

ECONOMIC MODELS OF BUSINESS DECISION

Constantin MILITARU¹, Marian-Sorin IONESCU²

Prezenta lucrare se concentrează pe probleme de cercetare, ce implica alocarea secvențială a deciziilor în cadrul stadiilor de dezvoltare a unei organizații economice noi create, tipul start up. Managerii iau decizii operaționale ce vor influența performanța economică ulterioară și strategiile de viitor ale organizației economice. În prezentul material se studiază astfel de secvențe, de decizie prin utilizarea conceptelor și metodelor de programare dinamică. Aceasta prezintă modul în care performanța în cadrul procesului de afaceri poate fi îmbunătățită. Ceea ce domină opțiunile posibile de a fi alese de către manager în cadrul procesului decizional inițial este problematica de alocarea efortului în timp, partajarea între un loc de muncă existent și angajamentul de a crea și dezvolta o nouă afacere. Se demonstrează că politica optimă de alocare a timpului este gestionată de manager cu toleranță pentru muncă dar și pentru modul în care se face repartizarea acestei resurse către organizația economică vizată. Este foarte important să înțelegem modul de alocare a resurselor pentru activități interne, cum ar fi dezvoltarea de produse și de recrutare a clienților. Lucrarea de față prezintă ceea ce un manager rațional va face atunci când se confruntă cu o alocare a efortului de la categorii diferite de clienți. Oferim, de asemenea, liniile directoare (tendințele) pentru îmbunătățirea performanțelor de către un manager care nu a atins încă decizia optimă. Se prezintă modul de calcul al timpului de optim pentru un produs nou și se descrie modul în care strategia aleasă este afectată de suma alocată pentru finanțare și de gardul incertitudinii sale.

This paper focuses on research issues involving sequential allocation decisions in the development stages of a newly created economic organization, such as start up, that invoke effort allocation in sequential decision-making at early development stages of a new venture creation. Managers make decisions that influence subsequent and future performance. We studied such sequences of decisions by using dynamic programming. It prescribes how business performance can be improved. What dominates the manager's decision process initially is the effort allocation problem in sharing time between an existing job and committing to the new venture. We show that the optimum time allocation policy is driven by manager's tolerance for work and by how returns behave with respect time allocation in the venture. It is very important to understand resources allocation to internal activities such as product development and customer recruitment. We present what a rational manager will do when faced with allocation to effort to different customer categories. We also provide guidelines for improving the performance of a manager who may not be acting optimally.

We prescribe the optimum release time for the new product and describe how this strategy is affected by the expected amount of funding and its uncertainty.

1. Introduction

We studied here the decisions that a manager must take in allocating time to exploiting a customer base in order to maximize profit. It is important to decide how to allocate time in a face of different costs or/and benefits accruable from

¹Professor at the University "POLITEHNICA" Bucharest

²Phd. Candidate, University POLITEHNICA of Bucharest, e mail:sorin_ionescu_07@yahoo.com

potential customer versus contacting loyal customers who have made purchases. Critical time points after which no more potential customers should be contacted are derived, also we can obtain critical time points after which loyal customers are more profitable to contact than potential customers. This is a risky investment, under two conditions all possible future outcomes of exploiting that market opportunity are known at the time decision is made, when the probability of each of these outcomes occurring is also known at the time a decision is made. Most economic models of business decision making are applicable to risky decisions.

Problem Description³

The total number of contacts will be limited to a positive integer, N , for each period of time. We denote by πL^c , the immediate expected profit from contacting a loyal customer, this being typically larger than the immediate profit from contacting a potential customer, denoted by πP^c . For not contacting a loyal customer, the immediate expected profit, denoted by πL , is positive since a proportion of loyal customers is expected to buy the same product again, without being contacted.

	Contacted	Non-Contacted
Loyal Customer	πL^c	πL
Potential Customer	πP^c	0

Assumption 1: Knowledge about managerial product is obtained only from contact with the manager. **Assumption 2:** The potential customers who do not buy in the period during they are contacted forget about the contact. **Assumption 3:** Loyal customers who have already bought the product are constantly informed about upgrading and do not need new information to repeat purchases. The probability of success for each contact with a potential customer is denoted by β , this is a conservative property, regardless of past contacts. We present a concrete illustration using a company providing **Enterprise Resource Planning (ERP)** that integrates internal and external management information across an entire organization, embracing finance-accounting, manufacturing, sales and service. Each quarter, the manager of ERP company contacts private company to offer services. Each quarter, the software company must decide who should be contacted in order to maximize its profits. Loyal customers are more profitable during the period of contact, the company has an incentive to contact potential customers because it is the only way of building a base of loyal customers.

2. Optimum Contact Policies, Basic Model, Markov Decision Model

The manager collects fixed wages each period allowing him or her to cover the cost of contacting N of the most costly potential buyers. The main target

³ M. Levesque, Model of Entrepreneurial Decision: A Dynamic Programming Approach, Ph.D. Thesis, The University of British Columbia, 1998

is to allocate time to the class of potential buyers that is most profitable in the long run.

2.1. Markov Decision Process

Markov processes have been of interest to marketers for a long time, in the context of frequently purchased goods. We consider computational aspects of decision-theoretic planning model by **Markov Decision Processes**. Our results show that their performance is not always superior and depends on the parameters of a problem and the parameters of the algorithms, such as number of iterations in the value determination procedure in MPI. The combined value policy iteration (CVPI) algorithm proposed in our work implements this criterion and generates an optimum policy faster than both policy and value iteration algorithms. A Markov Decision Processes is a stochastic automaton, decision making is based on **maximum expected utility** (MEU) approach. Given a policy, we compute an expected utility U of each state S_i , according to the formula:

$$U(S_i) = R(S_i) + \sum_j P_{ij}^{pi(s_i)} U(S_j) \quad (1)$$

The goal is to find an **optimum policy** π^* which maximizes the expected utility of each state. VI is an iterative procedure that calculates the expected utility of each state using the utilities of neighboring states until the utilities calculated on two successive steps are close enough: $\max_{S_i} |U(S_i) - U'(S_i)| < \varepsilon$, where ε is a predefined threshold value. The smaller the threshold, the higher is **the precision of the algorithm**. We can calculate a corresponding policy according to the **maximum expected utility** principle, choosing:

$$\pi(S_i)^* = \arg \max_a \sum_j P_{ij}^a U(S_j) \quad (2)$$

as an optimum policy. We propose a schematic description of the VI (VALUE ITERATION) algorithm :

```

function VALUE - ITERATION( $P, R$ ) returns a utility matrix
  inputs:  $P$ , a transition – probability matrix
            $R$ , a reward matrix
  local variables:  $U$ , utility matrix, initially identical to  $R$ 
                     $U'$ , utility matrix, initially identical to  $R$ 
  repeat
     $U \leftarrow U'$ 
    for each state  $i$  do
       $U(S_i) \leftarrow R(S_i) + \max_a \sum_j P_{ij}^a U(S_j)$ 
    end
  until  $\max_{S_i} |U(S_i) - U'(S_i)| < \varepsilon$ 
  return  $U$ 

```

2.2 Policy Iteration Alogritm

We called policy iteration (PI), another way of finding optimum policies. It iteratively performs two step: **value determination**, calculates the utility of each state given the current policy, **policy improvement**, updates the current policy if any improvement is possible:

```
function POLICY – Iteration ( $P, R$ ) returns a policy
```

```

inputs:  $P$ , a transition – probability matrix
           $R$ , a reward matrix
local variables:  $U$ , utility matrix, initially identical to  $R$ 
           $\pi$ , a policy, initially optimum with respect to  $U$ 
repeat
     $U \leftarrow VALUE - DETERMINATION (P, U, R, \pi)$ 
     $change \leftarrow false$ 
for each state  $i$  do
    if  $\max_a \sum_j P_{ij}^a U(Sj) > \sum_j P_{ij}^{\pi(sj)} U(Sj)$ 
    then
         $\pi_{(Si)} \leftarrow \arg \max_a \sum_j P_{ij}^a U(Sj)$ 
         $change \leftarrow true$ 
    end
    until  $changed = false$ 
return  $U$ 

```

Value determination of PI is done by solving a system of linear equations :

$$U(Si) = R(Si) + \sum_j P_{ij}^{\pi(Si)}, \quad (3)$$

where the reward R and the transitional probabilities matrix $P^{\pi(Si)}$ are given, and $U_{(Si)}$ are unknown. This is often the most efficient approach for small state spaces. Another way of implementing Value Determination is by iteratively calculating the utilities for the given policy as:

$$U(Si) \leftarrow R_{[i]} + \sum_j P_{ij}^{\pi(Si)} U(Sj) \quad (4)$$

We proposed the following options for the decision:fixed number of steps for all iterations, always performs only n approximation step where n is fixe,choosing the number according to some pre specified pattern, selecting this number adaptively until the utility values stabilize with a given precision as in VI. We can implement that version of PI calling it MPI and compare it to VI, subsequently, we can compare both algorithm to MPI with fixed numbers of approximation steps in Value Determination. We use notions of **threshold** and **precision** both for VI and MPI.

2.3. Convergence

We can give an analysis of convergence of MPI for **discounted** MDP's, therefore in practice, value iteration should never used. We did not follow blindly the description of VI and PI given in and did not calculate the values of two terminal states according to the general formula given there.

2.4. The model

We can study mailing decisions in the catalog sales industry using a sophisticated mathematical model with stochastic responses from potential buyers and with dynamic evolution of a customer base. Their customer base is divided according to a recency, when the consumer placed the last order, frequency, how many orders the consumer placed in a period of time, and monetary, how much money the consumer spent, (RFM) classification which is also used to define the

dynamics of their system. Let $S = \{0, 1, 2, \dots, \dots\}$ be the state space where an element s from S represents the number of loyal customers. The set of allowable actions given state s is: $A_{(S)} = \{(a_L, a_P) : a_L + a_P \leq N, 0 \leq a_L \leq s, a_P \geq 0\}$.

With a_L and a_P representing respectively the number of contacted loyal and potential customers, and N the maximum number of contacts each period. A randomized Markovian decision rule, denoted $d_t = (d_t, L, d_t, P)$ is a mapping of the state space S into the set probability distributions on the action space $A_{(S)}$. The components is associated with: d_t, L – loyal customers, d_t, P – potential customers. A contact policy $\varphi = \{d_t : t = 1, 2, 3, \dots\}$ prescribes at each period t a decision rule d_t . A randomized decision rule in which the probability distribution on the actions is degenerated, a non-randomized rule, will specify for each state how many loyal and potential customers are contacted so that, for any state s from S , $d_{t(S)} = (a_L, a_P)$ from $A_{(S)}$. The contacted potential customers who purchase during period t add to the loyal customer base at the beginning of period $t + 1$. Given the state s_t at period t and a non-randomized decision rule: $d_{t(s_t)} = (a_L(s_t), a_P(s_t))$ from S , one obtains the dynamic equation: $S_t + 1 = S_t + \sum_{i=1}^{a_P(s_t)} \xi_{it}$ where $\xi_{it} = 1$ if the $i - th$ contacted potential customer at period t buys during period t , and 0 if he does not. The probability of buying for a contacted potential customer is the same for any contacted potential customer, we denoted it by β . The ξ_{it} are independent identically distributed random variables with $E(\xi_{it}) = \beta$, for any i and any t . That $s_t + 1 - s_t$ has a binomial distribution with parameters β as the probability of success and $a_P(s_t)$ as the number of trials. Under decision rule $d_t(s_t) = (a_L(s_t), a_P(s_t))$, the expected number of loyal customers at the beginning of period $t + 1$ given the size of the loyal customer base at the beginning of period t , s_t is thus $E_{dt}(s_{t+1}|s_t) = s_t + \beta a_P(s_t)$.

The immediate expected reward or profit if s customers are in the loyal customer base, a_L loyal customers are contacted, and a_P potential customers are contacted is: $r^*(s, a_L, a_P) = s\pi_L + a_L[\pi_L^c - \pi_L] + a_P\pi_P^c$. The reward, immediate expected profit, $r^*(s, a_L, a_P)$ is unbounded on the state space. We define three gain quantities, G_P , G_P^c , and G_L . The first two are the expected net present value from contacting a potential customer with no follow-up in subsequent periods and contacting a potential customer and following up in each subsequent period if the potential customers buy. The third gain $G_L = \pi_L^c - \pi_L$, is the expected incremental profit obtained by contacting a loyal customer over and above the expected profit π_L which is obtained regardless of contact. Contacting a potential customers gives an immediate expected profit of π_L^c . There is an expected revenue of $\beta\pi_L$ in each subsequent period if there are no follow-up contacts. The future gains are discounted with a discount rate of λ so that the expected net present value of contacting a potential customer with no follow-up contacts is:

$$G_P = \pi_P^c + \sum_{l=1}^{\infty} \lambda^l \beta \pi_L = \pi_P^c + \{(\lambda \beta \pi_L) / (1 - \lambda)\}.$$

The expected net present value of contacting a potential customer with follow-up contacts is:

$$G_P = \pi_P^c + \sum_{l=1}^{\infty} \lambda^l \beta \pi_L = G_P + \sum_{l=1}^{\infty} \lambda^l \beta G_L = G_P + \{(\lambda \beta G_L) / (1 - \lambda)\}$$

Our model is a new formulation Markov Decision Process (MDP), the reward function is build like: $r(s, a_L, a_P) = a_L G_L + a_P G_P$, is bounded by $N_X\{G_L, G_P\}$. In our formulation, the entire expected net present value of a potential customer, G_P is accrued during the period of contact, whereas in the original formulation, a contact results in a revenue stream with the same expected net present value, G_P , but which is accrued over time. The expected net present value of using policy φ in the new formula when the are initially s loyal customers is:

$$V_\varphi(s) = E[\sum_{t=1}^{\infty} r(s_t, d_t(s_t) | s_1 = s)]. \text{ The difference must be equal to:}$$

$\sum_{t=1}^{\infty} \lambda^{t-1} s_1 \pi_L = \{(s_1 \pi_L) / (1 - \lambda)\}$, since the revenue derived from the first s_1 loyal customers are taken away in the new formulation. The objective is to find an optimal policy with value $V^*(s)$ where $V^*(s) = \sup_{\varphi} V_\varphi(s)$. We can define a non-randomized stationary policy φ that takes the following form: $\varphi = (d, d, \dots) = (d)^\infty$, where d is a decision rule such that $d(s) = (a_L, a_P)$ s from S.

3. A natural policy to implement

Divide the period into N time slots, for $s \leq N$ loyal customers available, we assume that if they will be contacted it will be happened during the first s time slots, if at all. The starting at time slot 1, contact a loyal customers, a potential customers, no contact, maximizes the immediate expected gain, ignoring any possible interactions with other time slots, do the same for time slot 2, and so one. A policy essentially manages each time slot independently of the others, will referred to as a “**time slot management policy**”. So that $V_\varphi(N, s)$ is the expected net present value of starting with s loyal customers and N time slots, if φ is of the time slot management type, the one has the decomposition:

$$V_\varphi(N, s) = V_\varphi(l, l) + (n - s)V_\varphi(l, 0), s \leq N$$

$$NV_\varphi(l, l), s > N$$

The optimum contact policy, denoted φ^* , will take one of the following forms, depending on the relative values of G_P , G_P^c , G_L and 0. The policy is stationary with the choice of a_L^* and a_P^* as a function of s only.

Theorem: Let G_P , G_P^c and G_L be define as above.

- (a) If $G_P > \max\{G_L, 0\}$, the potential customers are profitable to contact and are preferred to loyal ones so that the policy which chooses $a_P^* = N$ and $a_L^* = 0$ for all states is optimum.

- (b) If $G_P \leq 0$ and $G_L \leq 0$, then it is not profitable to contact either loyal or potential customers so that $a_L^* = a_P^* = 0$ for all states.
- (c) If $G_L > 0$ and $G_P \leq G_L$, then loyal customers are profitable to contact and preferred to potential ones so that $a_L^*(s) = \min\{N, s\}$.
- (i) If in addition $G_P^c > 0$, then potential customers are also profitable to contact and thus any remaining resources should be used to contact them so that $a_P^*(s) = \max\{N - a_L^*(s), 0\}$.
- (ii) If, on the other hand, $G_P^c \leq 0$ or $s \leq N$, then no potential customers are contacted.

Because of the bounded reward function and the finite number of actions available, the optimum value function (1) is the unique solution to the optimality equation:

$$V^*(N, s) = \max_{(a_L, a_P) \text{ from } A_{(s)}} \left\{ a_L G_L + a_P G_P + \lambda E V^* \left(N, s + \sum_{i=1}^{a_P} \xi_{it} \right) \right\} \quad (5)$$

We can verify the optimum computing the expected value function for given policy and check that (2) holds. Let φ^* be the policy where this is accomplished, $a_L^* = \min\{s, N\}$; in this case, the loyal customer will be contacted each period so that: $V_{\varphi^*}(l, l) = \{G_L/(l - \lambda)\}$ (6)

A customer should be contacted depends on the sign of G_P^c . Part (c.i) of theorem considers the case in which $G_P^c > 0$ and so potential customers should be contacted during any time slot for which loyal customers are not available, that is, $a_P^* = \max\{N - s, 0\}$. The value function for φ^* with one time slot can be computed by: $V_{\varphi^*}(l, 0) = G_P + \lambda[\beta V_{\varphi^*}(l, l) + (l - \beta)V_{\varphi^*}(l, 0)]$, then substitute using (6) and solving to obtain:

$$V_{\varphi^*}(1, 0) = \{(G_P + (\lambda/(1 - \lambda))\beta G_L)/(1 - \lambda + \lambda\beta)\} = \{G_P^c/(1 - \lambda + \lambda\beta)\} \quad (7)$$

Since $G_P^c > 0$, any value of (a_L, a_P) which achieves the maximum in (5) must have $a_L + a_P = N$, and then, equation (5) becomes:

$$V^*(N, s) = \max_{0 \leq k \leq s} \{(s - k)G_L + (N - s + k)G_P + \lambda E V^*(N, s + \sum_{i=1}^{N-s+k} \xi_{it})\} \quad (8)$$

To verify that policy φ^* , satisfies (8) is equivalent to show that the maximum is satisfied with $k = 0$. We suppose a positive $k \leq s$. The right hand side of (8), corresponds to a policy φ which contacts k more potential customers in the first period than φ^* would, and then use policy φ^* from period two onward. The incremental expected loss in the first period from using φ instead of φ^* is $kG_L - kG_P$. Policy φ thus gives up in the first period $k(G_L - G_P)$ in hope that this loss will be compensated by the additional loyal customers gained by the extra k contracts of potential customers. Let $X(n)$ be the number of (new) loyal customers which result from n contacts of potential customers; then $X(n)$ has a binomial distribution with parameters β as the probability of success and n as the number of trials. The number of additional loyal customers gained in the first period with policy φ is $X(N - s + k)$ which is equivalent to $X(N - s) + X(k)$ with $X(N - s)$ and $X(k)$ begin independent binomial variables of

parameters $(\beta, N - s)$ and (β, k) . The number of additional loyal customers gained in the first period with policy φ^* is also binomial, say $X(N - s)$, and has the same distribution as $X(N - s)$. We can assume that $X(N - s) = X(N - s)$, and then:

$$Z = \min\{X(N - s) + X(k), N - s\} - X(N - s),$$

Additional loyal customers gained in the second period if φ is used instead of φ^* . The incremental expected gain from the second period on due to the extra k contacts of potential customers is thus: $\beta E(Z)[V_{\varphi^*}(1,1) - V_{\varphi^*}(1,0)]$. Taking the difference between (6) and (7) gives:

$$V_{\varphi^*}(1,1) - V_{\varphi^*}(1,0) = \{(G_L - G_P)/(1 - \lambda + \lambda\beta)\} \quad (9)$$

The expected value of Z can be easily bounded by noting that $Z \leq X(k)$, with positive probability $Z < X(k)$, so that:

$$EZ < k\beta \quad (10)$$

The optimality equations (8) are satisfied by φ^* if the expected loss from contacting k fewer loyal customers during the first period is not covered by the expected gain from the extra loyal customers available in period two onward generated by these contacts:

$$\beta EZ[V_{\varphi^*}(1,1) - V_{\varphi^*}(1,0)] \leq k[G_L - G_P] \quad (11)$$

The left hand side is less than: $\lambda k \beta [(G_L - G_P)/(1 - \lambda + \lambda\beta)]$, since $\lambda\beta/(1 - \lambda + \lambda\beta) < 1$, (8) is verified. In the case when $G_P^c \leq 0$, there is no benefit derived from contacting a potential customer, one only needs to consider (a_L, a_P) with $a_P = 0$, that in case (c), $a_P = \min\{s, N\}$ dominates other choices, hence policy φ^* is optimal. The application of our theorem by returning to the example of the a company providing **Enterprise resource planning (ERP)** integrates internal and external management information. The company plans to contact 200 business organizations each quarter:

Company providing ERP's Immediate Expected Profits		
	Contacted	Non-Contacted
Loyal Customer	$\pi_L^c = 23$	$\pi_L = 15$
Potential Customer	$\pi_P^c = -2$	0

The company realizes that it may be in its best interest to contact them because 5% of them will buy and will be converted into loyal customers. For $\lambda = 0.95$, we can verify that $G_P > \max\{G_L, 0\}$, the company providing ERP is advised to contact 200 potential economic organizations potential customers and no loyal customers each quarter. The expected net present values of contacting a potential customer, G_P and G_P^c , would become time dependent. Let T be the total finite number of periods. We define: $G_P(t) = \pi_P^c + (T - t)^{\beta\pi_L}$ and $G_P^c(t) = G_P + (T - t)^{\beta GL}$ as the expected net present value of contacting a potential customer with no follow up contacts and with follow up contacts respectively.

4.A Theoretical Example⁴

Let $V_\varphi(t, N, s)$ be the expected net present value of a policy φ from starting at period t with s loyal customers and N contacts. Let φ^* be the policy where as many loyal customers as possible are contacted each period and let φ^1 be the policy where one more potential customer is contacted in the first period then φ^* would, and then using φ^* from the next period onward. If $G_L > G_P(1)$, one would expect that a maximum number of loyal customers should be contacted on the first period. One should have: $V_{\varphi^*}(t, N, s) \geq V_{\varphi^1}(t, N, s)$, for any combination of t , N , and s , we can show that this is not always true. Choose $N = 2$, $T = 2$, and $\lambda = 1$, for $t = 1$ and $s = 1$, $V_{\varphi^*}(1, 2, 1) = G_L + G_P(1) + \beta 2G_L + (1 - \beta[G_L + G_P(2)])$ and: $V_{\varphi^1}(1, 2, 1) = 2G_P(1) + \beta^2 2G_L + 2\beta(1 - \beta)2G_L + (1 - \beta)^2[G_L + G_P(2)]$. With $\beta = .5$ and $\pi_P^c = 0$, so that $G_P(2) = 0$, $V_{\varphi^*}(1, 2, 1) = 2.5G_L + G_P(1)$ and $V_{\varphi^1}(1, 2, 1) = 1.75G_L + 2G_P(1)$, and so: $V_{\varphi^*}(1, 2, 1) < V_{\varphi^1}(1, 2, 1)$, whenever $.75G_L < G_P(1)$. By choosing G_L and $G_P(1)$ so that $.75G_L < G_P(1) < G_L$, it is possible to prefer contacting potential customers even though, at period $t = 1$.

5.Cash Flow Model

The optimum policies of **Theorem** work well when survival is not affected by short run considerations such as cash flow. For a fixed amount of revenue, $\beta\pi_L$ at each period following the contact, this happens when the manager sells a product or a service that requires follow ups. We can have an example include software-ERP, that is sold with maintenance contract and natural water distributors. It is still optimum to contact as many potential customers as possible whenever loyal customers are not profitable to contact ($G_L \leq 0$). When a loyal customers becomes profitable to contact, numerical examples exist where a loyal customer is more profitable to contact than a potential customer, $G_L > G_P$, but potential customers only should be contacted. Let $S = \{(c, s) : c \text{ from } R, s \text{ from } I\}$ be the state space where c represents the cash balance, s is the number of loyal customers, R is the set of non-negative real numbers, and I the set of non-negative integers. The set of allowable actions given states (c, s) is thus:

$$A(c, s) = \{(a_L, a_P) : a_L + a_P \leq N, a_L\alpha_L + a_P\alpha_P \leq c, 0 \leq a_L \leq s, a_P \geq 0\}$$

With a_L and a_P representing the number of loyal and potential customer contacted, α_L and α_P representing respectively the cost of contacting a loyal potential customer, N the maximum number of contacts for each period. A contact

⁴ Y.Baba, R.F.Engle, D.Kraft, and K.F.Korner, Multivariante Simultaneous Generalized ARCH. Technical Report, Departament of Economics, University of California San Diego, 1989
M.Lévesque, Model of entrepreneurial decision:A dynamic Programming Approach, Ph.D.thesis, The University of British Columbia, 1998

policy, $\varphi = \{d_t : t = 1, 2, \dots\}$, possibly randomize, specifies an action $d(c, s) = (a_L, a_P)$ from $A(c, s)$ for each state (c, s) and time period t , where $(a_L, a_P) = (0, 0)$, whenever c reaches 0. Because the future stream of revenues from contacting a potential customer will be accrued in the period of contact, state $(0, s)$ is absorbent for any s from I . Given the state (c_t, s_t) at period t and a decision rule : $d(c_t, s_t) = (a_L(c_t, s_t), a_P)$, (c_t, s_t) from S , one obtains the dynamic equation $s_{t+1} = s_t + \sum_{i=1}^{a_P(c_t, s_t)} \xi_{it}$, where $\xi_{it} = 1$, if the $i - th$ contacted potential customer at period t buys during period t , and 0 if he does not. We can generate profit, contacted loyal and potential customers, that will add to the current cash balance. Let ρ be the fixed revenue from a sale, $\eta_{it} = 1$ if the $i - th$ contacted loyal customers buys, and 0 otherwise, $\xi_{it} = 1$, if the $i - th$ contacted potential customers buys, and 0 if he does not buy, and λ be the discount rate; one obtains a second dynamic equation:

$$c_{t+1} = c_t + \rho \sum_{i=1}^{aL(c_t, s_t)} \eta_{it} + [\rho + ((\lambda)/(1 - \lambda))\pi_L] \sum_{i=1}^{aP(c_t, s_t)} \xi_{it} - [\alpha_L a_L(c_t, s_t) + \alpha_P a_P(c_t, s_t)]$$

As defined earlier, the ξ_{it} s are independent identically distributed random variables with $E(\xi_{it}) = \beta$. Under decision rule $d_t(c_t, s_t) = (a_L(c_t, s_t), a_P(c_t, s_t))$, the expected number of loyal customer at the beginning of the period $t + 1$, given state (c_t, s_t) at the beginning of period t , is: $E_{dt}(s_{t+1}|c_t, s_t) = s_t + \beta a_P(c_t, s_t)$. The η_{it} s are independent identically distributed random variables with $E(\eta_{it}) = \gamma$ where γ represents the incremental portion of loyal customers who would not purchase if not contacted, the benefit from non contacted loyal customers who purchase is deterministically accounted at the period at which they move from potential customer base to loyal customer base.

$$G_P = \beta\rho - \alpha_P + \beta\{(\lambda)/(1 - \lambda)\}\pi_L = \pi_P^c + \{(\lambda)/(1 - \lambda)\}\beta\pi_L$$

It follows that the expected cash balance at the beginning of period $t + 1$, and given state (c_t, s_t) at the beginning of period t , is: $E_{dt}(c_{t+1}|c_t, s_t) = c_t + a_L(c_t, s_t)G_t + a_P(c_t, s_t)G_P$. The manager's objective is to maximize the total expected cash on hand at any period t , that is: $\text{Max}_{a_L, a_P} E_{c_t}$, for any t . If P potential customer should be contacted where P is the integer part of c/α_P , when the value $G_P \leq 0$, nobody should be contacted.

6. Study Case, Numerical Example

We can realize that upgrading is not profitable business⁵, but servicing is, $G_L \leq 0$, $G_P > 0$. If we apply the theorem above, the computer expert should use all of his cash reserves and time on contacting companies that did not acquire the software yet. The manager has a limit budget and a limited time for demos.

⁵ M. Levesque, Model of Entrepreneurial Decision: A Dynamic Programming Approach, Ph.D. thesis, The University of British Columbia, 1998

Assume that: $\pi_L^c = 8$, $\pi_L = 3$, $\pi_P^c = 2$, $\beta = .1$, $\alpha_L = 10$, $\alpha_P = 1$, $N \geq 0$, $\lambda = .8$, and, at the current period, $c_t = 10$ and $s_t \geq 1$. Corresponding to case (c) of the theorem above, $G_L = 5$, $G_P = 3.2$, so that $G_L > 0$ and $G_P \leq G_L$. The entrepreneur can only afford to contact one loyal customer who will generate a total surplus benefit of 5. By contacting potential customers, he can afford 10 contacts generating a total expected benefit of 32. We present the development of the VI algorithm as an AI⁶ engine for use in real time 2D, like computer applications. The human opponents were more difficult to anticipate and were more challenging, in comparison to their NPC counterparts. The cell containing the enemy will have a reward value of -1 and cell containing the home (or goal) will have a reward value of +1, regardless of cell's other properties. A schematic description of VI (value iteration) is given below: **function** VALUE ITERATION (M, R) **returns** a utility function

```

inputs M a transition model (transition probabilities)
        R a reward function on states
local variables: U, a utility function, initially identical to R
                    UI, a utility function, initially identical to R
                    AllStatesChanged, an all states changed flag initially equal
                    to false
                    NumStatesBelowZero, stores the number of states below zero
                    LastNumStatesBelowZero, stores the last number of states
                    below zero
repeat
for each state i do
     $U1[i] \leftarrow R[i] + \max \sum_j M_{ij} * U[j]$ 
end
     $U \leftarrow UI$ 
    LastNumStateBelowZero=NumStateBelowZero
for each state i do
    if  $U(i) \leq 0$  then
        if  $U(i) = \text{AnyTerminalStates}$  then
            NumStateBelowZero=NumStateBelowZero+1
        endif
    endif
end
if LastNumStateBelowZero=NumStateBelowZero then
    AllStatesChanged=true
end if
until AllStatesChanged=true
return U

```

The equation in that loop can also be seen below.

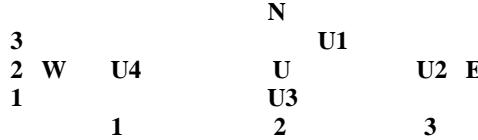
$$U1[i] \leftarrow R[i] + \max_a \sum_j M_{ij}^a U[j]$$

Where $U1[i]$ is the new utility value estimate for a cell in the grid and $R[i]$ is the reward value, \max_a is selection of the utility that returns the maximum

⁶ AI, artificial intelligence

value; I is the index of all cells in the grid and j is the index of the number of cells surrounding i . M is the transition model, the probability of moving in a certain direction, U is the current utilities.

$U = \text{Utility}, P = \text{Probability}$



$PA = 0.8$.

$PB = 0.1$.

$PC = 0.1$.

$PD = 0.0$.

To work out the utility of cell 2,2 the following will be connected:

$$\text{Action } N = PA * U1 + PB * U2 + PC * U4 + PD * U3$$

$$\text{Action } E = PA * U2 + PB * U3 + PC * U1 + PD * U4$$

$$\text{Action } S = PA * U * PB * U4 + PC * U2 + PD * U1$$

$$\text{Action } W = PA * U4 + PB * U1 + PC * U3 + PD * U2$$

The process is repeated until stopping criterion is met.

6.1. Experimental Results

For all experiments the following things were kept the same: there were two goal states, +1(home) and -1(enemy), and there was a cost of -0.0000001 for all non goal states. The probability of moving in the intended direction was 0.8 and the size of the software application was 10x10. The HD columns in tables 1 and 2 stand for hamming distance between the generated policy and the optimum policy.

The environment map and the results produced from experiments conducted on the map

Convergence Threshold	No of Iterations to Convergence	Agent Successfully Navigated Map	No of steps Agent took to Navigated Map	HD
1.00	1	No	N/A	70
0.5	4	No	N/A	61
0.25	8	No	N/A	39
0.125	12	No	N/A	11
0.0625	15	No	N/A	2
0.031250	19	Yes	18	0
0.015625	21	Yes	18	0
0.007812	23	Yes	18	0
0.003906	24	Yes	18	0
0.001953	25	Yes	18	0
0.0000977	26	Yes	18	0
0.000488	27	Yes	18	0
0.000122	29	Yes	18	0
0.000061	30	Yes	18	0
0.000031	30	Yes	18	0

0.000015	31	Yes	18	0
0.000008	32	Yes	18	0
0.0000000	59	Yes	18	1
VI	18	Yes	18	1

The environment map and the results produced from experiments conducted on the map

Convergence Threshold	No of Iterations to Convergence	Agent Successfully Navigated Map	No of steps Agent Took To Navigate Map	HD
1.00	1	No	N/A	27
0.5	4	No	N/A	27
0.25	7	No	N/A	27
0.125	10	No	N/A	26
0.0625	14	No	N/A	22
0.031250	19	No	N/A	15
0.015625	20	No	N/A	12
0.007812	22	No	N/A	14
0.003906	27	Yes	32	9
0.001953	34	Yes	36	3
0.00097	40	Yes	36	0
0.000488	46	Yes	36	0
0.000244	50	Yes	36	0
0.000122	53	Yes	36	0
0.000061	56	Yes	36	0
0.000031	60	Yes	36	0
0.000015	63	Yes	36	0
0.000008	66	Yes	36	0
0.000000	210	Yes	36	-
VI	32	Yes	36	4

The hamming distance is very small, but on a large map, the difference might become significant.

7. Conclusions

The managers involved in their first start ups may know a lot about their products and their characteristics, but often little understanding regarding prospecting for customers. The optimum policies were of the form, **among the contacts available during a time slot, choose the one which has the largest expected benefit**. This happened when the expected proportion of buyers among contacted potential customers (β), or discount rate (λ), or the immediate expected profit from not contacting a loyal customer (π_L) was large enough. If loyal customers were profitable to contact and preferred to potential ones, then a mixed contact policy where both loyal and potential customers are contacted was optimum whenever β, λ, π_L is large enough and the number of loyal customers is smaller than the number of time slots. The manager had access to enough capital to contact as many potential buyers as profitable. Only for the stochastic formulation, the managers had no unutilized credit and could not contact unless expected gains from buyers were accumulated. The formal deviations of the optimum policies will involve more decision variables since, to express the

dynamic of debit or credit, one will need know how many contacts are made with cash reserves and how many made with credit. To be of real value, one must validate them. Our contribution resides in proposing an approach where the lifetime value of each potential contact can be evaluated and a simple optimum policy can be implemented of each potential contact evaluated and a simple optimum contact policy can be implemented by ordering these lifetime values. This database could be used to estimate the number of contacts the manager can make over time, the probability of purchasing for a potential customer and the expected profit from each type of potential buyers. We present an analytical approach that provided a new focus to managerial research. Analytical methods are useful in that they provide sharp answers to focused questions and thus extend one's knowledge of the issue being studied. When the time dimension take center stage in the decision process, analytical methods permit to explicitly model the evolution of the business process as it varies over times. An analytical approach to management can also provide clearer hypothese to test, thus guiding empirical studies of various managerial processes. The optimization perspective allows decision makers to find the optimum strategy rather than limiting the search for the best action to a small set of possible actions. An analytical approach contributes to a reorientation of the management scientific literature which has been dominated by empirical work, we encourages researches to fulfil the need for a multidisciplinary emphasis in any research work on management science.

R E F E R E N C E S

- [1]. *F.Alizadeh and D.Goldfarb*, Second-order cone programming, Mathematical , 95:1(2003)
- [2]. *Y.Baba, R.F. Engle, D.Kraft, and K.F.Kroner*, Multivariante Simultaneous Generalized ARCH. Technical report, Department of Economics, University of California San Diego (1989)
- [3]. *J.R.Birge and F.Louveaux*, Introduction to Stochastic Programming, New York:Springer, 1997
- [4]. *F.Black and R.Litterman*, Global portfolio optimization, Financial Analysts Journal, 48:5 (1992)
- [5]. *J.Gondzio and R.Kouwenberg*, High performance computing for asset Liability management, Operations Research, 49 (2001)
- [6]. *E.Guslitser*, Uncertainty – immunized solutions in linear programming Master's thesis, The Technion, Haifa (2002)
- [7].*R.Kouwenberg*, Scenario generation and stochastic programming models for asset liability management, European Journal of Operational Research , 134 (2001)
- [8]. *M.Levesque*, Model of entrepreneurial decisions: A Dynamic Programming Approach, Ph.D. thesis, The University of British Columbia, 1998