RESEARCH DESCRIPTION

Quantities are ubiquitous and an important part of our understanding about the world – we talk of engine horsepower, size, mileage, price of cars; GDP, population, area of countries; wingspan, weight, surface area of birds, and so on. We live in a world of quantitative dimensions, and reasonably accurate estimation of quantitative values is necessary for understanding and interacting with the world. Our life is full of evaluations and rough estimates of all sorts. How long will it take to get there? Do I have enough money with me? How much of the load can I carry at once? These everyday, common sense estimates utilize our ability to draw a quantitative sense of the world from our experiences. This type of reasoning is particularly common in engineering practice and experimental sciences, in activities like evaluating the feasibility of a design, planning experiments, and setting up and double-checking detailed analyses. In these domains, one often comes across situations where a rough answer generated quickly is more valuable than waiting for more information or resources. Some domains like environmental science and biophysics are so complex that a rough estimate is all one can manage with the available knowledge and data. I believe that the same processes underlie both these common sense estimates and expert's reasoning to generate ballpark estimates. Specifically, the drawing upon experience to make such estimates, and the achievement of expertise in part by accumulating, organizing, and abstracting from experience to provide the background for such estimates, are the same fundamental processes. Such processes are at the heart of common sense reasoning.

Making such estimates, or, "back of the envelope" reasoning involves generating quantitative answers in situations where exact data and models are unavailable and where available data is often incomplete and/or inconsistent. A rough estimate generated quickly is more valuable and useful than a detailed analysis, which might be unnecessary, impractical, or impossible because the situation does not provide enough time, information, or other resources to perform one. As we try to increase the specificity in the answer, the analysis requires exponentially more resources in the form of time, information, formalization, and computation; and back of the envelope reasoning strikes an optimal tradeoff between complexity and specificity.

The goal of my research is to build computer programs that can do such reasoning, and in doing so, learn something about human cognition. Today's software, while good at calculation, has no physical intuition. For instance, in an evaluation of question-answering programs that mine text for answers, one program came up with 360 tons as the amount of folic acid that an expectant mother should have per day, and 14 feet as the diameter of the earth. (Alas, students sometimes come up with nonsensical answers as well, such as 3000 kilograms for the weight of the Earth.) My research is about building programs that have such a reality-check, a feel for what might be plausible, to reduce the brittleness and the garbage-in-garbage-out nature of today's computer systems. The problem-solving techniques that I am building are heavily based on similarity and experiential knowledge, and are equipped with a "feel for numbers." Research done over the last twenty years in the fields of qualitative reasoning and analogical reasoning provides the solid ground for my work.

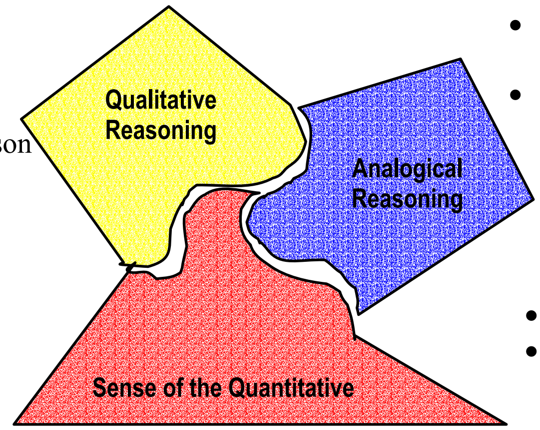
Research in analogical reasoning has led to computational models of retrieval from memory, matching and making inferences sanctioned by similarity, and generalizing from examples. These models are a success story of cognitive science, being great examples of computational models that have strong converging psychological evidence and are also employed in performance systems. However, they completely ignore the role of quantitative dimensions making them incomplete as cognitive models. My research highlights this inadequacy of existing models, as the goal is to make quantitative estimates. To this end, the questions that my research addresses are – 1) What role do quantitative dimensions play in computing similarity? 2) What can we infer about a missing quantitative value based on a similar example? 3) How do we develop a "feel for numbers" by exposure to a domain? These are important questions in cognitive science, and of practical importance in building systems to do back of the envelope reasoning.

One of the goals of qualitative reasoning research has been to understand human-like commonsense reasoning without resorting to the preciseness of models that consist of differential algebraic equations and parameters that are real-valued numbers. Qualitative reasoning has explored many different representations for expressing less precise and partial information about quantities[1]. The representations differ in the *kind* of distinctions that they allow us to make. The question of how to automatically build the necessary and relevant distinctions and tie them to numeric values has been largely ignored, as these representations were handcrafted by domain modelers. I have developed a set of systematic principles to automatically construct cognitively plausible representations of quantities, based on evidence and ideas from qualitative reasoning, cognitive psychology and linguistics.

Implementing back of the envelope reasoning provides a fertile ground to explore reasoning that involved a tight interweaving of qualitative and analogical reasoning, and cognitively plausible representations of quantity. Knowing a large number of examples of various problems and scenarios helps in building the estimation model. Given a new problem, we can solve it by retrieving a similar example from which we can borrow relevant modeling

[1] For example, status algebras (normal/abnormal); sign algebra (– , 0, +), which is the weakest representation that supports reasoning about continuity; quantity spaces, where we represent a quantity value by ordinal relationships with specially chosen points in the space; intervals and their fuzzy versions, among others.

assumptions, default values, etc. Exposure to a large number of examples involving various quantities in a domain gives rise to sense of the quantitative. The following picture shows the relation between the different fields and what they bring to back of the envelope reasoning.



- Provides analytic framework
- Facilitates comparison

- Borrow modeling assumptions
- Supply default, pre-computed information, parameter values
- Reality check

- What's reasonable, high, or low?
- Where do the qualitative abstractions of the quantity space come from?
- Grounds symbols to real world

Figure 1. Implementing back of the envelope reasoning: the pieces of the puzzle.

The sense of the quantitative, or, the "feel for numbers" addresses questions that have long been in the scope of qualitative reasoning, psychology, and linguistics – but have not been answered. One way to pose the question is – by exposure to examples in a domain (for example cars), what do we learn about quantitative attributes (like power, mileage, etc.). My claim is that we develop two kinds of abstract symbolic representations that are weaker, but useful abstractions of the space of values. The first kind is what I call distributional partitions, which amounts to dividing the space of values into low, medium and high based on the distribution of the quantitative values. This might explain how dimensional adjectives like "large", "tall", "expensive", acquire their meaning. The second kind of symbolic abstractions are what I call structural limit points. In contrast to the distributional partitions, these are not based on the quantitative values along the dimension, but on other structural aspects of the examples being considered. Consider "Freezing Point" – as we move across this point in temperature, there is a change in the structure of relationships between parameters, the underlying causal story. Another way to think about this is that things in world come in *structural bundles,* more than just bundles of correlated attributes. And so a change in underlying structure induces an interesting distinction on the space of quantity values. Structural limit points generalize the idea of phase transitions to everyday domains. Not always as crisply and rigorously defined as in physics, yet there exists such points on the space of values moving across which we have deep and structural implications. An example of such will be "poverty line." As we move across it, many other aspects of people – their lifestyle, the amount of time/money they spend on entertainment, education, the family and social climates in which they live, their expectations and relationships to the rest of the social structure, etc. change. I am working on programs that will generate such useful symbolic abstractions given a set of examples in a domain.

Back of the envelope reasoning is a powerful, flexible and useful problem-solving paradigm. Building programs that reason like that will lead to useful software that collaborates with human partners to accomplish tasks like damage control assessment, operations planning, sifting through on-line information for relevant data, teaching and tutoring, and developing complex scientific and engineering models. It has potential applications in education, especially engineering education, where estimation skills are crucial but rarely taught explicitly. There are studies that show that engineering undergraduates are surprisingly bad at this[2]. A better understanding of back of the envelope reasoning, and software that does it in human-like manner, will help in changing it.

[2] For example, more than 90% of mechanical engineering seniors (100 at MIT, and 250 from five other universities) came up with wrong order of magnitude estimates of value of energy stored in a 9-volt "transistor" battery. The responses varied by nine orders of magnitude excluding outliers [Linder, B. 1999. Understanding Estimation and its relation to engineering education, Ph.D. dissertation, Department of Mechanical Engineering, MIT].