How To Compare the Incomparable

Dick Grune
dick@cs.vu.nl

Vrije Universiteit
de Boelelaan 1081
1081 HV Amsterdam

ABSTRACT

A technique for deciding between two processes, one with cost (x_A, y_A) and the second with cost (x_B, y_B) where $x_A > x_B$ and $y_A < y_B$ and no further information is available, is derived, using dimension- and gauge-theory like arguments.

The technique can be summarized as: If you can't add them, multiply them.

Keywords: compiler construction, models, decision-making

Introduction

Imagine you are a maid serving a family on one of the Lesser Kalakola Islands. The family is going to have guests from the mainland tonight and you are sent to the market with money to buy one basket of fruits. On the market, which is very small, you are offered the choice between two baskets, both at the same price. One contains 3 yapyappas and 3 sakaronjas and the other one 5 yapyappas and 1 sakaronja. You have no idea of the preferences of the family's strange guests or even how to explain the options to them. Which basket would you buy? (Think of an answer now and compare it to the theory later.)

The Problem

Compiler writers are often confronted with the above problem. A source language has a construct that can be translated into object code in two ways, one (named *A*) costing x_A memory words and y_A CPU-seconds and the other one (*B*) x_B memory words and y_B CPU-seconds. All is well if $x_A > x_B$ and $y_A > y_B$ (or vice versa) since then the choice is obvious, but if $x_A > x_B$ and $y_A < y_B$ we have a problem.

We realize that we cannot afford just not to choose; in our naivety we approach the user and ask him which one he prefers. The user, who does not understand the implications and does not want to be bothered in the first place, tells us: "You're the compiler writer, do it right". When we then press him to tell us at least whether he prefers speed to memory or vice versa, he says: "Well, I really don't know, they're both important. Why don't you treat them as equally important?"

Some Observations

Treating two incommensurable quantities as equally important is easier said than done. We can observe the following:

1. We want a function

   $$F(x_A, y_A, x_B, y_B)$$

   which gives us either *{A}* if *A* is preferable, *{B}* if *B* is preferable, or *{A,B}* (the latter, e.g., when $x_A = x_B$ and $y_A = y_B$).
2. We are denied any yardstick in any of the quantities, both absolute and relative. Memory usage could be measured in bits, bytes or megabytes and the time in nanoseconds or centuries, and the result of *F* should not change. This implies that we have no constant of the dimension "memory usage", nor one of dimension "time". This restriction, although it seems the height of ignorance, is actually a source of strength. For one thing, it forbids any formula for *F* in which x_A and y_A are added.
3. When we exchange *x* and *y* in $F(x_A, y_A, x_B, y_B)$ the answer does not change.
4. When we exchange *A* and *B* in $F(x_A, y_A, x_B, y_B)$ the answer changes accordingly.

These observations strongly restrict the functions possible for F and will be called Restriction 1 to 4 in the sequel. We shall show that, under some reasonable assumptions and disregarding trivial variants, there is only one function satisfying all restrictions: comparison of $x_A y_A$ and $x_B y_B$.

The Solution

The answer required has the form {smaller, equal, larger}. Such an answer is supplied by comparing a numeric value to zero (this is Assumption 1); we introduce a function $V(x_A,y_A,x_B,y_B)$ with the mapping

$$V(x_A,y_A,x_B,y_B)<0 \quad <-> \quad F(x_A,y_A,x_B,y_B)=\{A\}$$
$$V(x_A,y_A,x_B,y_B)>0 \quad <-> \quad F(x_A,y_A,x_B,y_B)=\{B\}$$
$$V(x_A,y_A,x_B,y_B)=0 \quad <-> \quad F(x_A,y_A,x_B,y_B)=\{A,B\}$$

V must be continuous; any discontinuity, say at (x_A,y_A,x_B,y_B), would indicate knowledge of absolute values of x and y (restriction 2). The same applies to all its derivatives, which must likewise all be continuous. Consequently, the function V can be written as a Taylor series:

$$V(x_A,y_A,x_B,y_B)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}c_{ijkl}{x_A}^i{y_A}^j{x_B}^k{y_B}^l$$

Restriction 2 also says that $V(\alpha x_A,y_A,\alpha x_B,y_B)>0$ shall give the same result as $V(x_A,y_A,x_B,y_B)>0$, for any value of $\alpha>0$. Now,

$$V(\alpha x_A,y_A,\alpha x_B,y_B)=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\alpha^{i+k}c_{ijkl}{x_A}^i{y_A}^j{x_B}^k{y_B}^l$$

By changing α we can emphasize certain terms and play down other terms, thus changing V's behaviour, which is unacceptable. The problem is solved by requiring that there be only one power of α, i.e., $i+k=N$ is constant. This is essentially the dimension argument from physics: all terms added must have the same power of x (i.e., dimension in x). Obviously the same argument holds for the y's, and, for symmetry reasons (restriction 3), with the same value of N. This reduces V to:

$$V(x_A,y_A,x_B,y_B)=\sum_{i=0}^{N}\sum_{j=0}^{N}c_{ij(N-i)(N-j)}{x_A}^i{y_A}^j{x_B}^{N-i}{y_B}^{N-j} \qquad (1)$$

This means that given the first two indices of c we can calculate the last two, which are therefore redundant: we can write c_{ij}.

Restriction 3 tells us that $V(x_A,y_A,x_B,y_B)=V(y_A,x_A,y_B,x_B)$ for @i[all] values of x_A,y_A,x_B,y_B, which can only be realized if

$$c_{ij}=c_{ji} \qquad (2)$$

whereas restriction 4 implies that $V(x_A,y_A,x_B,y_B)=-V(x_B,y_B,x_A,y_A)$, again for @i[all] values of x_A,y_A,x_B,y_B, which can only be realized if

$$c_{ij}=-c_{(N-i)(N-j)} \qquad (3)$$

This turns V into the difference of two (symmetrical) sets of terms that are identical except that A and B have been exchanged.

If we attribute the first set of terms entirely to A and the second set to B, they represent in some sense the "costs" of A and B; this can be realized by requiring V to be the difference of two functions, one in A and one in B (this is Assumption 2). For our function in equation 1 this means that the only non-zero coefficients are c_{00} and c_{NN} and that they are equal:

$$V(x_A,y_A,x_B,y_B)=c_{NN}{x_A}^N{y_A}^N-c_{NN}{x_B}^N{y_B}^N$$

from which the cost function takes the form

$$C(x,y)=c_{NN}(xy)^N$$

or, since the values of N and c_{NN} are immaterial, as long as they are positive non-zero:

$$C(x,y)=xy \qquad (4)$$

Note:
Even if we do not make the Assumption 2, we still get the same result if we admit linear terms only: if N=1, the

equations 2 and 3 above alone already make sure that there are no cross products (i.e., products that contain values from both *A* and *B*). The only possible form of equation 1 for $N=1$ is:

$$V(x_A,y_A,x_B,y_B)=\alpha x_A y_A-\alpha x_B y_B$$

Unfortunately, this does not extend to values of $N>1$: for $N=2$

$$V(x_A,y_A,x_B,y_B)=\alpha x_B^2 y_B^2+\beta y_A x_B^2 y_B+\beta x_A x_B y_B^2-\beta x_A y_A^2 x_B-\beta x_A^2 y_A y_B-\alpha x_A^2 y_{A_2}$$

satisfies all restrictions, except that it does not allow itself to be viewed as the difference of two functions, one in *A* and one in *B* (unless $\beta=0$).

A Geometrical Interpretation

The function $V(x_A,y_A,x_B,y_B)$ divides 4-D space (actually a half-space) into two parts, where the results are *{A}* and *{B}* resp., separated by a 3-D space where the result is *{A,B}*. Rather than trying to visualize the 4-D space, we can draw a plane through it in which x_B,y_B are constant:

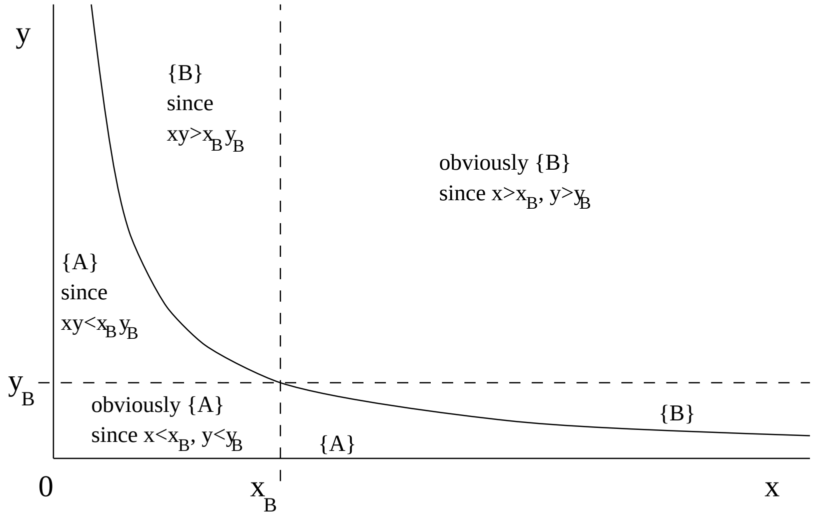

Figure 1.

In two of the four regions the answer is obvious. The curve

$$xy=C=x_B y_B$$

is the simplest one that divides the two undetermined regions without intersecting the *x*- or *y*-axes (which would betray knowledge of an absolute value of *x* or *y*). It should be noted that all more complicated boundaries have been ruled out by Assumption 2 above.

A Psychological Interpretation

The human mind still wants to be able to add and subtract happiness and misery linearly. This can be accomodated in the model by writing the expression

$$x_A y_A > x_B y_B$$

as

$$\log x_A + \log y_A > \log x_B + \log y_B$$

(a form which did not emerge in the above derivation since taking the logarithm of a quantity requires absolute knowledge of its scale). This suggests that the additive effect of a change from x_A to x_B is proportional to $\log x_A - \log x_B$.

Example: somebody having a salary of $10000 will experience a certain increase in happiness if his salary is raised to $20000. The model predicts that for a second, equally large increase in happiness a raise to $40000 will be required. (Note that this only holds in the absence of all other knowledge!)

Slightly more information

The model is easily adapted to some situations where slightly more information is available. Suppose property y is considered twice as important as property x; y is then considered as two properties, each with cost y and the cost function then becomes xy^2.

Quite often we have at least some vague idea about the scale of a property, particularly when a cost value is absurdly high. Although you may be able to sell tickets to travel a certain distance in one hour for $100, in 2.5 hours for $40 and in 5 hours for $20, nobody will pay $36 million for a one-second journey or accept a one-year journey for $0.01. We can incorporate this into the model by lowering the curve in fig. 1 over a small distance ε, so that it will intersect the x-axis in a far-away point X:

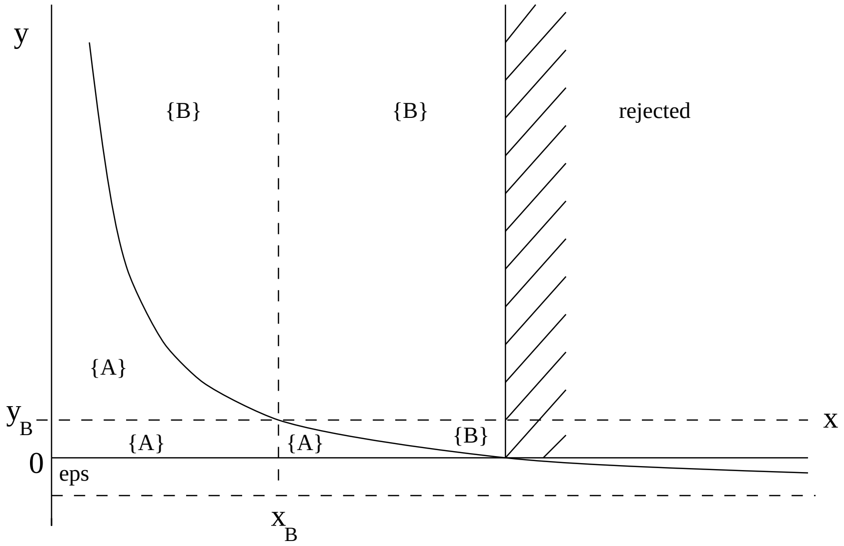

Figure 2.

The curve, which separates the A region from the B region, is described by:

$$U(x,y)=x(y+\varepsilon)-C=0$$

It passes through (x_B,y_B) so that $x_B(y_B+\varepsilon)=C$ and through $(X,0)$ so that $X\varepsilon=C$, where X is an absurdly high cost value of x. Eliminating C and ε yields:

$$U(x,y)=xy-x_By_B\left[\frac{X-x}{X-x_B}\right]$$

If $U(x_A,y_A)<0$ choose A, if $U(x_A,y_A)>0$ choose B. Again we can express this as the difference of two similar terms:

$$\frac{U(x_A,y_A)}{1-\frac{x_A}{X}}=\frac{x_Ay_A}{1-\frac{x_A}{X}}-\frac{x_By_B}{1-\frac{x_B}{X}}$$

provided all points with $x\geq X$ are rejected beforehand. The cost function is then:

$$C(x,y)=\frac{xy}{1-\frac{x}{X}} \tag{5}$$

For normal values of x ($\ll X$) this form is insensitive to the exact value of X. The formula is especially useful if a cost parameter can assume the value zero; it prevents the other parameter from running away.

Note that if both $x_A>X$ and $x_B>X$, both A and B are disqualified, and in addition to *{A}*, *{B}* and *{A,B}* the comparison may yield *{}*.

Summary

If

- we want to decide between two processes, one with cost (x_A, y_A) and the second with cost (x_B, y_B); and
- we have no further knowledge about *x* or *y*; and
- we want to base our decision on the comparison of two numerical values; and
- we want to be able to view these numerical values as the costs of *A* and *B*;

then

- the only cost function $C(x,y)$ to exhibit the necessary degree of symmetry is

$$C(x,y)=xy$$

Small increases in our knowledge about *x* and *y* can be accommodated as small changes in the form of $C(x,y)$.

The model leads to the conclusion that a basket with 3 yapyappas and 3 sakaronjas is preferable to one with 5 yapyappas and 1 sakaronja, as fits our intuition.