Half Orders of Magnitude

Jerry R. Hobbs
USC Information Sciences Institute
Marina del Rey, California
hobbs@isi.edu

1 The Intuition

Consider the following multiple-choice questions:

> About how many children are there in the average family?
> a) 1 b) 10 c) 100
>
> About how many children are there in the average elementary school classroom?
> a) 1 b) 10 c) 100

For both of these questions we want to say, "None of the above."

Now consider the following revisions of those questions:

> About how many children are there in the average family?
> a) 1 b) 3 c) 10 d) 30 e) 100
>
> About how many children are there in the average elementary school classroom?
> a) 1 b) 3 c) 10 d) 30 e) 100

We have no trouble answering the first with b and the second with d.

Now consider the question

> About how many oranges are there in a basket full of oranges?

Our first reaction to this question is that there's no way of telling. We don't know how big the basket is, and for that matter we don't know how big the oranges are. The packing of oranges in a basket will also depend on the shape of the basket. But suppose we are given the choices

> a) 1 b) 3 c) 10 d) 30 e) 100

It turns out that when forced to pick one of these answers, well over half the people asked pick 10 and most of the rest pick 30. No one picks 1 or 100.

We are generally not able to come up with precise values for quantities we encounter in everyday life. We can easily come up with order-of-magnitude estimates, but these estimates are so imprecise as to be uselessly uninformative in most instances. However, people find it nearly as easy to come up with half-order-of-magnitude estimates and these are very often just as informative as they need to be.

This suggests that there is some cognitive basis for thinking in terms of half orders of magnitude (HOMs). For scales that are isomorphic to the integers or to the reals, precise values are often not available. We need coarser-grained structures on scales. Qualitative physics (deKleer, 1985; Forbus, 1988) has used the division of scales into negative, zero, and positive regions, but this is often too little structure. There has been work on order-of-magnitude reasoning (Raiman, 1986), where scales are partitioned in a way that changes in one region have no effect on quantities in higher regions; for example, adding a stamp to a letter does not change its weight enough for more postage to be required. This is an improvement over the tripartite division, but is still too little structure for many contexts. In this paper, I propose half orders of magnitude as a more refined intermediate structure on scales, one that often is just the right sort of structure one needs. We want a rough logarithmic categorization scheme for quantities, in which the categories are large enough that aggregation operations have reasonably predictive results and normal variation does not cross category boundaries, but are small enough that our interactions with objects is predictable from their category. Often the most appropriate estimate of a quantity is to a half order of magnitude.

Let us consider one more example to prime our intuitions before proceeding. The coins and bills in many currencies are available in half order of magnitude denominations. In the United States, for example, it would not be enough to have only coins for $.01 and $.10, and bills for $1.00, $10.00, and $100.00. It takes too many of each to make the next larger. In addition, American currency provides the nickel ($.05), the quarter ($.25), and the $5 and $20 bills as intermediate denominations. These are roughly at half-order-of-magnitude levels. There have been other intermediate denominations—coins for $.02 and $.50 and $2 bills—but these fell out of use. $50 bills are in circulation, but not in numbers anything like the numbers of $20 bills.[1]

In Section 2, I examine what an arithmetic of HOMs would be like. In Section 3, I argue that there are certain natural HOMs, anchored on persons, that play an

[1]The new euro currency has been constructed essentially as a binary system, with denominations at the 1, 2, and 5 levels. It will be interesting to see if this lasts, or if one of the intermediate values begins to drop out.

important role in our lives; it is there we will see how to calculate the number of oranges in a basket. In Sections 4, 5 and 6, I examine three very different ways in which HOMs impact on lexical semantics—in the uses of "several", "where", and the "approximately" sense of "about".

2 Shallow Defeasible HOM Arithmetic

One of the reasons for making estimates is that it allows us to do rough calculations. These are necessarily shallow, because each arithmetic operation increases the likely error, and they are defeasible in that we readily accept corrections when more precise measurements are made available. Nevertheless, they are often adequate for the purposes at hand. In this section, we will look at the operations of addition and multiplication and ask what the most reasonable rough calculations would be. That is, given two quantities where we know only their HOMs, what would be the best estimate for the HOMs of their sum and their product.

The square root of 10 is about 3.16. The geometric mean between that and 1 is about 1.8, and between it and 10 is about 5.5. Suppose we divide the positive reals up into HOM intervals as follows:

$$\ldots, [.55,1.8], [1.8,5.5], [5.5,18], [18,55], [55,180], \ldots$$

The first of these intervals is the numbers close to 1, the second those close to $\sqrt{10}$ (think of this interval as representing "several"), the third those close to 10, and so on. More generally, we divide the positive reals into the intervals $[10^{h-\frac{1}{4}}, 10^{h+\frac{1}{4}}]$, h a positive multiple of .5. Each of these intervals may be called an HOM, and referred to as the HOM around 10^h.

Now suppose all we know about quantities is their half order of magnitude. What sort of operations may we perform on them and how certain are our conclusions?

It is reasonable to assume that quantities are uniformly distributed throughout an HOM, since we have no knowledge to the contrary, and that is what we will assume for the remainder of this section.

We first consider the case of adding two numbers from the same HOM S, the HOM around 10^h. Suppose we add two arbitrary numbers from S. What HOM is the sum most likely to belong to, and what is the probability of this.

Consider Figure 1. If x and y correspond to a point in the shaded triangle, their sum will lie in S. If they correspond to a point in the other part of the square, their sum will be one half order of magnitude higher than S. Thus, the probability of x and y adding up to a number in S is the ratio of the area of the triangle to the area of the square. Each side of the triangle is $10^{h+\frac{1}{4}} - 2 \cdot 10^{h-\frac{1}{4}}$. The area of the triangle is thus

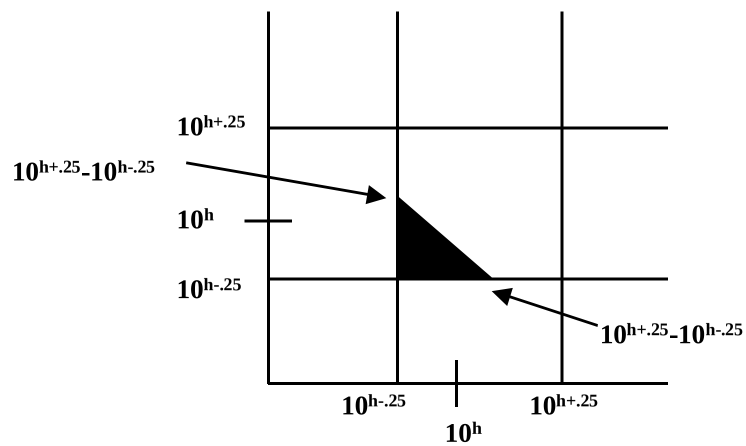

Figure 1: Sum of two identical half orders of magnitude.

$$\frac{1}{2}(10^{h+\frac{1}{4}} - 2 \cdot 10^{h-\frac{1}{4}})^2$$

The area of the square is

$$(10^{h+\frac{1}{4}} - 10^{h-\frac{1}{4}})^2$$

Thus, the ratio of the area of the triangle to the area of the square is

$$\frac{14-4\sqrt{10}}{22-4\sqrt{10}}$$

or about .1445.

Thus, if we add two numbers from the same HOM there will be a probability of about .8555 that the result will be one HOM larger. That is, defeasibly, several plus several equals about ten.

Now let S_1 and S_2 be two distinct HOMs, around 10^{h_1} and 10^{h_2}, respectively, where $h_1 < h_2$. Suppose we add a number from S_1 to a number from S_2. What HOM is the sum most likely to belong to, and what is the probability of this?

Consider Figure 2. If x and y correspond to a point in the shaded trapezoid, their sum will lie in S_2. If they correspond to a point in the other part of the rectangle, their sum will be one half order of magnitude higher than S_2. Thus, the probability of x and y adding up to a number in S_2 is the ratio of the area of the trapezoid to the area of the rectangle. The long vertical side of the trapezoid is $10^{h_2+\frac{1}{4}} - 10^{h_1-\frac{1}{4}} - 10^{h_2-\frac{1}{4}}$. The short vertical side of the trapezoid is $10^{h_2+\frac{1}{4}} - 10^{h_1+\frac{1}{4}} - 10^{h_2-\frac{1}{4}}$. The horizontal side is $10^{h_1+\frac{1}{4}} - 10^{h_1-\frac{1}{4}}$. The area of the trapezoid is thus

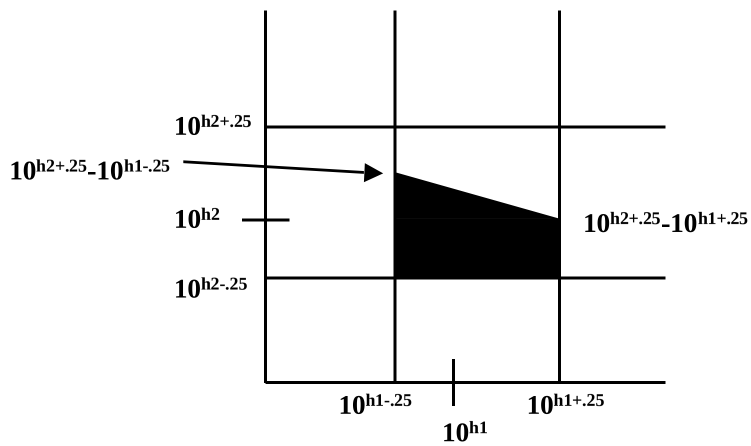

Figure 2: Sum of two different half orders of magnitude.

$$\frac{1}{2}(2 \cdot 10^{h_2+\frac{1}{4}} - 10^{h_1+\frac{1}{4}} - 10^{h_1-\frac{1}{4}} - 2 \cdot 10^{h_2-\frac{1}{4}})(10^{h_1+\frac{1}{4}} - 10^{h_1-\frac{1}{4}})$$

The area of the rectangle is

$$(10^{h_2+\frac{1}{4}} - 10^{h_2-\frac{1}{4}})(10^{h_1+\frac{1}{4}} - 10^{h_1-\frac{1}{4}})$$

The ratio of the area of the trapezoid to the area of the rectangle works out to

$$1 - \frac{1}{2} \cdot 10^{h_1-h_2} \frac{\sqrt{10}+1}{\sqrt{10}-1}$$

This is approximately

$$1 - .9625 \cdot 10^{h_1-h_2}$$

This means that there is about a 70% chance of the sum being in the larger HOM if the larger is one HOM larger than the smaller. So it is a reasonable estimate if there is corroborating evidence, and the best estimate even if there isn't. If the larger HOM is two HOMs larger than the smaller, there is about a 90% chance of the sum being in the larger HOM. If it is three HOMs larger, the probability is about 98%, and beyond that it is over 99%.

In brief, it is a good heuristic to take the sum of two identical HOMs to be the next larger HOM, and to take the sum of two distinct HOMs to be the larger of the two.

When we discard all information about a quantity other than its HOM, we lose information, and when we perform addition on two such quantities, we lose more information. We could have defined a logarithmic scale with some base other than $\sqrt{10}$, which, after all, was chosen because of the accidental fact that we use a decimal number system. It is reasonable to ask what base would be optimal in the sense that the least information is lost in doing addition. Hobbs and Kreinovich (2001) show that the optimal base with respect to addition is about 3.9. This is reasonably close to $\sqrt{10}$ in view of the fact that we approximate an HOM with values from 2 to 5 in everyday problems.

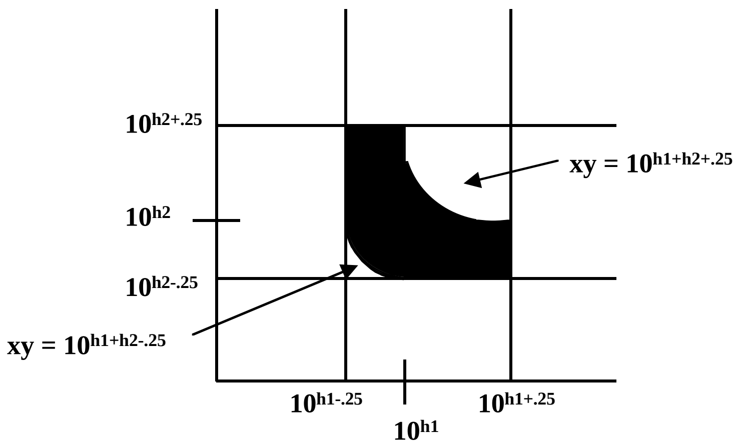

Figure 3: Product of two half orders of magnitude.

Now let us consider multiplication. (See Figure 3.) Of all the pairs of numbers in the region defined by

$$10^{h_1-\frac{1}{4}} \leq x \leq 10^{h_1+\frac{1}{4}}, 10^{h_2-\frac{1}{4}} \leq y \leq 10^{h_2+\frac{1}{4}}$$

those whose products do not fall into the HOM around $10^{h_1+h_2}$ are those under the curve defined by

$$xy = 10^{h_1+h_2-\frac{1}{4}}$$

and those over the curve defined by

$$xy = 10^{h_1+h_2+\frac{1}{4}}$$

That is, the probability P that the product of a number from the HOM around 10^{h_1} and a number from the HOM around 10^{h_2} is a number in the HOM around $10^{h_1+h_2}$ is

$$P = \frac{A+B}{C}$$

where

$$A = 10^{h_2+\frac{1}{4}}(10^{h_1} - 10^{h_1-\frac{1}{4}}) - 10^{h_1+h_2-\frac{1}{4}} \int_{10^{h_1-\frac{1}{4}}}^{10^{h_1}} \frac{dx}{x}$$

$$B = 10^{h_1+h_2+\frac{1}{4}} \int_{10^{h_1}}^{10^{h_1+\frac{1}{4}}} \frac{dx}{x} - 10^{h_2-\frac{1}{4}}(10^{h_1+\frac{1}{4}} - 10^{h_1})$$

and

$$C = (10^{h_2+\frac{1}{4}} - 10^{h_2-\frac{1}{4}})(10^{h_1+\frac{1}{4}} - 10^{h_1-\frac{1}{4}})$$

A is the area of the shaded region up to 10^{h_1}, B is the area of the shaded region after 10^{h_1}, and C is the area of the whole rectangle. The probability P then works out to

$$P = \frac{10^{\frac{1}{4}} - 2 + 10^{-\frac{1}{4}} + \frac{10^{\frac{1}{4}} - 10^{-\frac{1}{4}}}{4} ln 10}{10^{\frac{1}{2}} - 2 + 10^{-\frac{1}{2}}} \approx .7037$$

Thus, there is about a 70% chance that the product of a number in the HOM around 10^{h_1} and a number in the HOM around 10^{h_2} will be in the HOM around $10^{h_1+h_2}$. So it is a reasonable estimate if there is corroborating evidence, and the best estimate even if there isn't.

These results allow us to draw arithmetic inferences when we know quantities only approximately. Of course, the arithmetic we do must necessarily be shallow. After one or two operations, the probabilities become too low and the estimates are unreliable. However, this is sufficient for many practical applications, such as estimating the number of oranges in a basket.

3 Natural HOMs

The average adult human is about six feet or 180 cm tall. Certainly the vast majority of people are within one foot or 30 cm of that. Let us take this "Person Size" as the basis for a system of half orders of magnitude in everyday life. The half orders of magnitude above and below this correspond to functionally characterizable categories of objects and spaces, and part of what we know about the types of objects in our world is the natural HOM category they typically fall into.

Among the entities of Person Size are the major items of furniture, such as chairs, sofas, benches, tables, counters, and single beds. They can be moved by one person, but with difficulty. A space of this size can accommodate only one person comfortably.

One HOM below this measures about two feet or 60 cm in linear dimension. Entities of about this size can be held in two arms. They include baskets, cardboard boxes, sacks, TV sets, microwave ovens, laptops, bookshelves, dogs, and watermelons. A person's horizontal dimensions are of this HOM, which is why hugs are possible.

One HOM below this measures about eight inches or 20 cm. Entities of about this size can be held in one hand. Books, footballs, cantelopes, and desk telephones fall into this category.

One HOM below this measures about three inches or 8 cm. Entities in this category can be manipulated with the fingers. Examples are pens, mice (of both kinds), oranges, hamburgers, cell phones, and cups.

One HOM below this measures about one inch or 2.5 cm. It includes things that can be bitten, such as french fries, the chunks we cut our meat into before eating it, peppermint candies, erasers, and AA batteries. They can be manipulated easily with two fingers and a thumb.

One HOM below this measures one quarter inch or 1 cm, and includes things that must be handled with care between two fingers, such as diamonds, M&Ms, and thumb tacks.

One HOM below this is the size of a grain of rice, and we have little everyday experience with individual entities whose size is below this. (A speck of dust in your eye is an exception.)

One HOM above Person Size measures about twenty feet or six meters in linear dimension. Individual persons can move around in spaces of this size, and several people can fit into this space and engage in a joint activity that does not require much movement. Individual offices and typical rooms in houses are in this natural HOM category. Cars are at the lower end of the category.

The measures and typical examples of HOMs above this are as follows:

20 yards/meters	house, restaurant, small yard, 10-100 people
60 yards/meters	commercial building, bank, post office, large yard
200 yards/meters	small factory, small bridge, field
600 yards/meters	large factory, large bridge, dam
1 mile/1.5 km	town, airport
3 miles/5 km	small city
10 miles/15 km	large city, small county
30 miles/50 km	large county
100 miles/150 km	small state
300 miles/500 km	large state, small nation
1000 miles/1500 km	typical large European nation
3000 miles/5000km	the United States, China

Part of what we know about physical objects is the natural HOM category it typically belongs to. This, together with defeasible HOM arithmetic, is why we know that a basket full of oranges has about ten oranges in it, a living room full of people has about 10 people in it, and a hamburger is not eight feet in diameter. A basket has a natural HOM of two feet, and an orange a natural HOM of 3 inches. This is two HOMs difference, so there are about 10 oranges in a basket.

We can also define basic HOM categories for time, anchored on the units of time: 1 second, 5 seconds, 15 seconds, 1 minute, 5 minutes, 15 minutes, 1 hour, 3 hours, 12 hours, 1 day, 3 days, 1 week, 1 month, 3 months, 1 year, and so on. We know for various types of events which HOM categories they fall into. Thus, a cough lasts one second, a lecture lasts one hour, and a course lasts three months. We know that if we are told that John missed a course because he coughed, this requires elaborate explanation.

Rieger (1974) proposed encoding this kind of knowledge about the typical durations of events. Dahlgren (1988) proposed encoding this kind of knowledge about the size of everyday objects and using it to disambiguate prepositional phrase attachment ambiguities, as in

John drove down a street in a car.

4 "Several"

The most obvious place to look for a notion of half orders of magnitude in the English lexicon is the word "several".

I examined 25 occurrences of the word "several" in a wide variety of texts, including news articles, scientific articles, a novel, poetry, song lyrics, and transcripts of a meeting. For each instance, using my knowledge of the world and the context, I made an estimate of the range of numbers that might be counted as "several" in that context. For example, in

> Several women walked into the cafe.

I picture between three and five women. The statement would be misleading if the real number were two or if it were more than six.

When we use the word, there is generally an implicit comparison set from which the referenced entities are drawn. For example, the women who walk into the cafe must be drawn from the set of women who were in the neighborhood.

It was found that if the comparison set had an HOM of ten or fewer members, the word "several" referred to three to five. The above sentence is an example. If the comparison set had an HOM of thirty or more, "several" could refer to three to eight. For example, in the sentence

> About 80,000 people lost their long-distance service and *several* communities lost their 911 emergency phone.

my intuition is that as many as eight or so communities could have lost their 911 emergency phone.

This characterization covered 24 of the 25 examples. In 13 of 25 cases the comparison set was small and "several" meant three to five. In 11 of 25 examples, the comparison set was large and "several" meant three to eight. The one exception was a reference to

> ... criminal investigation of GE and several of its employees.

My feeling was that "several" could cover a range from three to about twelve.

Of course in a more rigorous study we would want to test a number of subjects, who were not theory-laden, and see what agreement there was among them.

5 "Where"

I examined 74 occurrences of the word "where" in the same corpus, seeking constraints on what can be where. The natural HOMs turned out to be a convenient way of characterizing that data.

Syntactically, the word "where" occurs adverbially, nominally, as a relative clause, and as a question, but in all of these, it places a figure X at or in a ground Y. When it occurs as a relative clause, as in

> farms where corn is grown

the head noun Y is the ground, and the relative clause is the figure X—the corn is grown at the farm. When it occurs adverbially, as in

> Where corn is grown, farmers have prospered.

the "where" phrase names a ground Y such that both the complement of "where" and the modified clause are figures X located at Y—there is a Y such that corn is grown there and farmers have prospered there. When it occurs nominally, as in

> The Midwest is where corn is grown.

the "where" phrase again names a ground Y such that the complement of "where" is a figure X located at Y. In "where" questions, as in

> Where is corn grown?

the answer is a ground Y such that the event described by the rest of the sentence is a figure X located at Y.

The figure and the ground are very often not physical objects but properties of physical objects, or events, conditions, activities, or situations involving physical objects. Thus, only 7 of the 74 examples concerned the physical location of a physical object. But 61 more of the 74 examples involved properties, events, conditions, activities, or situations of physical objects. Only 6 of the 74 involved abstractions at abstractions. When the figure and ground were properties, events, activities, and so forth of physical objects, I assumed their spatial extent was the same as the spatial extent of the physical objects.

I then asked for each of these 68 examples what the sizes of the figure and the ground were.

In 36 of the 68 cases, the figure and the ground were of the same HOM, as in

> Right here beside me is where you belong.

In 13 cases, the ground was one HOM larger than the figure, as in

> the counter where slabs of meat were kept
> The front room was where Marvin stayed.

In 5 cases, the ground was two HOMs larger than the figure, as in

> the houses where the workers live

Thus, in 54 of 68 cases, the ground was at least as big as the figure and no more than two HOMs larger.

There were 11 cases where the ground was more than two half orders of magnitude larger than the figure. One was a reference to a jewel (one-quarter inch) hidden in a chest (two feet). In this case, the disparity in size was one factor in the figure's being hidden.

The other 10 cases involved the activities of persons, groups of persons, or organizations, where these activities may take the persons to a wide set of locations within the ground, or where some property specifically of the ground is relevant to the activity.

> the town where he lives
> the laws of New York, where the business is based.

In these cases we can say that the spatial extent of the activity is greater than the spatial extent of the participants, and the former lies within two HOMs of the size of the ground.

Finally, there were 3 cases out of 68 where the HOM size of the ground was smaller than the HOM size of the figure. All three were in poems. One was metaphorical and concerned having a person in one's heart. One involved a global property of the person, beauty, being located in a body part, the eyes. The other involved a person's image being located in a painting.

What this exercise has shown is that overwhelmingly the figure and the ground are of comparable approximate sizes, where comparable means that the ground is from zero to two half orders of magnitude larger than the figure.

6 "About", "Approximately", and "Nearly"

I examined 86 examples of the word "about" in the same corpus. 52 of these were the "topic" sense of "about", as in

> He didn't know what they were talking about.

These uses will concern us no further.

Six of the examples involved a physical neighborhood, perimeter, or circuit around a physical entity, as in

> the town and all the country about

Eight more involved a quality, activity, or event of a physical entity, whose spatial extent could be estimated. In all 14 of these physical cases of "Y about X", the relative sizes of X and Y were similar to the cases of "where", described above. Y was either the same HOM as or one or two HOMs larger than X. In the above example, if the town is one mile in diameter, "the country about" probably refers to a region of about ten miles in diameter. In

Her face had a tense quality about it.

the spatial extent of the tense quality is either coincident with the face or extends not much beyond it.

The remaining twenty examples involved the "approximately" sense of "about". This sense can also be best analyzed in terms of half orders of magnitude.

When we describe a quantity as "about N", there is first of all an implicit precision g. Suppose the number of attendees at a meeting was 920. Then the first two of the following sentences is true, the third one false.

There were about 1000 people at the meeting.
There were about 900 people at the meeting.
There were about 980 people at the meeting.

In the first, we are using a precision of, say, 200, 250 or 500. In the second, we are using a precision of 100. In the third, we are using a precision of 10.

We have strong coarse-grained intuitions about what range of numbers an "about" statement is and is not true for. For example, in

About 80,000 people lost their long-distance service.

my intuition is that the real number lies between 77,000 and 84,000. It is certainly not 87,000, and probably not 75,000.

Each of the twenty examples I labelled with an interval in which the real number most likely lay. (In a proper study, I would have asked a number of subjects to make these judgments.) I took these judgments to be the data to be explained.

The general characterization of these uses of "about" is as follows: If X is about N, then $N = n \times g$, for some integer n and some HOM g, its precision, and $N - \frac{1}{2}g < X < N + \frac{1}{2}g$. In "about 900", $n = 9$ and $g = 100$. A reasonable guess about X is that is between 850 and 950. In "about 980", $n = 98$ and $g = 10$. A reasonable guess about X is that it is between 975 and 985.

There are a number of complications. The HOM between 1 and 10 will be something between 2 and 5. Generally, 5 is chosen because of its divisibility properties. Sometimes 2 is chosen. 3 would perhaps be better, being closer to $\sqrt{10}$, but it does not have good divisibility properties. Between 10 and 100, sometimes 25 is chosen, having good divisibility properties and being close to $10^{\frac{3}{2}}$.

It is not always obvious what g should be. In "about 1000", if g is 500, X could be between 750 and 1250. If g is 250, X could be between 875 and 1125. If g is 200, X could be between 900 and 1100. If g is 100, X could be between 950 and 1050. For each of these, it is easy to imagine contexts in which it is the appropriate choice. The example "There were about 1000 people at the meeting" would be true for 920 people if g is 500, 250, or 200, but not if it is 100.

Some complications arise. When g is 5, even multiples of g are likely to grab larger regions. Thus, "about 35" is probably from 33 to 37, while "about 40" is from 37 to 43.

When N is a multiple of 10, it is often difficult to know if g is 5 or 10. In general, if n is less than 5, g is more likely to be 5. If n is greater than 5, g is more likely to be 10. For example, "about 30" is probably between 27 and 33, while "about 60" is between 55 and 65.

Another complication is that X often gets rounded down, simply because of the way numbers are represented. 86,000 simply looks more like 80,000 than 90,000, so it is conceivable that 86,000 would be considered "about 80,000", whereas 74,000 would not.

This fact can be manipulated. My intuition about the sentence

> An industry spokesman said about 19,000 animals were killed in testing in the past decade, mostly mice and rats.

is that the actual number of animals could be anywhere from 18,600 to 19,800. The spokesman is trying to minimize the stated number, which is why he picked a g of 1000 rather than 10,000, and rounding down would help as well.

Finally, it should be said that the variation of X from N must be causally irrelevant. Someone might be "about 21" and still not be able to buy a drink.

All of these complications, however, are minor.

With these caveats, everyone of the twenty examples in the corpus fell roughly within the limits given by the formula.

I also looked at 10 examples of "approximately" and 13 examples of "nearly". Two of the "approximately" examples involved algebraic formulas.

> For $R \gg 0$, we have approximately $S = e^{-R_0}$.

The other 8 all fit the pattern for "almost".

All 13 of the uses of "nearly" fit the pattern for "almost", except that "nearly" also conveys the information that the real number is less than the estimate.

Two examples of "nearly" provided data about the real value. The first was

> Advancing issues outnumbered declining ones by nearly 2 to 1 on the NYSE, with 1,097 up, 581 down and 484 unchanged.

It is reasonable to assume g is .5. If the ratio had been near 1.5, the author would probably have said "3 to 2"; it is unlikely that a larger number would have been used for the denominator. The real number is 1.89, well within the window of width g. The other sentence was

> The number of filings soared by 29 percent in January, to 71,970, and experts predict that nearly 900,000 bankruptcies will be filed nationwide in 1991.

Assuming the rate of bankruptcies per month would be constant over the year, then in twelve months there would be 863,640 of them. If the precision g in the experts'[2] prediction is 100,000, then this value is well within the window of width g.

It would be interesting to try to derive this formula of usage from the characterization of the spatial uses of "about". Under what circumstances would we be willing to say X *is* N as opposed to X is *about* N? That would in a sense tell us the "spatial extent" of N. For example, if 901 people attended the meeting, we might be perfectly happy saying that the attendance was 900 rather than about 900. If the "spatial extent" of N is conceived of as one half order of magnitude less than g, then an "about" neighborhood around N of $\frac{1}{2}g$ each direction is exactly what one would expect.

7 Future Directions

One of the things we know about common physical objects is their approximate sizes. In this paper I have proposed a system for characterizing the approximate sizes of objects—half orders of magnitude, anchored on Person Size. HOMs provide a logarithmic scale which is coarse enough that aggregation operations have reasonably predictable results and yet which is fine enough that our interactions with objects can be predicted from knowledge of their HOM category. By examining the uses of the words "several", "where", "about", "approximately", and "nearly", I have shown that this system has utility in several very different linguistic contexts. Moreover, I sketched a kind of defeasible arithmetic that allows us to draw reasonable shallow conclusions from our approximate knowledge of the size and number of things.

There are a number of other words whose uses may be illuminated by an examination in this framework. There is the spatial sense of the word "near". What are the common relations between figure and ground when something is described as *near* something else? Are there HOM differences in the uses of "at" and "in"? How effective would HOM congruence be in disambiguating prepositional phrase attachment ambiguities? Does this framework illuminate the uses of quantitative adjectives? What about characterizations of shape for objects without rough radial symmetry, such as "tall"? It may be that for an object to be called tall, its height must be a half order of magnitude greater than its width. These questions await further work.

---

[2] Presumably experts in multiplication.

Acknowledgements

I have profited in this research from discussions with Vladik Kreinovich and Srini Narayanan. This research was supported in part by the National Science Foundation under Grant Number IRI-9619126 (Multimodal Access to Spatial Data) and in part by the Defense Advanced Research Projects Agency under Air Force Research Laboratory contract F30602-00-C-0168. Any opinions, findings, and conclusions or recommendations expressed in this article are those of the author and do not necessarily reflect the views of DARPA or the National Science Foundation.

References

[1] Dahlgren, Kathleen, 1988. *Naive Semantics for Natural Language Understanding*, Kluwer Academic Press, Boston, Massachusetts.

[2] De Kleer, Johann, and John Seely Brown. 1985. "A Qualitative Physics Based on Confluences." *Formal Theories of the Commonsense World*, ed. by Jerry R. Hobbs and Robert C. Moore, Ablex Publishing Corp., pp. 109-184.

[3] Forbus, Kenneth D., 1988. "Qualitative Physics: Past, Present, and Future", in *Exploring Artificial Intelligence*, Howard Shrobe, ed., pp. 239-296. Republished in *Readings in Qualitative Reasoning about Physical Systems*, Daniel S. Weld and Johan de Kleer (Eds.), Morgan Kaufmann Publishers, Inc., San Mateo, California, pp. 11-39, 1989.

[4] Hobbs, Jerry R. and Vladik Kreinovich, 2001. "Optimal Choice of Granularity in Commonsense Estimation: Why Half Orders of Magnitude", *Proceedings*, Joint 9th IFSA World Congress and 20th NAFIPS International Conference, Vancouver, British Columbia, pp. 1343-1348, July 2001.

[5] Raiman, Olivier, 1986. "Order of Magnitude Reasoning", *Proceedings, AAAI-86, Fifth National Conference on Artificial Intelligence*, Philadelphia, Pennsylvania, pp. 100-104.

[6] Rieger, Charles J., III., 1974. "Conceptual Memory: A Theory and Computer Program for Processing the Meaning Content of Natural Language Utterances", Memo AIM-233, Stanford Artificial Intelligence Laboratory, Stanford University.