Title
An evaluative framework for research and development strategies

Authors
Rausser, Gordon C.
Yassour, Joseph

Publication Date
1980-08-01

AN EVALUATIVE FRAMEWORK FOR RESEARCH AND DEVELOPMENT STRATEGIES

Gordon C. Rausser and Joseph Yassour

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
August 1980

AN EVALUATIVE FRAMEWORK FOR RESEARCH AND DEVELOPMENT STRATEGIES

Gordon C. Rausser
Professor and Chairman, Department of Agricultural and Resource Economics
University of California, Berkeley.

Joseph Yassour
Associate Professor, Department of Economics and Management of the Kibbutz
Ruppin Institute, Israel.

ABSTRACT

This paper begins with a distinction between an "active-learning" framework and a "passive-learning" approach when employed as an evaluation of R D strategies. A simple decision-tree formulation is employed to gain insights on the essential differences of these two approaches. Numerical differences between the active- and passive-learning frameworks are derived. It is shown that in a sequence of decisions under uncertainty, where the probabilities of the occurrence of the unknown events are also unknown, much can be gained by taking into consideration the effect of learning along the horizon plan on decisions to be made in the future.

Subject areas: active learning, research and development, decision theory.

AN EVALUATIVE FRAMEWORK FOR RESEARCH AND DEVELOPMENT STRATEGIES

INTRODUCTION

In the research and development process, uncertainty is a key and pervasive characteristic. Unfortunately, in both the public and private sector, these processes are generally unstructured. Formal and operational methods for structuring these processes are not generally implemented. Cumbersome frameworks, whose associated cost of implementation appears to exceed by a wide margin in their associated benefits, occupy most of the available literature. In this paper an operational framework is advanced to capture inherent uncertainties and manage the risk of resource allocations to alternative R&D processes. This framework is based on the notion of adaptive or dual control [1] [4].

One of the key elements of the framework to be advanced relates to the determination of more accurate probabilities of "success." The manager of an R&D department is visualized as attempting to determine how many teams should be assigned to the development of a new technology. The individual teams are presumed to operate independently with a given complex of manpower and equipment. There is an underlying probability of success which is fixed but unknown. The more teams that are assigned to a particular project, the greater the chances that at least one team will succeed. Given a prior probability of success and a specified length of the planning horizon, along with a specified criterion function, an adaptive control formulation may be employed to determine the optimum number of teams to support.

Our analysis begins with a distinction between an "active-learning" framework and a "passive-learning" approach. The latter is typically employed as an evaluation of R&D strategies. A simple decision-tree formulation is employed to gain some insight on the essential differences of these two approaches. Once this distinction is drawn, we proceed to demonstrate how both methods can be used to evaluate the development of a new product to which multiple-research teams in a dynamic setting can be assigned. Numerical differences between the active- and passive-learning frameworks are derived and general conclusions are drawn. This is followed by a problem involving development of complementary products or a single-product with more than one attribute. Here, again, numerical examples are offered and results for the active-learning versus the passive-learning frameworks are distinguished. Finally, a general "active" learning framework is structured for the "multiple teams," "multiple product" attribute case.

ACTIVE VERSUS PASSIVE LEARNING

The R&D process is fraught with formidable uncertainties. It is indeed difficult to capture the relevant probabilities that should be assigned to associated unknown events. Often research administrators assign judgmental probabilities to the unknown events as a first approximation. In many situations, the research administrator may be prepared to incur some costs to capture more accurate measurements of the relevant probabilities. It is well known that this may be accomplished by formal sequential designs of experiments [2]. Such experimentations should be undertaken if the expected value of information generated, however determined, is higher than the cost of experiment.

In many real-life situations, the cost of R&D experiments is formidable. To mitigate such costs, a history of past actions can be viewed as "passive experiments"; results generated can be employed as a basis for estimating the relevant probability distributions. Under these circumstances, "passive learning" takes place with the research administrator utilizing only historical information. No future measurements conditioned on current decisions are allowed. In the case of active learning, the research administrator updates probability measurements according to past history but, in addition, makes future measurements for "ex ante history" conditioned on current decision. In essence, actions are selected which maximize the sum of both current period gains and the present value of experimentally expected future gains. A trade-off between current and experimentally expected future gains is explicitly recognized. Specifically, some current period gain is forsaken in order to obtain improved probability measurements to allow more nearly optimum actions to be selected during future periods of the planning horizon. In this setting, current gains--as well as the rate of learning during early periods of planning horizon--are jointly optimized.[1]

The distinction between the passive- and active-learning frameworks can be drawn from a simplistic example. Suppose a person is faced with a game in which he can invest $1.00 in a slot machine which may be either of type A or type B, each with a probability of .5. If the machine is of type A, the return is $1.50; and if the machine is of type B, the return is 0. He is allowed to play the game 10 times. Assuming the person is risk-neutral, should he play the game?[2] In the passive-learning framework, the expected value of the game is .5 (.5) + .5 (-1) = -.25. Hence, the decision for this framework is to not play the game. If, however, an active-learning approach is employed, we find that the expected value of the information obtained from the first game is greater than the cost of the experiment (25 cents). Obviously, there is a probability of .5 that a total of $5.00 will be earned. Since there is also a .5 probability that the machine will be of type B, the expected value of the first trial is .5 (5) + .5 (-1) = $2.00.

The expected value of the information obtained from this trial is .5 (4.5) + .5 (0) = \$2.25; and, thus, the total value of the first game is the expected immediate gain plus the expected value of the information obtained or \$2.00.

In terms of decision analysis, the active-learning framework can be described as:

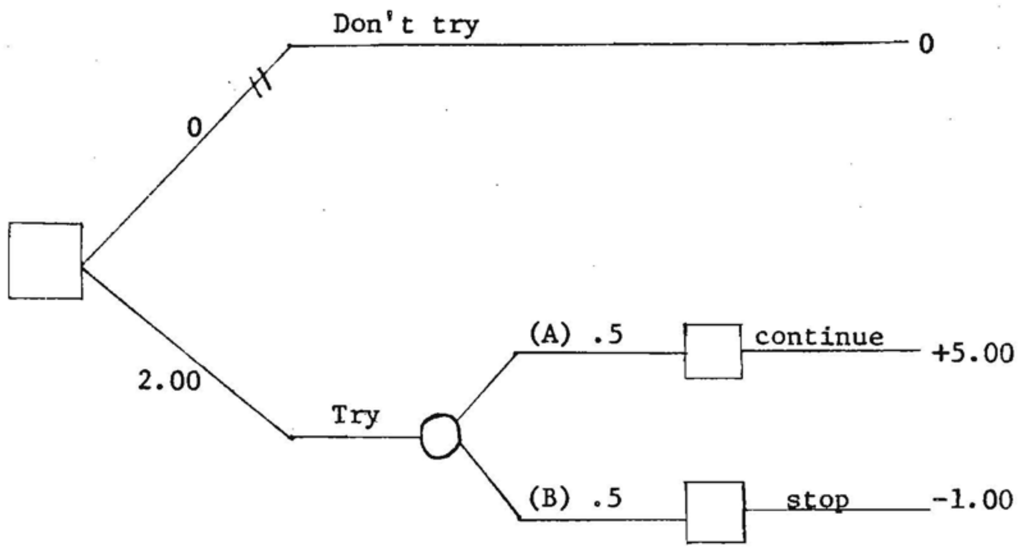

Figure 1

where the posterior probability of the machine--as of type A--is assigned 1 if the first trial is successful and 0 if it is a "failure." For the myopic approach of passive learning, the decision tree is instead

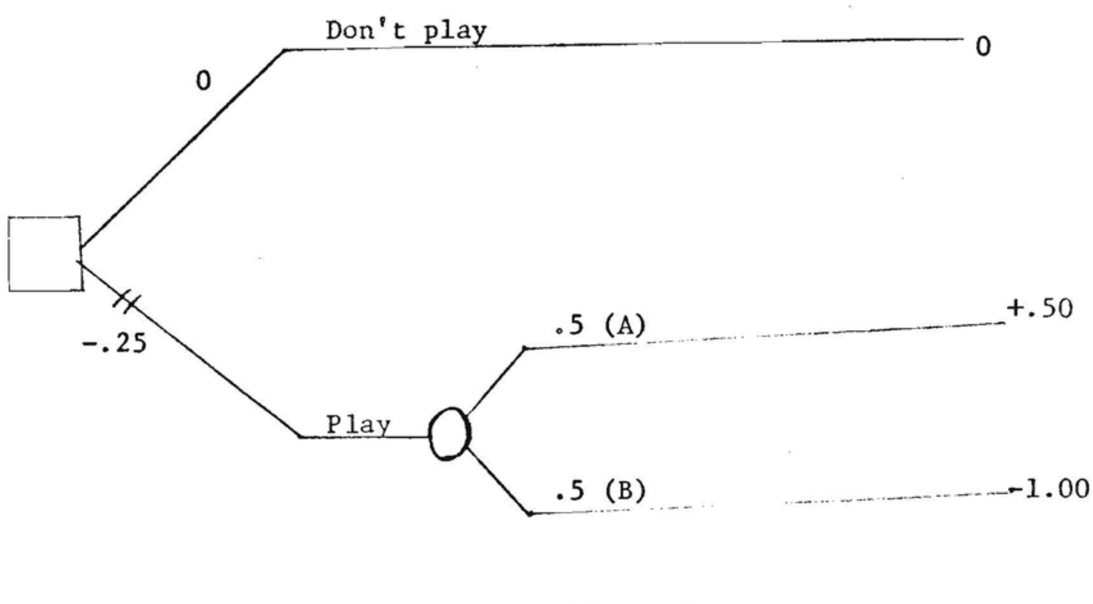

Figure 2

The above simple example can be generalized in the case of evaluating R&D strategies. Suppose that the head of an R&D department has been assigned the responsibility of attempting to develop a new product. The research head may want one or more research units allocated to this task. Suppose, in addition, that the probability that any one team would succeed to develop the product is not known with certainty.

To proceed with formal analysis, the following assumptions will be imposed:

1. The process which generates a success or failure for a particular team is specified as a Bernoulli with an unknown probability, $\tilde{P}$.
2. The unknown $\tilde{P}$ has a beta distribution with prior parameters n_0 and r_0.[3]
3. The total gains generated each year, $\tilde{y}_t$, consist of three components:
   a. a dynamic product demand effect, represented by last year's actual gains (y_{t-1});
   b. sale revenues generated from the introduction of the new product to the market; and
   c. the variable cost associated with the research teams being assigned to the product development in question where the cost is proportional to the number of such teams.

Given the above assumptions and defining n_t as the number of teams assigned to the project in year t ($n_t = 1, 2, . . .$); r_t denoting the number of successful teams in year t, $r_t = 0, 1, . . ., n_t$, c_t variable cost associated with one team; and $\tilde{z}_t$ revenue resulting from a success or failure in year t, i.e.,

$$z_t = \begin{cases} z_t \text{ if } r_t > 0 \\ 0 \text{ if } r_t = 0. \end{cases} \tag{1}$$

With these definitions, the total gain in year t may be represented by

$$\tilde{y}_t = a_t\, y_{t-1} + \tilde{z}_t - c_t\, n_t \tag{2}$$

where a_t is the rate of inertia in revenues from one year to another.

The formal decision problem is:

$$\max_{n_t} \sum_{t=1}^{T} \rho^{-t}\; E_{t-1}(\tilde{y}_t) \tag{3}$$

where T is the number of years in the planning horizon, ρ is the rate of discount, and $E_{t-1}(\tilde{y}_t)$ is expected value of total gains in year t which is given by

$$E_{t-1}(\tilde{y}_t) = a_t\, y_{t-1} + [1 - f_{hb}(0 \mid n_t, N_t', R_t')]\, z_t - c_t\, n_t \tag{4}$$

where $N_t' = \Sigma_{t=0}^{t-1} n_t$, $R_t' = \Sigma_{t=0}^{t-1} r_t$ and f_{hb} is the hyperbinomial distribution which follows from the assumed prior beta distribution and binomial sampling. The probability of r_t successful teams, given an allocation of n_t teams and the history on past team allocations and successes, is given by

$$P(\tilde{r}_t = r_t \mid n_t, N_t', R_t') = f_{hb}(r_t \mid n_t, N_t', R_t')$$

$$= \frac{(r_t + R_t' - 1)!\,(n_t + N_t' - r_t - R_t' - 1)!\,n_t!\,(N_t' - 1)!}{r_t!\,(R_t' - 1)!\,(n_t - r_t)!\,(N_t' - R_t' - 1)!\,(n_t + N_t' - 1)!} \tag{5}$$

The above formulation can be formally solved by the use of dynamic programming. In year $t = T$, i.e., the last year of the planning horizon, n_T must be selected such that the expected value of gains for the last period is maximized taking into account the state variables reflecting the history of past actions and $\Sigma_{t=0}^{T-1} n_t$ and successes $\Sigma_{t=0}^{T-1} r_t$. Proceeding sequentially in year $T - 1$, n_{T-1} must be selected such that the discounted expected revenue of year T, conditioned on n_{T-1} plus the expected gain of year $T - 1$, is maximized. This sequential analysis proceeds until we reach the first period of the planning horizon at which the optimal n_1 is determined. More formally, for the active-learning formulation, the maximization problem for a two-period horizon is:

$$\underset{n_1}{\text{Max}} \left\{ E_0 \, (a_1 \, y_0 + \tilde{z}_1 - c_1 \, n_1) + \underset{n_2}{\text{Max}} \; \rho \left(E_1 \left[a_2 \, (a_1 \, y_0 + \check{z}_1 - c_1 \, n_1) + \tilde{z}_2 - c_2 \, n_2 \right] \right) \right\} \tag{6}$$

where $\tilde{z}_1$ is the yet unknown z_1; $\check{z}_1$ is the unknown z_1, realized in the second period after r_1 is observed.

For the passive-learning formulation, the two-period maximization problem is:

$$\underset{n_1}{\text{Max}} \left\{ E_0 \, (a_1 \, y_0 + \tilde{z}_1 - c_1 \, n_1) + \underset{n_2}{\text{Max}} \left[E_0 \left(a_2 \, [a_1 \, y_0 + \check{z}_1 - c_1 \, n_1] + \tilde{z}_2 - c_2 \, n_2 \right) \right] \right\} \tag{7}$$

Notice that in (6) the second expectation operator is E_1, namely, the anticipated results of the first period are taken into account in the active learning, whereas in (7) the expectation operator for the second period is the same as that for the first period, E_0. This means that in passive learning the anticipated results of the first period are not taken into account and, thus, future probability measurements are not conditioned on current decisions.

For a stationary state, the probability of success, $\tilde{P}$, will stabilize at its mean, and the two formulations will converge. However, the larger the distance from the stationary state (the longer the planning horizon), the larger the marginal benefit of learning and thus, presumably, the greater the difference between the active- and passive-learning formulations. Moreover, for very short planning horizons, the difference between the two formulations will be small due to the inability to exploit any active learning that might occur during the first or second period of the planning horizon.

Product Development and the Allocation of Multiple Teams

The potential value of the active-learning approach can be demonstrated for a problem involving a four-period horizon and the possibility of allocating multiple teams to the development of a new product. For this illustrative numerical example, we shall presume that $n_0 = 3$, $r_0 = 1$, $z_t =$ \$1,000, $c_t = \$350$, $a_t = .9$, and $\rho = 1$. Given these numerical values, the passive-learner and active-learner strategies are determined for the first period along with contingency plans.[4] That is, after the optimal decision for the first period was found (n_1^*), a simulation model based on the "optimal" decision and on the prior information (n_0, r_0) determines the number of successful teams (r_1^1) out of the n_1^* teams. Based on this result, the "optimal" decision for the second period was determined by the same procedure as for the first period. This process was followed as well for all remaining periods.[5] It was found that the "optimal" decision for the first period was to assign two teams to the project, and the total expected profits for the four years was \$1.945 million.[6]

In the case of active learning, the probability of success differed from one period to another in accordance with the history (n_1, r_1, n_2, r_2, etc.) for each period.[7] The optimal decision for the first period in the active-learning case was to assign three teams to the project, and the total expected profits for the four years was $2.086 million. Hence, an increase of 7.3 percent in the expected profit is due to the additional learning which resulted from assigning one more team to the project in the first year. It should be noted that the zero percent rate of return inflated the difference between the two approaches but, nevertheless, the above example demonstrates the additional gains resulting from the active-learning approach.

Joint Product Attribute Development

The differences between the active- and passive-learning formulations can also be illustrated by another R&D problem which is frequently encountered. Suppose the research administrator desires to improve the attributes of an existing product, and alternative research projects can be undertaken in the hopes of achieving such improvements. In the numerical example to be presented below, we shall presume that two of the product attributes are in need of improvement, and there are two alternative projects--each associated with one attribute--that might be pursued. The first project, labeled A, has a cost of c_a; its estimated probability of success is P_a; and its contribution, if successful, is z_a. Similar notation is employed for the second project, labeled B. It is possible to undertake each project separately, sequentially, or simultaneously. If both projects are successful, the gain is z, not necessarily the sum of z_a and z_b. The probability of success in both projects, if pursued simultaneously, is $P_a \cdot P_b$. The success over both projects, however, is not statistically independent.

If project B is pursued after project A, its probability of success is not P_b but, rather, $P_{b|a+}$ or $P_{b|a-}$ or the conditional probability of success in B given A was a success (a+) or a failure (a-), respectively. Moreover, the probability of success for each of the projects in period t + 1 depends upon the outcome in period t.

The problem is to determine which projects should be undertaken and in what order. The prior probabilities of P_a and P_b will be defined by x/w and u/v, respectively. For this problem, Bayes' rule allows the relevant updated probabilities from the assumed beta prior distributions on P_a and P_b and binomial sampling. Specifically, the updated probabilities of success, given any outcome, would require that the denominators increase by 1, 2, or 3 if, during the prior period, project A was pursued, project B was pursued, or both projects A and B were pursued, respectively. The numerator is simply increased by 0, 1, 2, or 3 if both A and B were failures; if A was a success but B was not; if B was a success but A was not; and if both A and B were successful during the previous period, respectively. For example,

$$P_{a|a+} = \frac{x+1}{w+1}, \quad P_{a|a-} = \frac{x}{w+1}, \quad P_{a|a-,b+} = \frac{x+2}{w+3},$$

$$P_{b|a+,b-} = \frac{u+1}{v+3}, \text{ and } P_{ab|a-,b-} = \frac{x}{w+3} \cdot \frac{u}{v+3}. \tag{8}$$

For this numerical example, the parameters will be specified as: C_a = 100, C_b = 200, R_a = 1,000, R_b = 2,000, R = 2,500, x = 1, w = 2, u = 1, v = 5, and a = 0, i.e., no dynamic product demand effect exists.

The relevent decision tree for the active-learning formulation over a two-period planning horizon is represented in Figure 3. With no learning, the probabilities of success in any project do not change in the second period. The pseudo-optimal policy for this framework is to undertake projects A and B simultaneously, with expected returns of 695. If, however, learning takes place through future-updated-probability measurements, the optimal policy is to undertake project A only with expected returns of 719.

Concluding Remarks

The two examples can be combined such that there are M research teams to be assigned to project A or B. It is possible to assign all to one project or divide them between the two projects. Obviously, the greater the number of teams in one project, the higher the probability of success, the more information to be gathered on the probability of success in that project, but the lower the probability of success and learning from the other project.

The approach taken in this paper can be extended to the case where the projects are complementary, and success in each project is necessary for the entire program to be successful. In this case, the program probability of success depends on the inherent probability of success of the project as a whole, on the human capital of the teams assigned to each project and on the success in the past of other teams and other complementary projects. In this setting, equation (6) can, thus, be generalized to:

$$\max_{b_{ikt}} E_t\left\{\left(a_{t+1}\, y_t + \tilde{z}_t - \sum_{i,k} b_{ikt}\right) + \max_{b_{ikt+1}} \rho \left(E_{t+1}\left[a_{t+2}\left(a_{t+1}\, y_t + \check{z}_t - \sum_{i,k} b_{ikt}\right) + \tilde{z}_{t+1} - \sum_{i,k} b_{ikt+1}\right]\right)\right\} \quad \text{for all } t = 1, 2, \ldots, T \tag{9}$$

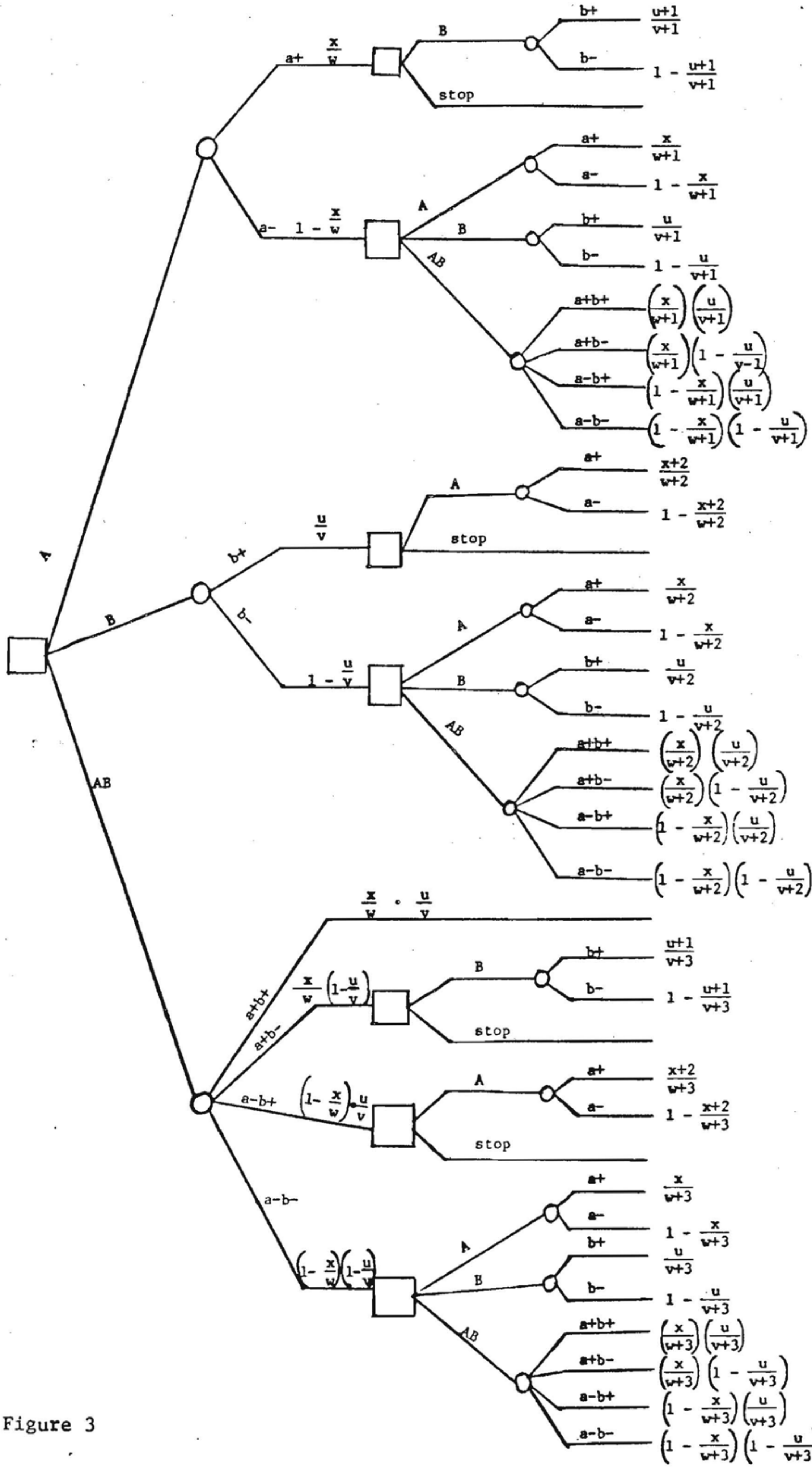
A
B
AB
a+
a-
b+
b-
stop
a+b+
a+b-
a-b+
a-b-
x/w
1 - x/w
u/v
1 - u/v
(u+1)/(v+1)
1 - (u+1)/(v+1)
x/(w+1)
1 - x/(w+1)
u/(v+1)
1 - u/(v+1)
(x+2)/(w+2)
1 - (x+2)/(w+2)
x/(w+2)
1 - x/(w+2)
u/(v+2)
1 - u/(v+2)
(u+1)/(v+3)
1 - (u+1)/(v+3)
(x+2)/(w+3)
1 - (x+2)/(w+3)
x/(w+3)
1 - x/(w+3)
u/(v+3)
1 - u/(v+3)

Figure 3

where b_{ikt} is the budget assigned to the team i on project k in year t, and z_t is the total gain as a result of successes in all complementary projects.

There are, of course, other possible ways to formulate the multiproject, multiteam case. For example, the project administrator might be interested in minimizing the time of the completion of the project as whole, or alternatively to maximize to probability of overall success.

The principal result of our analysis is simply that, when making decisions under uncertainty, it is indeed worthwhile to obtain information about the unknown event as long as the expected value of the information is higher than the cost of the information. In a sequence of decisions under uncertainty, where the probabilities of the occurrence of the unknown events are also unknown, much can be gained by taking into consideration the effect of learning along the horizon plan on decisions to be made in the future. When making a current decision, the decision-maker should find the alternative which maximizes not only the immediate outcomes of that decision but also the value by which improved measurements can lead to optimal future outcomes.

Footnotes

[1]This framework has been referred to in the mathematics and electrical engineering literature as dual control (Feldbaum, Tse). The special case of a two-period planning horizon has been referred to in the literature as preposterior analysis (Pratt, Raiffa, and Schlaifer).

[2]This type of problem has been referred to as the "one-armed bandit" problem. See e.g., DeGroot.

[3]It is well known that the beta distribution has two particularly advantageous features. First, almost any desirable shape of the probability distribution is admitted; and, second, the updating of the unknown parameter estimates is indeed simple. If the prior parameters are n_0 and r_0, then an experiment of sample size n_1 is taken. With r_1 successes observed; then the updated parameters are simply $n_0 + n_1$ and $r_0 + r_1$.

[4]Computer programs for both of these algorithms are available upon request from the authors.

[5]In the fourth period, the simultation model was not used; but, rather, the highest expected value determined the profit and the optimal decision.

[6]This result is based on 1,000 simulations.

[7]The hyperbinomial distribution was used to calculate the probabilities for both the passive- and active-learning algorithms.

References

[1] Bellman, R. E. *Adaptive Control Processes: A Guided Tour*. Princeton: Princeton University Press, 1961.

[2] Chernoff, H. *Sequential Analysis and Optimal Design*. Philadelphia: *SIAM*, 1972.

[3] DeGroot, M. H. *Optimal Statistical Decisions*. New York: McGraw-Hill, 1970.

[4] Feldbaum, A. A. *Optimal Control System*. New York: Academic Press, 1965.

[5] Pratt, J. W., H. Raiffa, and R. O. Schlaifer. *Introduction to Statistical Decision Theory*. New York: McGraw-Hill, 1965.

[6] Tse, E. "Adaptive Dual Control Methods," *Annals of Economic and Social Measurement*. Vol. 3 (1974), pp. 65-83.