our resources are limited,
  - it makes sense to limit our search to highly promising regions, i.e., to select a high threshold α.

66. Alpha-Cuts (cont-d)

- To adequately describe the expert knowledge irrespective of an application, we therefore need to know:
  - the α-cuts
  - corresponding to different thresholds α.
- Each α-cut $\mathbf{x}_i(\alpha)$ describes the set of values which are possible with degree of confidence at least α.
- By definition, α-cuts corresponding to different α are *nested*:
  - when $\alpha \leq \alpha'$,
  - then $\mu_S(x_i) \geq \alpha'$ implies $\mu_S(x_i) \geq \alpha$ and thus,

$$\mathbf{x}_i(\alpha') = \{x_i : \mu_S(x_i) \geq \alpha'\} \subseteq \mathbf{x}_i(\alpha) = \{x_i : \mu_S(x_i) \geq \alpha\}.$$

67. From Alpha-Cuts to Membership Functions

- It is worth mentioning that:
  - if we know the α-cuts $\mathbf{x}_i(\alpha) = \{x_i : \mu_S(x_i) \geq \alpha\}$ corresponding to all possible $\alpha \in (0, 1]$,
  - then we can uniquely reconstruct the corresponding membership function $\mu_S(x_i)$.
- The possibility for such a reconstruction follows from the fact that:
  - every real number r
  - is equal to the largest largest value α for which

$$r \geq \alpha.$$

- In particular, for every x_i, the value $\mu_S(x_i)$ is equal to the largest value α for which $\mu_S(x_i) \geq \alpha$.
- By definition of the α-cut, the inequality $\mu_S(x_i) \geq \alpha$ is equivalent to $x_i \in \mathbf{x}_i(\alpha)$.

68. From Alpha-Cuts to Membership Functions (cont-d)

- Thus, for every x_i, the value $\mu_S(x_i)$ can be reconstructed as the largest value α for which $x_i \in \mathbf{x}_i(\alpha)$.
- So, we can alternatively view a membership function as a nested family of α-cuts.

69. Possible Exam Questions

- What is an α-cut?
- How to select α? Give examples.
- What is the relation between α-cuts corresponding to different α?
- For $\mu(x) = \max(0, 1 - |x|)$ and $\alpha = 0.6$, what is the α-cut?
- Here, $1 - |x| \geq 0.6$ means $|x| \leq 1 - 0.6 = 0.4$, i.e., that $-0.4 \leq x \leq 0.4$.
- Thus, $\mathbf{x}(0.6) = [-0.4, 0.4]$.
- *For students to do on their own:* what is $\mathbf{x}(0.6)$ for $\mu(x) = \max(0, 1 - |x - 1|)$?

70. Fuzzy Numbers

- In most practical situations, the membership function:
  - starts with 0,
  - continuously increases until a certain value and then
  - continuously decreases to 0.
- Such membership function describe usual expert's expressions such as:
  - "small",
  - "medium",
  - "reasonably high",
  - "approximately equal to a with an error about σ", etc.

71. Fuzzy Numbers and Intervals

- Membership functions of this type are actively used in expert estimates of number-valued quantities.
- So, they are usually called *fuzzy numbers*.
- For a fuzzy number $\mu_i(x_i)$, every α-cut $\mathbf{x}_i(\alpha)$ is an interval.
- Thus, a fuzzy number can be viewed as:
  - a nested family of intervals $\mathbf{x}_i(\alpha)$
  - corresponding to different degrees of confidence α.

72. Possible Exam Questions

- What is a fuzzy number? Give examples.
- How to describe a fuzzy number in terms of α-cuts?

73. Simplest "And"- and "Or"-Operations: Reformulation in Terms of Sets and Alpha-Cuts

- The main formulas for fuzzy computations (i.e., for processing fuzzy data) were derived by using:
  - the "and"-operation $f_\&(d_1, d_2) = \min(d_1, d_2)$
  - and the "or"-operation $f_\vee(d_1, d_2) = \max(d_1, d_2)$.
- Let us reformulate these "and"- and "or"-operations in terms of α-cuts.
- Let us assume that we have two properties A and B which are described:
  - by the membership functions $\mu_A(x)$ and $\mu_B(x)$ and,
  - correspondingly, by the α-cuts

$$\mathbf{x}_A(\alpha) = \{x : \mu_A(x) \ge \alpha\} \text{ and } \mathbf{x}_B(\alpha) = \{x : \mu_B(x) \ge \alpha\}.$$

74. "And"- and "Or"-Operations in Terms of Alpha-Cuts

- If we use the simplest "and"-operation $f_\&(d_1, d_2) = \min(d_1, d_2)$, then:
  - the composite property $A \,\&\, B$ ("A and B")
  - is described by the membership function $\mu_{A\,\&\,B}(x) = \min(\mu_A(x), \mu_B(x))$.
- What are the α-cuts corresponding to this membership function $\mathbf{x}_{A\,\&\,B}(\alpha) = \{x : \mu_{A\,\&\,B}(x) \geq \alpha\}$?
- The minimum of two real numbers:
  - is greater than or equal to α if and only if
  - both of these numbers are greater than or equal to α.
- Thus, $\mu_{A\,\&\,B}(x) = \min(\mu_A(x), \mu_B(x)) \geq \alpha$ is equivalent to "$\mu_A(x) \geq \alpha$ and $\mu_B(x) \geq \alpha$".

75. "And"- and "Or"-Operations in Terms of Alpha-Cuts (cont-d)

- So, the set $\mathbf{x}_{A\,\&\,B}(\alpha)$ of all x for which $\mu_{A\,\&\,B}(x) = \min(\mu_A(x), \mu_B(x)) \ge \alpha$ is an intersection of:
  - the set of all x for which $\mu_A(x) \ge \alpha$ and
  - the set of all x for which $\mu_B(x) \ge \alpha$.
- In other words, for every α, we have

$$\mathbf{x}_{A\,\&\,B}(\alpha) = \mathbf{x}_A(\alpha) \cap \mathbf{x}_B(\alpha).$$

- Therefore:
  - to perform the simplest "and"-operation $f_\&(d_1, d_2) = \min(d_1, d_2)$,
  - we simply take the intersection of the corresponding α-cuts.

76. "And"- and "Or"-Operations in Terms of Alpha-Cuts (cont-d)

- This is a very natural operation.
- Indeed, for exactly defined properties, the set of all the elements which satisfy $A \,\&\, B$ is the intersection of:
  - the set of all elements which satisfy property A and
  - the set of all elements which satisfy property B.
- Similarly, for the simplest "or"-operation $f_\vee(d_1, d_2) = \max(d_1, d_2)$,
  - the composite property $A \vee B$ ("A or B")
  - is described by the membership function

$$\mu_{A \vee B}(x) = \max(\mu_A(x), \mu_B(x)).$$

- We want to find the α-cuts corresponding to this membership function

$$\mathbf{x}_{A \vee B}(\alpha) = \{x : \mu_{A \vee B}(x) \geq \alpha\}.$$

77. "And"- and "Or"-Operations in Terms of Alpha-Cuts (cont-d)

- The maximum of two real numbers:
  - is greater than or equal to α
  - if and only if one of these numbers is greater than or equal to α.
- Thus, $\mu_{A\vee B}(x) = \max(\mu_A(x), \mu_B(x)) \ge \alpha$ is equivalent to "$\mu_A(x) \ge \alpha$ or $\mu_B(x) \ge \alpha$".
- Hence:
  - the set $\mathbf{x}_{A\vee B}(\alpha)$ of x for which the condition $\mu_{A\vee B}(x) = \max(\mu_A(x), \mu_B(x)) \ge \alpha$ is satisfied
  - is the union of the set of all x for which $\mu_A(x) \ge \alpha$ and the set of all x for which $\mu_B(x) \ge \alpha$.
- In other words, for every α, we have

$$\mathbf{x}_{A\vee B}(\alpha) = \mathbf{x}_A(\alpha) \cup \mathbf{x}_B(\alpha).$$

78. "And"- and "Or"-Operations in Terms of Alpha-Cuts (cont-d)

- Therefore:
  - to perform the simplest "or"-operation $f_\vee(d_1, d_2) = \max(d_1, d_2)$,
  - we simply take the union of the corresponding α-cuts.
- This is also a very natural operation.
- Indeed, for exactly defined sets and properties: the set of all the elements which satisfy the property $A \vee B$:
  - is equal to the union of
  - the set of all elements which satisfy property A and
  - the set of all elements which satisfy property B.

79. Possible Exam Questions

- How to describe "and" in terms of α-cuts? Why is this description natural?
- How to describe "or" in terms of α-cuts? Why is this description natural?
- $\mu_1(x_1) = \max(0, 1 - |x_1|)$, $\mu_2(x_2) = \max(0, 1 - |x_2 - 1|)$, $y \stackrel{\text{def}}{=} x_1 \cup x_2$, $z \stackrel{\text{def}}{=} x_1 \cap x_2$, and $\alpha = 0.6$.
  - What is the α-cut for y?
  - What is the α-cut for or z?

80. Fuzzy Computations Can Be Reduced to Interval Computations: Derivation

- The main problem of fuzzy computation can be described as follows:
- We know an algorithm $y = f(x_1, \ldots, x_n)$ that relates:
  - the value of the desired difficult-to-estimate quantity y
  - with the values of easier-to-estimate auxiliary quantities $x_1, \ldots, x_n$.
- We also know, for every i from 1 to n, a membership function $\mu_i(x_i)$.
- This function describes the expert knowledge about the i-th input quantity x_i.
- Our objective is to compute the function

$$\mu(y) = \sup\{\min(\mu_1(x_1), \ldots, \mu_n(x_n)) : y = f(x_1, \ldots, x_n)\}.$$

81. Reduction to Intervals (cont-d)

- Let us now describe this relation in terms of α-cuts.
- This relation was first discovered and proved by Hung T. Nguyen in 1978.
- To describe this result in precise terms, let us first make some mathematics-related remarks.
- The function $y = f(x_1, \ldots, x_n)$ describes the relation between physical quantities.
- In physics, such a relation is usually continuous.
- Sometimes, we have a seemingly discontinuous transitions, e.g., in phase transitions when, e.g.:
  - the density of water changes into
  - a much smaller density of steam.

82. Reduction to Intervals (cont-d)

- However, it is not really a discontinuous transition.
- It is simply a very fast but still continuous one.
- In view of this observation, we will assume that the function $y = f(x_1, \ldots, x_n)$ is continuous.
- We will also assume the membership functions $\mu_i(x_i)$ are continuous.
- If we had exact knowledge, then continuity would make no sense.
- Indeed, then, the corresponding degree of confidence would abruptly go:
  - from 1 for possible values
  - to 0 for impossible ones,
  - without ever attaining any intermediate degrees.

83. Reduction to Intervals (cont-d)

- However, for fuzzy knowledge, continuity makes perfect sense:
  - if there is some degree of confidence that a value x_i is possible,
  - then it makes sense to assume that values close to x_i are possible too,
  - with a similar degree of belief.
- In practice, as we mentioned earlier, membership functions are indeed usually continuous.
- It is important to mention that:
  - for continuous membership functions $\mu_i(x_i)$,
  - the α-cuts $\{x_i : \mu_i(x_i) \geq \alpha\}$ are closed sets (i.e., sets which contain all their limit points).

84. Reduction to Intervals (cont-d)

- Finally, we require that for every i and for every $\alpha > 0$, the α-cut is a compact set.
- For real numbers:
  - since we have already assumed that the α-cuts $\{x_i : \mu_i(x_i) \geq \alpha\}$ are closed sets,
  - it is sufficient to require that these sets are bounded.
- This is true, e.g., if we assume that all the membership functions correspond to fuzzy numbers.
- In this case, all α-cuts are intervals.
- Suppose that we know the α-cuts $\mathbf{x}_i(\alpha)$ corresponding to the inputs.
- We want to find the α-cuts $\mathbf{y}(\alpha)$ corresponding to the output.

85. Reduction to Intervals (cont-d)

- By definition of an α-cut, $y \in \mathbf{y}(\alpha)$ means that $\mu(y) \ge \alpha$, i.e., that

$$\sup\{\min(\mu_1(x_1), \ldots, \mu_n(x_n)) : y = f(x_1, \ldots, x_n)\} \ge \alpha.$$

- By definition of the supremum, this means that:
  - for every integer $k > 2/\alpha$,
  - there exists a tuple $(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$ for which $y = f(x_1^{(k)}, \ldots, x_n^{(k)})$ and

$$\min(\mu_1(x_1^{(k)}), \mu_2(x_2^{(k)}), \ldots) \ge \alpha - 1/k.$$

- The minimum of several numbers is $\ge \alpha - 1/k$ if and only if all these numbers are $\ge \alpha - 1/k$, i.e.,

$$\mu_i(x_i^{(k)}) \ge \alpha - 1/k \text{ for all } i.$$

- Since $k > 2/\alpha$, we have $1/k < \alpha/2$ and $\alpha - 1/k > \alpha/2$.

86. Reduction to Intervals (cont-d)

- Thus, for each i and all k, the value $x_i^{(k)}$ belongs to the compact $(\alpha/2)$-cut $\mathbf{x}_i(\alpha/2)$.
- The tuples $(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})$ belong to the compact set
$$\mathbf{x}_1(\alpha/2) \times \mathbf{x}_2(\alpha/2) \times \ldots \times \mathbf{x}_n(\alpha/2).$$
- Thus, the sequence of these tuples has a convergent subsequence converging to some tuple $(x_1, x_2, \ldots, x_n)$.
- Since both f and μ_i are continuous, for this limit tuple, we get $y = f(x_1, \ldots, x_n)$ and $\mu_i(x_i) \geq \alpha$.
- In other words, every element $y \in \mathbf{y}(\alpha)$ can be represented as $y = f(x_1, \ldots, x_n)$ for some values $x_i \in \mathbf{x}_i(\alpha)$.

87. Reduction to Intervals (cont-d)

- Conversely, if $x_i \in \mathbf{x}_i(\alpha)$ and $y = f(x_1, \ldots, x_n)$, then $\mu_i(x_i) \geq \alpha$.
- So, $\min(\mu_1(x_1), \mu_2(x_2), \ldots, \mu_n(x_n)) \geq \alpha$, hence

$$\sup\{\min(\mu_1(x_1), \ldots, \mu_n(x_n)) : y = f(x_1, \ldots, x_n)\} \geq \alpha.$$

- Thus, $\mu(y) \geq \alpha$ and $y \in \mathbf{y}(\alpha)$.
- So, the desired α-cut $\mathbf{y}(\alpha)$ consists of exactly values $y = f(x_1, \ldots, x_n)$ for $x_i \in \mathbf{x}_i(\alpha)$:

$$\mathbf{y}(\alpha) = \{f(x_1, \ldots, x_n) : x_1 \in \mathbf{x}_1(\alpha), \ldots, x_n \in \mathbf{x}_n(\alpha)\}.$$

- For fuzzy numbers, all α-cuts are intervals, so this is exactly the interval computations range:

$$\mathbf{y}(\alpha) = f(\mathbf{x}_1(\alpha), \ldots, \mathbf{x}_n(\alpha)).$$

88. Fuzzy Computations Can Be Reduced to Interval Computations: Conclusion

- If the inputs $\mu_i(x_i)$ are fuzzy numbers and the function $y = f(x_1, \ldots, x_n)$ is continuous, then
  - for each α, the α-cut $\mathbf{y}(\alpha)$ of y is equal to
  - the range of possible values of $f(x_1, \ldots, x_n)$ as x_i ranges over $\mathbf{x}_i(\alpha)$ for all i:

$$\mathbf{y}(\alpha) = f(\mathbf{x}_1(\alpha), \ldots, \mathbf{x}_n(\alpha)).$$

- Thus, from the computational point of view,
  - the problem of processing data under fuzzy uncertainty
  - can be reduced to several problems of data processing under interval uncertainty,
  - as many problems as there are α-levels.

89. Fuzzy Computations Can Be Reduced to Interval Computations: Conclusion (cont-d)

- So, fuzzy computations can be reduced to interval computations.
- As we have mentioned, this is not just a theoretical observation.
- This is exactly how fuzzy data processing is usually performed.
- This is why interval computations techniques are explained in fuzzy textbooks.

90. Possible Exam Questions

- Assume that $u = x_1 + x_2$, $\mu_1(x_1) = \max(0, 1 - |x_1|)$, and $\mu_2(x_2) = \max(0, 1 - |x_2 - 1|)$. What is $\mathbf{y}(0.6)$?
- We have $\mathbf{y}(0.6) = \mathbf{x}_1(0.6) + \mathbf{x}_2(0.6)$.
- We already know that $\mathbf{x}_1(0.6) = [-0.4, 0.4]$ and $\mathbf{x}_2(0.6) = [0.6, 1.4]$.
- So, $\mathbf{y}(0.6) = [-0.4, 0.4] + [0.6, 1.4] = [0.2, 1.8]$.
- For students to do on their own: what is $y = x_1 - x_2$? what if $\alpha = 0.7$?

91. Intervals are Necessary to Describe Degrees of Belief

- Above, we described an idealized situation, in which we can describe degrees of belief by exact real numbers.
- In practice, the situation is more complicated.
- Reason: experts cannot describe their degrees of belief precisely.
- Indeed, let us start by reviewing the above-described methods of eliciting degrees of belief.
  - if an expert describes his or her degree of belief by selecting, e.g., 8 on a scale from 0 to 10,
  - this does not mean that his or her degree of belief is exactly 0.8
  - if instead, we ask him or her to select on a scale from 0 to 9,

- then whatever he or she chooses, after dividing it by 9, we will never get 0.8.

- If an expert chooses a value 8 on a 0 to 10 scale, then:
  - the only thing that we know about the expert's degree of belief
  - is that it is closer to 8 than to 7 or to 9,
  - i.e., that this degree of belief belongs to the *interval* $[0.75, 0.85]$.

- Another possible source of interval uncertainty is when we have *several* experts, and their estimates differ:
  - if, e.g., two equally good experts point to 7 and 8,
  - then, if we are cautious, we would rather describe the resulting degree of belief as the interval $[0.7, 0.8]$
  - or, in view of the above remark, as the interval $[0, 65, 0.85]$.

- If we determine the degree of belief by polling, then the same argument shows that the resulting numbers are not precise:
  - if 8 out of 10 experts voted for A,
  - then we cannot say that the actual degree of belief is exactly 0.8,
  - because, if we repeated this procedure with 9 experts,
  - we will never get exactly 0.8.
- In this case, there are two other sources of uncertainty:
  - First, picking experts is sort of a random procedure,
  - so, the result of voting is a statistical estimate that is not precise,
  - just like a statistical frequency estimate of probability;
  - better description will be to give an *interval* of possible values of $d(A)$.

- The polling method of estimating the degree of belief is based on the assumption that an expert can always tell whether he believes in a given statement S or not.
- Then, we take the ratio $d(S) = N(S)/N$ of the number $N(S)$ of experts who believe in S to the total number N of experts as the desired estimate.
- For $\neg S$, we thus have $N(\neg S) = N - N(S)$, so $d(\neg S) = N(\neg S)/N = 1 - d(S)$.
- In reality, an expert is often unsure about S; in this case:
  - instead of dividing the experts into two categories: those who believe in S and those who do not,
  - we must divide them into *three* categories:
    * those who believe that S is true (we will denote their number by $N(S)$),
    * those who believe that S is false (we will denote their number by $N(\neg S)$), and

    * those who do not have the definite opinion about S; there are $N - N(S) - N(\neg S)$ of them.

- In this situation, one number is not sufficient to describe the experts' degree of belief in S.
- We need at least two numbers.
- There are two ways to describe it.
- We can describe the degree of belief in S as $d(S) = N(S)/N$ and the degree of belief in $\neg S$ as $d(\neg S) = N(\neg S)/N$.
- These two numbers must satisfy the condition $d(S) + d(\neg S) \le 1$.
- This description is known as *intuitionistic fuzzy logic*.
- Alternatively, we can describe:
  - the degree of belief $d(S)$ in S and

- the degree of *plausibility* of S estimating as the fraction of experts who do not consider S impossible, i.e., as $pl(S) = 1 - d(\neg S)$.

- Thus, we get an interval $[d(S), pl(S)]$.

- This representation corresponds to the so-called *Dempster-Shafer formalism*.

- So, to describe degrees of belief adequately, we must use *intervals* instead of real numbers.

92. Possible Exam Questions

- Why do we need interval-valued degrees of belief?
- If an expert marked his/her degree of confidence as 4 on a 0 to 5 scale, what is the corresponding interval?
- What is intuitionistic fuzzy logic?
- What is Dempster-Shafer approach?
- If out of 10 experts, 3 say "yes", 4 say "no", and 3 are undecided, how will this be represented:
  - in intuionistic fuzzy logic?
  - in Dempster-Shafer approach?

93. Interval Computations for Processing Interval-Valued Degrees of Belief

- For an expert system with interval-valued degrees of belief, the following problem arises.
- Suppose that we have an expert system whose knowledge base consists of statements $S_1, \ldots, S_N$.
- We have an algorithm $f(Q, d_1, \ldots, d_N)$ (called *inference engine*) that, for any given query Q:
  - transforms the degrees of belief $d(S_1), \ldots, d(S_N)$ in the statements from the knowledge base
  - into a degree of belief in Q:

$$d(Q) = f(Q, d(S_1), \ldots, d(S_N)).$$

- For example, if $Q = S_1 \& S_2$, then $N = 2$ and $f(d_1, d_2) = f_\&(d_1, d_2)$.

Classification of . . .
Need to Process Fuzzy . . .
Degrees of Belief
Membership Functions
Need for "And"- and . . .
Fuzzy Data Processing
Zadeh's Extension . . .
Fuzzy Numbers
Reduction to Interval . . .
Home Page
Title Page
Page 99 of 103
Go Back
Full Screen
Close
Quit

94. Interval Computations for Processing Interval-Valued Degrees of Belief (cont-d)

- Suppose now that we know only the *intervals* $\mathbf{d}(S_1)$, $\ldots$, $\mathbf{d}(S_N)$ that contain the desired degree of belief.
- Then, the degree of belief in Q can take any value from:

$$f(Q, \mathbf{d}(S_1), \ldots, \mathbf{d}(S_N)) \stackrel{\text{def}}{=} \{f(Q, d_1, \ldots, d_N) \mid d_i \in \mathbf{d}(S_i)\}.$$

- Computing such an interval is a typical problem of *interval computations*.
- Since $f_\&$ and $f_\vee$ are increasing in both arguments:

$$f_\&([\underline{x}, \overline{x}], [\underline{y}, \overline{y}]) = [f_\&(\underline{x}, \underline{y}), f_\&(\overline{x}, \overline{y})],$$

$$f_\vee([\underline{x}, \overline{x}], [\underline{y}, \overline{y}]) = [f_\vee(\underline{x}, \underline{y}), f_\vee(\overline{x}, \overline{y})].$$

- For example,

$$\min([\underline{x}, \overline{x}], [\underline{y}, \overline{y}]) = [\min(\underline{x}, \underline{y}), \min(\overline{x}, \overline{y})],$$

$$\max([\underline{x}, \overline{x}], [\underline{y}, \overline{y}]) = [\max(\underline{x}, \underline{y}), \max(\overline{x}, \overline{y})].$$

95. Possible Exam Questions

- For two statements S_1 and S_2, we have $\mathbf{d}_1 = [0.5, 0.7]$ and $\mathbf{d}_2 = [0.3, 0.4]$.
  - What is the degree of confidence in $S_1 \,\&\, S_2$?
  - What is the degree of confidence in in $S_1 \vee S_2$?

96. Conclusion

- We have explained intrinsic and useful relations between interval computing and fuzzy data processing.
- The main relation is that a fuzzy set (membership function) can be viewed as a nested family of intervals.
- Its α-cuts corresponding to different levels of uncertainty α.
- From the computational viewpoint:
  - fuzzy data processing can be (and usually is)
  - reduced to level-by-level interval computations with the corresponding α-cuts.

97. Conclusion (cont-s)

- Another relation comes from the fact that
  - it is usually difficult to describe experts' degrees of certainty by exact real numbers;
  - a more adequate description of expert's uncertainty is by an interval.
- Processing interval-valued degrees of uncertainty also requires interval computations.