Adaptive Optimization in Stochastic Systems via the Variational Technique

Nicholas A. Nechval Konstantin N. Nechval Boris M. Heimanis

Department of Applied Mathematics
Aviation University of Riga
Lomonosov Street 1, LV-1019 Riga, Latvia
E-mail: rau@rau.lv

Abstract

This paper deals with the stochastic adaptive linear quadratic optimal control problems which have been an active area of research for many years. It has been known that these problems could be treated by dynamic programming. However, it has been conceded that explicit solution of the dynamic programming equations for these problems is generally not possible and that numerical solution of these equations is a difficult computational procedure. This has led to many approximation techniques. In the paper, a variational approach is used to obtain optimality conditions for the stochastic linear quadratic adaptive control problems. These conditions lead to an algorithm for computing optimal control laws which differs from the dynamic programming algorithm. If the unknown parameters enter into the state equation additively, and the prior distribution of the unknown parameters is normal, the algorithm can be carried out in closed form. The examples are given to illustrate the proposed technique.

Keywords: Stochastic System, Adaptive Optimization, Variational Technique

1 Introduction

For a long time control theorists and control engineers have dreamt of a controller that does not need to be tuned. This type of controller has been given many different names, for instance adaptive, self-organizing, self-optimizing and learning controller. The ultimate solution has not yet been found and it is questionable whether it exists. Many different solutions to the adaptive control problem have been suggested. Some solutions are designed from a very practical point of view, while others are based on highly technical theory.

In this paper, the above problem is considered in terms of the standard stochastic linear quadratic discrete-time adaptive control problem. Stochastic adaptive liner quadratic optimal control problems have been an active area of research for many years. Excellent surveys of work on discrete-time adaptive control problems are given in Nechval (1984). It has been known that these problems could be treated by dynamic programming (see Aoki, 1967; Yakowitz, 1969; Åström, 1970; Bertsekas, 1976; Nechval, 1984, 1988; Nechval et al., 1997). However, it has been conceded that explicit solution of the dynamic programming equations for these problems is generally not possible and that numerical solution of these equations is a difficult computational procedure. This has led to many approximation techniques.

The aim of the present paper is to use a variational approach in order to obtain optimality conditions for the stochastic adaptive control problems. These conditions lead to an algorithm for determining optimal controls which differs from the dynamic programming algorithm. The algorithm obtained requires integration of functions with respect to normal densities and solving some equations at each step.

2 Problem Statement

Consider the standard stochastic linear quadratic discrete-time adaptive control problem in which the state equation of the system is given by

$$x_i = A(\theta)x_{i-1} + B(\theta)u_{i-1} + C(\theta) + w_i, \quad i=1, \dots ,N, \tag{1}$$

where it is desired to find a control sequence $\mathbf{u}=(u_0, u_1, \dots ,u_{N-1})$ which minimizes the performance criterion

$$J(\mathbf{u}) = E\left\{ x_N' G_N x_N + \sum_{i=0}^{N-1}(x_i' G_i x_i + u_i' H_i u_i) \right\} \tag{2}$$

over all choices of admissible controls. In (1), x_i is the n-dimensional state vector of the system. The initial state x_0 is assumed to be a given random vector. The quantity u_i is an m-dimensional control vector. The w_i's are the disturbances, which for convenience are taken to be normal random vectors with mean zero and covariance the identity matrix. The quantities $A(\theta)$, $B(\theta)$ are respectively $n \times n$- and $n \times m$-dimensional matrices and $C(\theta)$ is an n-dimensional vector. Each of these depends on a q-dimensional vector of unknown parameters θ. The quantities G_i, $\forall i=0,1, \dots ,N$, are non-negative definite $n \times n$-matrices and H_i, $\forall i=0,1, \dots ,N-1$, are non-negative definite $n \times n$-matrices.

The class of admissible controls at time i is all Borel measurable functions $u_i=u_i(x_0, x_1, \dots ,x_i)$ of past states $(x_0, x_1, \dots ,x_i)$. It is assumed that the unknown parameters can be modeled as a random vector with prior probability measure $P(d\theta)$ and that the form of this prior probability is known to the controller. We shall let $P(dx_0)$ denote the probability measure of the initial state x_0 and assume that θ, x_0, and w_i, i=1, ... ,N, are mutually independent.

The problem is adaptive in that the control u_i is to be chosen as a function $u_i=u_i(x_0, x_1, \ldots, x_i)$ of the past states but without knowledge of the parameter θ.

3 Optimality Conditions

The optimality conditions for the above problem are given by the following theorem.

Theorem 1. A necessary condition for $\mathbf{u}^* = (u_0^*, u_1^*, \ldots, u_{N-1}^*)$ to be an optimal control is that for each $j \in \{0, \ldots, N-1\}$

$$E\left\{\left(2H_j u_j^* + B(\theta)' w_{j+1}\left[x_N' G_N x_N + \sum_{i=j+1}^{N-1}(x_i' G_i x_i + u_i^{*\prime} H_i u_i^*)\right]\right); x_0, \ldots, x_j\right\} = 0. \tag{3}$$

Proof. Our assumptions imply that the joint probability distribution of $(\theta, x_0, w_1, \ldots, w_N)$ is given by

$$P(d\theta)P(dx_0)\prod_{i=1}^{N}(2\pi)^{-n/2}\exp\left(-\frac{1}{2}\|w_i\|^2\right)dw_i. \tag{4}$$

Consider the mapping which takes $(\theta, x_0, w_1, \ldots, w_N)$ into $(\theta, x_0, x_1, \ldots, x_N)$ in which, for $i \geq 1$, x_i and w_i are related through the state equation (1). Since the determinant of the Jacobian of this mapping equals one, change of variables rules for probability densities and (4) imply that the joint probability distribution of $(\theta, x_0, x_1, \ldots, x_N)$ is given by

$$P(d\theta)P(dx_0)\prod_{i=1}^{N}(2\pi)^{-n/2}\exp\left(-\frac{1}{2}\|x_i - A(\theta)x_{i-1} - B(\theta)u_{i-1} - C(\theta)\|^2\right)dx_i. \tag{5}$$

It follows from (5) that the conditional density of $x_1, \ldots, x_N$ given θ and x_0 is given by

$$P^{\mathbf{u}}(x_1, \ldots, x_N; \theta, x_0) = (2\pi)^{-nN/2}\exp\left(-\frac{1}{2}\sum_{i=1}^{N}\|x_i - A(\theta)x_{i-1} - B(\theta)u_{i-1} - C(\theta)\|^2\right). \tag{6}$$

Then from (2), (5) and (6)

$$J(\mathbf{u})=\int\cdots\int\left[x_N'G_Nx_N+\sum_{i=0}^{N-1}(x_i'G_ix_i+u_i'H_iu_i)\right]P^{\mathbf{u}}(x_1,\ldots,x_N;\theta,x_0)P(d\theta)P(dx_0)dx_1\ldots dx_N. \tag{7}$$

We shall now proceed to deduce our optimality conditions by taking variations of $J(\mathbf{u})$ as is done in ordinary calculus of variations. Let j be some integer between 0 and N-1, and let $\mathbf{v}=\{v_i(x_0,\ldots,x_i)\}$ be an admissible control which satisfies $v_i(x_0,\ldots,x_i)=0$ if $i\neq j$. Consider the control $\mathbf{u}^*+\varepsilon\mathbf{v}$ where ε is a scalar and $\mathbf{u}^*$ is an optimal control. Since $\mathbf{u}^*$ is an optimal control, $J(\mathbf{u}^*+\varepsilon\mathbf{v})$ considered as a function of ε must have a minimum at $\varepsilon=0$. Thus its derivative with respect to ε must vanish there. Differentiating (7) with $\mathbf{u}$ replaced by $\mathbf{u}^*+\varepsilon\mathbf{v}$ under the integral sign, taking into account (6), gives

$$0=\frac{d}{d\varepsilon}J(\mathbf{u}^*+\varepsilon\mathbf{v})\Big|_{\varepsilon=0}=\int\cdots\int\Big(2v_j'H_ju_j^*+\left[x_N'G_Nx_N+\sum_{i=0}^{N-1}(x_i'G_ix_i+u_i^{*\prime}H_iu_i^*)\right]$$

$$\cdot\, v_j'B(\theta)'(x_{j+1}-A(\theta)x_j-B(\theta)u_j^*-C(\theta))\Big)P^{\mathbf{u}^*}(x_1,\ldots,x_N;\theta,x_0)\,P(d\theta)P(dx_0)dx_1\ldots dx_N. \tag{8}$$

Equation (8) can be written using (1) as

$$E\left\{v_j'\left(2H_ju_j^*+B(\theta)'w_{j-1}\left[x_N'G_Nx_N+\sum_{i=0}^{N-1}(x_i'G_ix_i+u_i^{*\prime}H_iu_i^*)\right]\right)\right\}=0. \tag{9}$$

Since w_{j-1} and $\theta, x_1,\ldots,x_j$ are independent, it follows that

$$E\left\{v_j'B(\theta)'w_{j+1}\sum_{i=0}^{j}(x_i'G_ix_i+u_i^{*\prime}H_iu_i^*)\right\}=E\left\{v_j'B(\theta)'\sum_{i=0}^{j}(x_i'G_ix_i+u_i^{*\prime}H_iu_i^*)\right\}E\{w_{j+1}\}=0. \tag{10}$$

Thus from (9) and (10)

$$E\left\{v_j'\left(2H_ju_j^*+B(\theta)'w_{j+1}\left[x_N'G_Nx_N+\sum_{i=j+1}^{N-1}(x_i'G_ix_i+u_i^{*\prime}H_iu_i^*)\right]\right)\right\}=0. \tag{11}$$

Since (11) must hold for every $v_j(x_0,\ldots x_j)$, a standard argument using the definition of conditional expectation implies (3) follows from (11). □

4 Determining Optimal Controls

Consider trying to use (3) to determine **u*** in an explicit form. To begin this, define

$$g_j(\theta, x_0, \dots, x_j) = E\left\{ x'_N G_N x_N + \sum_{i=j}^{N-1} (x'_i G_i x_i + u_i^{*\prime} H_i u_i^*); \theta, x_0, \dots, x_j \right\}. \tag{12}$$

Notice that $g_j(\theta, x_0, \dots, x_j)$ is the conditional remaining cost from stage j onward, given both the unobserved parametric vector θ and the past states $x_0, \dots, x_j$. We easily see from the law of iterated conditional expectations that

$$g_j(\theta, x_0, \dots, x_j) = x'_j G_j x_j + u_j^{*\prime} H_j u_j^* + E\{g_{j+1}(\theta, x_0, \dots, x_{j+1}); \theta, x_0, \dots, x_j\}. \tag{13}$$

and that the terminal condition

$$g_N(\theta, x_0, \dots, x_N) = x'_N G_N x_N \tag{14}$$

is satisfied.

The law of iterated conditional expectations and (1) imply the left side of (3) is given by

$$2H_j u_j^* + E\{B(\theta)' w_{j+1} g_{j+1}(\theta, x_0, \dots, x_j, A(\theta)x_j + B(\theta)u_j^* + C(\theta) + w_{j-1}); x_0, \dots, x_j\}. \tag{15}$$

The quantity under the conditional expectation sign in (15) is a function of $x_0, \dots, x_j$, θ, u_j^*, and w_{j-1}. The control u_j^* is to be chosen as a function of $x_0, \dots, x_j$. Thus this conditional expectation can be evaluated by integration with respect to the conditional probability distribution of θ and w_{j-1} given $x_0, \dots, x_j$. Since w_{j+1} is independent of θ and $x_0, \dots, x_j$, the joint conditional probability distribution of θ and w_{j+1} is the product of the probability distribution of w_{j+1} and the conditional probability distribution of θ given $x_0, \dots, x_j$.

The following theorem holds for the conditional probability distribution of θ given $x_0, \dots, x_j$ when $A(\theta)$, $B(\theta)$ and $C(\theta)$ are affine. We shall need to define some notation for affine functions used in the statement of the theorem. An $n \times m$ -matrix valued function $D(\theta)$ of a q-dimensional vector variable θ will be called affine if there are $n \times m$ -dimensional matrices D_1 and D_{2i}, such that

$$D(\theta) = D_1 + \sum_{i=1}^{q} D_{2i}\theta_i. \tag{16}$$

This can be written in matrix notation by

$$D(\theta) = D_1 + D_2\Theta, \tag{17}$$

where

$$D_2 = (D_{21}, D_{22}, \dots, D_{2q}) \tag{18}$$

is the $n \times mq$ -matrix,

$$\Theta = (\theta_1 I, \theta_2 I, \dots, \theta_q I)' \tag{19}$$

is the $mq \times m$ -matrix, and I is the $m \times m$ -identity matrix.

Theorem 2. If $A(\theta)$, $B(\theta)$, $C(\theta)$ are affine and the prior distribution $P(d\theta)$ of θ is normal with mean vector μ and covariance matrix Ω, then the conditional probability distribution of θ given $x_0, \dots, x_j$, $P(d\theta; x_0, \dots, x_j)$, is normal with mean vector $\hat{\theta}_j$ and covariance matrix $\hat{Q}_j$ which satisfy the difference equations,

$$\hat{\theta}_j = \hat{\theta}_{j-1} + \hat{Q}_j(A_2X_{j-1} + B_2U_{j-1} + C_2)'\left[(A_2X_{j-1}+B_2U_{j-1}+C_2)(\theta-\hat{\theta}_{j-1})+w_j\right], \tag{20}$$

$$\hat{Q}_j = \hat{Q}_{j-1}\left[I + (A_2X_{j-1}+B_2U_{j-1}+C_2)'(A_2X_{j-1}+B_2U_{j-1}+C_2)\hat{Q}_{j-1}\right]^{-1}, \tag{21}$$

with initial conditions $\hat{\theta}_0 = \mu$ and $\hat{Q}_0 = \Omega$, where x and X, u and U are related through (19).

Proof. This follows in the same manner as that of the Kalman filter and so it is omitted here. □

Now (13), (14), (15) and Theorem 2 allow one to compute $u_j(x_0, \dots, x_j)$ and $g_j(\theta, x_0, \dots, x_j)$ in order to find a control which satisfies (3).

5 Examples

Three simple examples will be given which illustrate the approach described above.

Example 1. Consider the scalar linear stochastic system with state equation

$$x_j = x_{j-1} + \theta u_{j-1} + w_j, \tag{22}$$

where θ is an unknown constant and $\{w_j\}$ is a sequence of independent equally distributed gaussian random variables with zero mean values and standard deviation $\sigma=1$. We want to select u_{N-1} in order to minimize the loss function

$$J(u_{N-1}) = E\{x_N^2\}. \tag{23}$$

The unknown parameter θ in (22) can be estimated using the least squares method (see Åström and Eykhoff, 1971),

$$\hat{\theta}_{N-1} = E\{\theta; x_0, \dots, x_{N-1}, u_0, \dots, u_{N-2}\} = \sum_{j=1}^{N-1}(x_j - x_{j-1})u_{j-1} \Big/ \sum_{j=1}^{N-1} u_{j-1}^2, \tag{24}$$

$$\hat{Q}_{N-1} = \mathrm{Var}\{\theta; x_0, \dots, x_{N-1}, u_0, \dots, u_{N-2}\} = \sigma^2 \Big/ \sum_{j=1}^{N-1} u_{j-1}^2 = \left[\sum_{j=1}^{N-1} u_{j-1}^2\right]^{-1}. \tag{25}$$

If θ were known then the optimal loss is given as

$$\min_{u_{N-1}} J(u_{N-1}) = \min_{u_{N-1}} E\{(x_{N-1} + \theta u_{N-1} + w_N)^2\} = \min_{u_{N-1}} \{(x_{N-1} + \theta u_{N-1})^2 + \sigma^2\} = \sigma^2 = 1. \tag{26}$$

To obtain this we have used that w_N is independent of $\theta, x_0, \dots, x_{N-1}, u_0, \dots, u_{N-2}$. The optimal control law is

$$u_{N-1}^\circ = - x_{N-1}/\theta. \tag{27}$$

If the estimated value, $\hat{\theta}_{N-1}$, is used in (27) instead of the true value we get

$$\hat{u}_{N-1}^\circ = - x_{N-1}/\hat{\theta}_{N-1}, \tag{28}$$

i.e. we have assumed that the certainty equivalence principle can be used. The loss when using (28) will be

$$J(\hat{u}^{\circ}_{N-1}) = E\left\{\left(x_{N-1} - \frac{\theta}{\hat{\theta}_{N-1}}x_{N-1} + w_N\right)^2\right\} = \frac{\hat{Q}_{N-1}}{\hat{\theta}^2_{N-1}}x^2_{N-1} + \sigma^2 = \frac{\hat{Q}_{N-1}}{\hat{\theta}^2_{N-1}}x^2_{N-1} + 1. \quad (29)$$

To get the last equality the standard formula

$$E\{\theta^2\} = \left(E\{\theta\}\right)^2 + \mathrm{Var}\{\theta\} \quad (30)$$

has been used. The loss has increased with the term

$$\frac{\hat{Q}_{N-1}}{\hat{\theta}^2_{N-1}}x^2_{N-1} \quad (31)$$

compared with the optimal loss when θ was known. The control law (28) does not minimize (23) because

$$\min_{u_{N-1}} J(u_{N-1}) = \min_{u_{N-1}} E\left\{\left(x_{N-1}+\theta u_{N-1}+w_N\right)^2\right\} = \min_{u_{N-1}} \left\{\left(x_{N-1}+\hat{\theta}_{N-1}u_{N-1}\right)^2 + \hat{Q}_{N-1}u^2_{N-1}+\sigma^2\right\}$$

$$= \frac{\hat{Q}_{N-1}}{\hat{\theta}^2_{N-1}+\hat{Q}_{N-1}}x^2_{N-1} + \sigma^2 = \frac{\hat{Q}_{N-1}}{\hat{\theta}^2_{N-1}+\hat{Q}_{N-1}}x^2_{N-1} + 1 \quad (32)$$

and the minimum is assumed for the control law

$$u^*_{N-1} = -\frac{\hat{\theta}_{N-1}}{\hat{\theta}^2_{N-1}+\hat{Q}_{N-1}}x_{N-1}. \quad (33)$$

The loss in eqn. (32) is less than in (29) since $\hat{Q}_{N-1} \geq 0$. The first term in (32) is the loss due to the uncertainty of the parameter and the second term is due to the process noise w_{N-1}.

The optimal controller (33) is cautious since it considers the inaccuracy of the estimate of θ. If $\hat{Q}_{N-1}\to 0$ then (28) and (33) will be the same and the loss approaches the optimal loss for known θ, (26).

Another way to obtain the optimal control law (33) is to use the variational approach presented here. Then the estimation equation (20) and (21) are

$$\hat{\theta}_j = \hat{\theta}_{j-1} + \hat{Q}_j\left[u_{j-1}^2(\theta-\hat{\theta}_{j-1}) + u_{j-1}w_j\right], \tag{34}$$

$$\hat{Q}_j = \hat{Q}_{j-1}\left(1 + u_{j-1}^2\hat{Q}_{j-1}\right)^{-1}. \tag{35}$$

If $\hat{\theta}_0=\mu=0$ and $\hat{Q}_0=\Omega\to\infty$, it follows from (22), (34) and (35) that

$$\hat{\theta}_{N-1} = \frac{\sum_{j=1}^{N-1} u_{j-1}(x_j-x_{j-1})}{\sum_{j=1}^{N-1} u_{j-1}^2}, \tag{36}$$

$$\hat{Q}_{N-1} = \left[\sum_{j=1}^{N-1} u_{j-1}^2\right]^{-1}. \tag{37}$$

Now (33) follows immediately from (3). It can be shown by using (33), (36) and (37) that

$$u_{N-1}^* = -x_{N-1}\left[\frac{\sum_{j=1}^{N-1} u_{j-1}(x_j-x_{j-1})}{\sum_{j=1}^{N-1} u_{j-1}^2} + \frac{1}{\sum_{j=1}^{N-1} u_{j-1}(x_j-x_{j-1})}\right]^{-1}. \tag{38}$$

Example 2. Consider the scalar linear stochastic system whose state equation is

$$x_j = \theta x_{j-1} + u_{j-1} + w_j. \tag{39}$$

in which θ is an unknown parameter. It is desired to minimize the performance criterion

$$J(u_{N-1}) = E\{x_N^2\}. \tag{40}$$

Assume that the prior density of θ is normal with mean μ and variance Ω, the disturbances w_i, i=1, ... ,N, normal with mean zero and variance 1 and θ, w_i, i=1,... ,N are mutually independent.

In this example the estimation equations (20) and (21) are

$$\hat{\theta}_j = \hat{\theta}_{j-1} + \hat{Q}_j x_{j-1}\left[x_{j-1}(\theta-\hat{\theta}_{j-1}) + w_j\right], \tag{41}$$

$$\hat{Q}_j = \hat{Q}_{j-1}\left(1 + x_{j-1}^2\hat{Q}_{j-1}\right)^{-1}. \tag{42}$$

If $\hat{\theta}_0=\mu=0$ and $\hat{Q}_0=\Omega\to\infty$, it follows from (39), (41) and (42) that

$$\hat{\theta}_{N-1} = \frac{\sum_{j=1}^{N-1} x_{j-1}(x_j-u_{j-1})}{\sum_{j=1}^{N-1} x_{j-1}^2}, \tag{43}$$

$$\hat{Q}_{N-1} = \left[\sum_{j=1}^{N-1} x_{j-1}^2\right]^{-1}. \tag{44}$$

It can be shown in this case by using (3), (43) and (44) that

$$u_{N-1}^* = -\frac{\hat{\theta}_{N-1}x_{N-1}}{2} = -\frac{x_{N-1}\sum_{j=1}^{N-1} x_{j-1}(x_j-u_{j-1})}{2\sum_{j=1}^{N-1} x_{j-1}^2}. \tag{45}$$

Example 3. Consider the scalar linear stochastic system whose state equation is

$$x_j = x_{j-1} + u_{j-1} + \theta + w_j. \tag{46}$$

in which θ is an unknown parameter. Suppose it is desired to minimize the performance criterion

$$J(\mathbf{u}) = E\left\{x_N^2 + \sum_{i=0}^{N-1}(x_i^2 + u_i^2)\right\}. \tag{47}$$

Assume that the prior density of θ is normal with mean μ and variance Ω, the disturbances w_i, i=1, ... ,N, normal with mean zero and variance 1 and θ, w_i, i=1, ... ,N, are mutually independent.

In this example the estimation equations (20) and (21) are

$$\hat{\theta}_j = \hat{\theta}_{j-1} + \hat{Q}_j(\theta - \hat{\theta}_{j-1} + w_j), \tag{48}$$

$$\hat{Q}_j = \hat{Q}_{j-1}(1 + \hat{Q}_{j-1})^{-1}. \tag{49}$$

If $\hat{\theta}_0=\mu=0$ and $\hat{Q}_0=\Omega\rightarrow\infty$, it follows from (46), (48) and (49) that

$$\hat{\theta}_j = \sum_{i=1}^{j}(x_i - x_{i-1} - u_{i-1})\Big/ j, \tag{50}$$

$$\hat{Q}_j = 1/j. \tag{51}$$

It can be shown in this case that

$$u^*_{N-j} = -\left[(F_{2j-1} - 1)\hat{\theta}_{N-j} + F_{2j}x_{N-j}\right]\Big/ F_{2j+1}, \quad j=1(1)N, \tag{52}$$

where F_j is the jth Fibonnaci number.

6 Conclusion

The authors hope that this work will stimulate further investigation using the approach on specific applications to see whether obtained results with it are feasible for realistic applications.

7 Acknowledgements

This work was supported in part by Grant GR-96.0213 from the Latvian Council of Sciences and the National Institute of Mathematics and Informatics of Latvia.

References

Aoki, M. (1967). *Optimization of Stochastic Systems.* New York: Academic Press.

Åström, K.J. (1970). *Introduction to Stochastic Control Theory.* New York: Academic Press.

Åström, K.J. and P. Eykhoff, P. (1971). System identification: a survey. *Automatica,* 7, pp. 123-164.

Bertsekas, D.P. (1976). *Dynamic Programming and Stochastic Control.* New York: Academic Press.

Nechval, N.A. (1984). *Theory and Methods of Adaptive Control of Stochastic Processes.* Riga: RCAEI.

Nechval, N.A. (1988). A new efficient approach to constructing the minimum risk estimators of state of stochastic systems from the statistical data of small samples. *Preprint of the 8th IFAC Symposium on Identification and System Parameter Estimation,* pp. 71-76, Beijing, P.R. China.

Nechval, N.A., Nechval, K.N. and Heimanis, B.M. (1997). Optimal rescheduling of airplanes under uncertainty. *Proceedings of the Nordic Operations Research Conference (NOAS '97),* pp. 213-224, Denmark, University of Copenhagen.

Yakowitz, S.J. (1969). *Mathematics of Adaptive Control Processes.* New York: American Elsevier.