

## CS 4365/CS 5351 Interval Computations Fall 2017

**Class times:** Mondays and Wednesdays, 3-4:20 pm, in CCSB 1.0202.

**Instructor:** [Vladik Kreinovich](mailto:vladik@utep.edu), email [vladik@utep.edu](mailto:vladik@utep.edu), office CCSB 3.0404, office phone (915) 747-6951.

- The instructor's office hours are Mondays and Wednesdays 1:30-3 pm and 4:30-5 pm.
- If you want to come during office hours, there is no need to schedule an appointment.
- If you cannot come during the instructor's office hours, please schedule an appointment in the following way:
  - use the instructor's appointments page <http://www.cs.utep.edu/vladik/appointments.html> to find the time when the instructor is not busy (i.e., when he has no other appointments), and
  - send him an email, to [vladik@utep.edu](mailto:vladik@utep.edu), indicating the day and time that you would like to meet.

He will then send a reply email, usually confirming that he is available at this time, and he will place the meeting with you on his schedule.

### Prerequisites:

- for graduate students: no special pre-requisites; graduate level standing is sufficient
- for undergraduate students: ideally, Statistics and Numerical Methods; some of these pre-requisites can be waived on an individual basis.

**Main objective:** to learn theoretical and applied aspects of interval computations.

**Course description** An overview of interval computations that take into account how input uncertainties influence the computation result. A review of the main ideas behind interval computations, main interval techniques, and applications to practical problems such as robotics, computer graphics, control, and bioinformatics.

### What are interval computations?

**Need for indirect measurements.** In many practical situations, we are interested in the value of the quantity  $y$  that is difficult or even impossible to measure directly. To estimate the value of  $y$ , we:

- measure the values of all easier-to-measure quantities  $x_1, \dots, x_n$  which are related to  $y$  by a known dependence  $y = f(x_1, \dots, x_n)$ , and then
- use the measured values  $X_i$  of these quantities to compute the estimate  $Y = f(X_1, \dots, X_n)$  for the desired quantity  $y$ .

In engineering this procedure is called *indirect measurement*. In computer science, it is called *data processing*. The need for extensive data processing is the main reason why computers were invented in the first place.

**Problem: measurements are never 100% exact.** The problem is that measurements are never 100% exact. For example, a person is not exactly 5 feet 7 inches or 170 pounds, the actual height or weight

are slightly different. because of the difference  $\Delta x_i = X_i - x_i \neq 0$  between measured value  $X_i$  and the (unknown) actual value  $x_i$ , the result  $Y = f(X_1, \dots, X_n)$  of data processing is somewhat different from the actual value  $y = f(x_1, \dots, x_n)$  of the quantity of interest.

Traditionally, computers simply return a number  $Y$  without specifying how accurate is this number, i.e., how big the inaccuracy  $\Delta y = Y - y$  can be. From the practical viewpoint, however, it is important to know this accuracy. For example, in geophysical applications, if we estimated the amount of oil in a field as 100 million ton, then if it is  $100 \pm 10$ , this is good news, we should start exploiting this area; however, if it is  $100 \pm 200$ , then maybe there is no oil at all, so it would be better to perform additional measurements before investing into this area.

**Traditional approach assumes that we know probabilities.** Traditional engineering approach to this uncertainty (the one that students learn, e.g., in engineering labs) assumes that we know the probabilities of different values of measurement error  $d_i$ . Usually, it is assumed that there errors are independent and normally distributed.

**In practice, sometimes, we do not know these probabilities.** In many practical situations, we only know the upper bounds  $\Delta_i$  on the measurement errors:  $|\Delta x_i| \leq \Delta_i$ .

This means that after we measure the value as  $X_i$ , the actual value of the measured quantity can be anywhere within the interval  $[x_i] = [X_i - \Delta_i, X_i + \Delta_i]$ .

For example, if a thermometer shows 86 F, and its accuracy is half a degree, it does not mean that the temperature is exactly 86.000: it could be anywhere between 85.5 and 86.5.

For a single measurement result, this inaccuracy is easy to handle. However, when we process thousands of data points, these inaccuracies add up.

**Main problem of interval computations.** When the only thing we know about each input  $x_i$  is the interval  $[x_i]$  of possible values, then the natural question is: what is the range  $[y]$  of possible values of  $y = f(x_1, \dots, x_n)$ ? The problem of computing this range is called the problem of *interval computations*.

**Where are interval computations used?** They started with the Department of Defense trying to make sure that the intercontinental missiles fall within the desired range. They continued with NASA designing a trajectory which is guaranteed to hit the Moon.

Nowadays, they are used everywhere: in manufacturing, in robotics, in geoinformatics (processing geophysical data), in bioinformatics (processing bioinformatics data), in economics, in computer graphics, in computer security and privacy (where intervals are introduced on purpose, to avoid disclosing exact values).

**Non-CS students.** Due to the inter-disciplinary character of potential applications, we welcome not only CS students, but also students from geosciences, environmental sciences, computational sciences, and from other disciplines in which fuzzy techniques can be (and are) useful such as bioinformatics, engineering (especially control), etc. The presence of such students will help application problems.

We realize that some non-CS students may not be as good in programming different fuzzy-related algorithms as our own CS students. With this in mind, we plan to tailor homeworks and other

assignments to such students, so that these tailored assignments are more focused on applications and somewhat less on programming.

### What will be in the course

- Formulation of the main problem.
- General (software) methods of solving this problem. In general, the problem of finding the range  $[y]$  is NP-hard (intractable), and therefore, we can expect only estimates for  $Y$ . For each estimation method, we will describe the required computation time, the possibility of parallelization, etc.
- A brief survey of applications.

**Project.** After learning the basic interval techniques, students will start working on an (individual or group) project. Ideally, a project should be related to the main topic of student research (we can help with that) and help the student in working on his or her dissertation, thesis, or project.

For those who have not yet selected a topic, or who are interested in making interval computations their research topic, there are many interesting open research problems in which they can help.

### Learning Outcomes

#### 1. Knowledge and Comprehension

- Understand why data processing and imprecise measurements are needed in practice
- Understand the main practical problem solved by interval computations: estimating uncertainty of the results of data processing and imprecise measurements
- Understand the main ideas behind the statistical approach to the problem of estimating uncertainty of the results of data processing -- the approach which is mainly used in science and engineering
- Understand why in many practical situations, it is impossible to use the traditional statistical approach
- Explain why in many practical situations in which traditional statistical methods are not applicable, we need to estimate the range of a function on given intervals
- Know linearization techniques for range estimation; understand the limitations of such techniques; know practical problems where approximate methods such as linearization are not sufficient
- Know calculus-based techniques for solving the problem of range estimation; understand why these techniques are not always sufficient
- Know the formulas for interval arithmetic -- estimating ranges for arithmetic operations
- Understand the main idea behind straightforward interval computations: parsing the function and replacing each elementary operation with the corresponding operation of interval arithmetic
- Understand how to take round-off errors into account when performing interval computations
- Understand the drawbacks of straightforward interval computations

- l. Understand the main ideas behind improved interval computations: bisection, centered form, constraint propagation techniques, generalized (affine) interval computations; understand the advantages and disadvantages of these ideas
- m. Know other problems solved by interval computations techniques, such as solving systems of equations and optimization under uncertainty; understand the practical need for these problems, how they are solved now, and how interval techniques help
- n. Understand how interval techniques can be parallelized
- o. Understand the computational complexity of the main problem of interval computations (it is NP-hard)
- p. Know applications of interval computations to practical problems, e.g., to robotics, computer graphics, control, and bioinformatics
- q. Be familiar with the practical need to combine interval uncertainty with probabilistic and expert uncertainty, and have a general understanding of how this combination can be done

## 2. Application and Analysis

- a. Apply calculus-based techniques to estimate ranges of simple functions
- b. Parse a given function
- c. Apply straightforward interval computations to estimate the range of a given function
- d. Apply advanced state-of-the-art interval computation techniques to produce a better estimate for the range of a given function
- e. Apply traditional techniques to other practically useful problems such as solving systems of equations and optimization
- f. Apply interval techniques to other practically useful problems such as solving systems of equations and optimization

## 3. Synthesis and Evaluation

- a. Derive formulas for interval arithmetic
- b. Design programs which use traditional and interval techniques to solve reasonable size problems
- c. Apply interval techniques to solve a practice-oriented problem (ideally related to the student's area of research): from extracting an appropriate sub-problem from the general problem to selecting the most appropriate interval techniques to actually using these techniques

**Textbook:** R. E. Moore, R. B. Kearfott, and M. J. Cloud, Introduction to Interval Analysis, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 2009.

**Additional source of information:** Interval computations webpage <http://www.cs.utep.edu/interval-comp> (by the way, hosted by UTEP).

**Tests and grades:** There will be two tests and one final exam. Each topic means home assignments (mainly on the sheets of paper, but some on the real computer). Some of them may be graded.

Maximum number of points:

- first test: 10
- second test: 25
- home assignments: 10
- final exam: 35
- project: 20

(smart projects with ideas that can turn into a serious scientific publication get up to 40 points).

A good project can help but it cannot completely cover possible deficiencies of knowledge as shown on the test and on the homeworks. In general, up to 80 points come from tests and home assignments. So:

- to get an A, you must gain, on all the tests and home assignments, at least 90% of the possible amount of points (i.e., at least 72), and also at least 90 points overall;
- to get a B, you must gain, on all the tests and home assignments, at least 80% of the possible amount of points (i.e., at least 64), and also at least 80 points overall;
- to get a C, you must gain, on all the tests and home assignments, at least 70% of the possible amount of points (i.e., at least 56), and also at least 70 points overall.

**Standards of Conduct:** You are expected to conduct yourself in a professional and courteous manner, as prescribed by the [UTEP Standards of Conduct](#).

Graded work, e.g., homework and tests, is to be completed independently and should be unmistakably your own work (or, in the case of group work, your team's work), although you may discuss your project with other students in a general way. You may not represent as your own work material that is transcribed or copied from another person, book, or any other source, e.g., a web page.

Academic dishonesty includes but is not limited to cheating, plagiarism and collusion.

- Cheating may involve copying from or providing information to another student, possessing unauthorized materials during a test, or falsifying data (for example program outputs) in laboratory reports.
- Plagiarism occurs when someone represents the work or ideas of another person as his/her own.
- Collusion involves collaborating with another person to commit an academically dishonest act.

Professors are required to -- and will -- report academic dishonesty and any other violation of the Standards of Conduct to the Dean of Students.

**Disabilities:** If you feel you may have a disability that requires accommodation, contact the The Center for Accommodations and Support Services (CASS) at 747-5148, go to Room 106 E. Union, or [e-mail to cass@utep.edu](mailto:cass@utep.edu).

*See you in the class!*