Global Journal of Science Frontier Research: E Interdiciplinary
Volume 14 Issue 2 Version 1.0 Year 2014
Type : Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Neuro-Dynamic Optimization of Biotechnological Process

By Tatiana Ilkova & Mitko Petrov

Institute of Biophysics and Biomedical Engineering, Bulgaria

Abstract- In this paper we have investigated the whey fermentation at the production of white brine cheese with strain Kluyveromyces marxianus lactis MC 5, using non-conventional substrate whence for receiving of unicellular protein by an ecological clean and wasteless technology. This process is used for dynamic optimisation Neuro-Dynamic Programming theory. With this optimization procedure the quantity product is increased at the end of the process, simultaneously fermentation time is decreased. The producing of lactose with using of cheese whey, which is a waste product at the production of white brine cheese leads to the receiving one of close cycle and an ecological clean and wasteless technology.

Keywords: *neuro-dynamic programming, whey fermentation.*

GJSFR-E Classification : *FOR Code : 27089*

Strictly as per the compliance and regulations of :

Neuro-Dynamic Optimization of Biotechnological Process

Tatiana Ilkova [α] & Mitko Petrov [σ]

Abstract- In this paper we have investigated the whey fermentation at the production of white brine cheese with strain *Kluyveromyces marxianus lactis MC 5*, using non-conventional substrate whence for receiving of unicellular protein by an ecological clean and wasteless technology. This process is used for dynamic optimisation Neuro-Dynamic Programming theory. With this optimization procedure the quantity product is increased at the end of the process, simultaneously fermentation time is decreased. The producing of lactose with using of cheese whey, which is a waste product at the production of white brine cheese leads to the receiving one of close cycle and an ecological clean and wasteless technology.

Keywords: neuro-dynamic programming, whey fermentation.

I. Introduction

Dynamic Programming (DP) is an method for optimization and optimal control that has applied to surmount these restriction. The irnteest in this approach is due to the fact that too many hypotheses are based on the analytical structure of the equations, system and criteria. Therefore, it is possible to develop algorithms for solution of the optimal control problems independently of the used model and optimization criterion. DP should be applied parameterize the optimal decision as a function of the system state. Unfortunately, from the very beginning it was apparent that an increase of the dimensionality of the problem, i.e. an addition of reservoirs, caused an exponential increase in the time required to find a decision. This conducts to the "curse of dimensionality" (Bertsekas & Tsitsiklis, 1996).

Neuro-dynamic programming (NDP) is suggested as an alternative to lighten the *"curse of dimensionality" of the DP. NDP is a contemporary approach of the dynamic programming methods for optimization and optimal control and decision making under uncertainty. NDP combines ideas from the scopes of neural networks, artificial intelligence, reinforcement learning cognitive science, simulation, and approximation theory (Driessens &. Dzeroski, 2004).*

Using common artificial intelligence terms, the methods allow the systems to "learn how to make good decisions by observing their own behavior and use built-in mechanisms for improving their actions through a reinforcement mechanism". In more mathematical meaning "observing their own behavior" relates to simulation and "improving their actions through a reinforcement mechanism" relates to the iterative schemes for improving the quality of approximation of the optimal cost function, the Q-factors or the optimal policy. There has been a gradual realization that the reinforcement learning techniques can be fruitfully motivated and interpreted in terms of classical DP concepts such as the value and policy iteration, (Sutton, 1988).

The optimal control of this process usually depends on the presence of complex, non-linear dynamic model of the system because of this is difficult to realize to working-out of the problem, which is very important to practical realize.

The general disadvantage in optimization and optimal control methods of the bioprocesses is that the fermentation finishing time should be fixed. When we want to find the optimal fermentation finishing time, many various fermentation finishing times should be assumed, for each one of them, an efficiency optimal control problem should be determined. This is an extremely demanding task in terms of calculations. Approach such as control parameterisation can be used for free-end-time tasks to obtain open loop optimal profiles. Another hindrance of these approach checks from the circumstance that they are open loop optimal, which signifies that each time an initial condition shifts, a new and different optimal control problem should be determined. Moreover, the resulting fixed strategies do not take into advantage the likely process disturbances (Vlachos et. all 2006), (Lewis & Derong, 2013).

The fermentation of lactose oxidation from natural substratum in fermentation of Kluyveromyces marxianus lactis MC 5 uses non-conventional whence for receiving of unicellular protein. This process is not well studied. Therefore it does not exist a general mathematical model of the microbial synthesis, because of the extreme complexity and great variety of living activity of the microorganisms, although various models of biotechnological process and of different parts of the whey fermentation exist. On this cheese whey, which is waste product at the production of white brine cheese, oneself with what oneself reception one close cycle (Nono et. all 2006).

The aim of this study is to develop optimal feed rate strategy of biotechnological process in whey fermentation using Neuro-dynamic control.

Author α σ: "Bulgarian Academy of Sciences.
e-mails: tanja@biomed.bas.bg, mpetrov@biomed.bas.bg

II. Mathematical Model of the Process

Six fermentations were carried out in aerobic batch cultivation of Kluyveromyces lactis. A laboratory bioreactor ABR 02M with capacity 2 l has been used. The strain Kluyveromyces marxianus lactis MC5 is cultivated under the following conditions (Anastassiadis, 2007):

1. Nutrient medium with basic component – whey ultra filtrate with lactose concentration 44 g/l. The ultra filtrate is derived from whey separated in production of white cheese and deproteinisation by ultra filtration on LAB 38 DDS with membrane of the type GR 61 PP under the following condition:

- T=40-43 °C;
- input pressure 0.65 MPa;
- out put pressure 0.60 MPa.

The ultra filtrate is used in native condition with lactose concentration 44 g/l. Nutrient medium consist of:

- (NH) 4HPO 0.6%;
- yeast's autolisate 5%;
- yeast's extract 1%;
- pH 5.0 - 5.2.

2. The air flow rate QG is 60 l/l/h up to the 4th hour and 120 l/l/h up to the end of the process under continuous mixing n=800 min-1.
3. Temperature is 29 °C.
4. Duration of the cultivation is tf=12 hours.

The following figure describes the mass balance of the pressed-dough cheese and whey drink process.

The changes of the microbiological process (lactose conversion in yeast's cells to protein) are studied during the strain growth:

a) lactose concentration in fermentation medium in oxidation and assimilation of lactose by Kluyweromyces Marxianus lactis MC5 is determined by enzyme methods by UV tests (Boehringer Manheim, Germany, 1983);
b) concentration of cell mass and protein contents are determined on the basis of the nitrogenium contents (Kjeltek system 1028);
c) concentration of the dissolved oxygen in the fermentation medium in the process of oxidation and assimilation of lactose is determined by oxygen sensor.
d) For the measurement of the oxygen concentration in the fermentation middle the oxygen sensor that is produced by LKB firm, is used.

Six fermentations where carried out in aerobic batch cultivation of Kluyveromyces lactis. The experimental investigations are carried out on the computer controlled laboratory bioreactor 2L-M with magnetic coupling.

The model of the batch processes includes the dependences between the concentrations of the basic variables of the process: cell mass concentration, substrate concentration and oxygen concentration in the liquid phase:

$$\frac{dX}{dt} = \mu(S, C_L)\, X$$
$$\frac{dS}{dt} = -Y_1\, \mu(S, C_L)\, X \qquad (1)$$
$$\frac{dC_L}{dt} = \frac{k_l a}{1-\varepsilon_G}\left(C^* - C_L\right) - Y_2\, \mu(S, C_L)\, X$$

Where:

$$\mu(S, C_L) = \mu_m \frac{S^2}{\left(k_S + S^2\right)} \frac{C_L}{\left(k_C + C_L + C_L^2 / k_i\right)}$$

$\mu(S,C_L)$–specific grown rate of the cells, h^{-1}; μ_m - maximal rate of the cells, h^{-1}; X–cell mass concentration, g/l; S–concentration of substrate, g/l; C*–mean oxygen concentration, g/l; CL–dissolved oxygen concentration in liquid phase, g/l; kla–volumetric mass transfer coefficient, h^{-1}; ε_G–gas hold up; k_S, k_C, k_i, Y_1 and Y_2–coefficients; t–time, h.

The initial conditions and coefficients in model are given as follows:

X (0) = X0=0.2 g/l; S (0) = S0=44 g/l; CL (0) = C0=6.0x10-3 g/l; C*=C0, g/l;
μ_m =0.89; k_S=1.62; kC=3.37x10-3; kI=0.47; Y_1=2.238 and Y2=3.24x10-3.

The mass transfer coefficient kla and gas hold-up are determined by:

$$k_l a = 52\left(P/V\right)^{0,38} W_G^{0,23},\ \varepsilon_G = 0.53\left(Q_G/(nd^3)\right)^{-0.014} \qquad (2)$$

where: P–power input, $P = 60.9\rho n^3 d^5 \mathrm{Re}^{-0.4}$, W; V–volume, m3; ρ–liquid density, kg/m3; n–agitation speed, s-1; d–impeller diameter, m; Re–Reynolds number; WG–gas velocity, $W_G = 4\,Q_G/\pi\, D^2$, m/s; QG–air flow rate, m3/s; D–bioreactor diameter, m.

The model of the process (1)-(2) is used for optimization of batch fermentation of lactose oxidation from natural substratum in fermentation of the strain Kluyveromyces marxianus lactis MC 5.

III. Optimization of Fermentation Process

For determination of the optimal control problem of fed-batch fermentation processes maximizing of the optimization criterion at the end of the process max J on

the used substrate S is accepted. Thus the optimization problem is reduced to find a profile of the control variable.

The optimization criterion is accepted the value of the functional X (t) at the end the process (T=12) that means of the quantity formed biomass after 12 hour fermentation.

The criterion of quality has a type:

$$\max_{\mathbf{u}} Q = \int_{t_0}^{t_f} X(t)V(t)\,dt$$

where: t_0 – initial time, t_f – final time of the fermentation. The objective of this work is to find the optimal feed flow rate (F(t)) of a fed-batch process, such as the biomass, that will raise the biomass at the end of the process, i.e.:

$$\max_{\mathbf{u}} Q = \int_{t_0}^{t_f} X(t)V(t)\,dt \qquad (3)$$

A general dynamic optimization problem can be defined as follows:

$$\max_{\mathbf{u}_0,\ldots,\mathbf{u}_{k-1}} \sum_{i=1}^{k-1} f\,(\mathbf{W}_i, \mathbf{u}_i) \qquad (4)$$

Where: W is a vector of the variables that describe process, u - vector of control variables, k is the current stage.

The objective is to maximize the combination of the total span and the stagewise, together with the terminal costs subject and the terminal constrains.

DP includes a stagewise calculation of the cost-to-go function to reach the solution for the general initial state. The cost-to-go (10) at each stage is defined by (*Anastassiadis,* 2007):

$$B_i\left(W(t_i), t_i\right) = \max_{u_{\min} \le u_k \le u_{\max}} \Delta t \sum_{k=1}^{N-1} f_k\left(\mathbf{W}_k, \mathbf{u}_k\right) \qquad (5)$$

Then the calculation of the cost-to-go function at each stage can be done as:

$$B_i\left(W(t_i), t_i\right) = \max_{\mathbf{u}_{\min} \le \mathbf{u}_k \le \mathbf{u}_{\max}} \left\{ f_i\left(W(t_i), \mathbf{u}_i\right) + B\left(W(t_{i+1}), t_{i+1}\right) \right\} \qquad (6)$$

Once obtained the cost-to-go function, represents a convenient vehicle to obtain the optimal solution for the general stage.

By continuing the cost-to-go iteration of (6) until convergence within the procedure it can be seen that the infinite horizon cost-to-go function B_∞, satisfying the following "Bellman equation" can be obtained:

$$B_\infty(W) = \max_{\mathbf{u}} \left\{ f(W, \mathbf{u}) + B(W, \mathbf{u}) \right\} \qquad (7)$$

Unfortunately, in very few cases the problem can be solved through the stagewise optimization in order to analytically obtain a closed-form expression for the cost-to-go problem. The conventional numerical approach to the problem involves gridding the state space, calculating and storing the cost-to-go for each grid points as one marches backward from the first (or last) stage to the lest (first). For an invite horizon problem the number of iterations required for convergence can be very big. Such an approach is seldom practically feasible due to the exponential growth of the computation with respect to the state dimension.

The traditional approach for solving the Bellman equation involves gridding of the state space, solving the optimization (10) for each grid point and performing the stagewise optimization until convergence is achieved. The comprehensive sampling of the state space can be avoided by identifying the relevant regions of the state space by simulation under judiciously chosen suboptimal policies (Vlachos et. al. 2006).

The policy improvement theorem states that a new policy that is greedy (a greedy policy is one whose current cost is the least) with respect to the cost-to-go function of the original policy is as good as or better than the original policy, so the new policy can be defined as follows:

$$\mathbf{u}(W) = \arg\max_{\mathbf{u}} f\ (\mathbf{W}, \mathbf{u}) + B(\mathbf{W}, \mathbf{u})$$

Where $\arg\ G(u, x, i) \in R_{m+n+r}$ is an improvement over the original policy and $u \in R_m$, $W \in R_n$ and $i \in R_r$.

When the new policy is as good as the original policy the above equation becomes the same as Bellman equation (7).

The relevant regions of the state space are identified by simulation of NDP control and the initial suboptimal cost-to-go function is calculated from the simulation data. In this survey a functional approximator is used to interpolate between this data. The improvement is obtained through the iteration of the Bellman equation. When the iteration Converge this off-line computed cost-to-go function can be used for an on-line optimal control calculation for the bioreactor (Xiong & Zhang, 2005).

NDP uses neural network approximations for the approximation of cost-to-go function. The cost-to-go function was not used to generate an explicit control law; instead, it was used in an on-line optimization to reduce the large (or infinite) horizon problem to a relatively short horizon problem. The method was found to be robust to approximation errors. Both deterministic (step changes in kinetic parameters) and stochastic problems (random variations in kinetic parameters and feed composition) were explored (Peroni et. all 2009), (Krishnamoorthy K. et al 2011).

The following notations are used for description of the algorithm:

B – Bellman equation;

$\widetilde{B}(x)$ - approximated Bellman equation corresponding to state W;

$()^i$ – iteration index for cost iteration loop;
k – discrete time.
Finally:

$$\widetilde{B}(k) \equiv \widetilde{B}(W(k)) \text{ and } f(k) = f\big(W(k), \mathbf{u}(k)\big)$$

The general simulation-approximation scheme involves computation of the converged cost-to-go approximation off-line. The architecture of the scheme is shown in Figure 2. Step 1, Step 2, Step 3 and Step 4 represent the "Simulation part", and 5 and 6 the "Cost Approximation Part".

The simulation-based approach involves computation of the converged profit-to-go approximation off-line. The following steps describe the general procedure of NDP algorithm:

1. Starting with a given policy (some rule for choosing a decision u at each possible state i), and approximately evaluate the cost of that policy (as a function of the current state) by least-squares-fitting a scoring function $\widetilde{J}^j(X)$ to the results of many simulated system trajectories using that policy;
2. The solution of one-stage-ahead cost plus cost-to-go problem, results in improvements of the cost values;
3. The resulting deviation from optimality depends on a variety of factors, principal among which is the ability of the architecture $\widetilde{J}^j(X)$ to approximate accurately the cost functions of various policies;
4. Cost-to-go function is calculated using the simulation data for each state visited during the simulation, as for each closed loop simulation (simulation part).
5. A new policy is then defined by minimization of Bellman's equation, where the optimal cost is replaced by the calculated scoring function, and the process repeats. This type of algorithm typically generates a sequence of policies that eventually oscillates in a neighbourhood of an optimal policy;
6. Fit a neural network function approximator to the data to approximate cost-to-go function as a smooth function of the states;
7. The improved costs are again fitted to a neural network, as described above, to obtain subsequent iterations $\widetilde{J}^1(X), \widetilde{J}^2(X)$, and so on …, until convergence.

The NDP algorithm block- scheme is shown in Figure 1.

Take into consideration that when starting with a fairly good approximation of the cost-to-go (which has to be a result of using a good suboptimal policy), the cost iteration has to converge fairly fast – faster than the conventional stagewise cost-to-go calculation.

The next values of F are examined - *F=0.1, 0.2, …, 0.7*, that can cover the possible rang of variations.

The bioreactor was started at three different W (0) values for each of the parameter values around the low product yield steady state.

A functional approximation relating cost-to-go with augmented state was obtained by the neural network – with five hidden nodes, six input nodes and two output nodes. The neural network presented a good fit with a mean error of 10-3 after training for 1000 epoch.

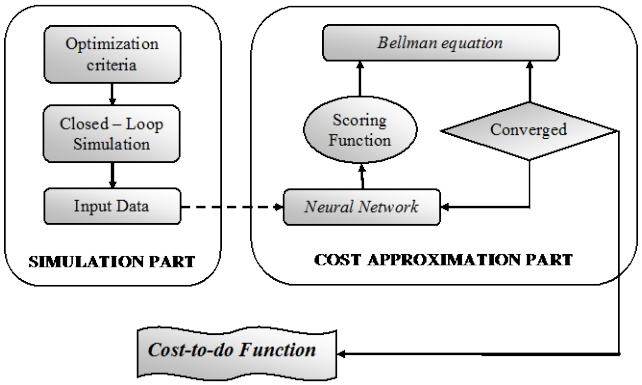

Figure1: NDP algorithm.

Improvement of the cost-to-go is obtained through the iterations of the Bellman equation (13). This method is known as a value iteration (or value iteration). The solution of the one-stage-ahead cost plus cost-to-go problem, results in the improvement of the cost values. The improved prices were again fitted to the neural network, described above to obtain subsequent iterations $\widetilde{B}^1(k)$, $\widetilde{B}^2(k)$ and so on ..., until they are converged. Cost is said to be "converged" if the sum of the absolute error is less than 5% of the maximum cost. The cost is converged in 7 iterations for our system.
The converged cost-to-go function from above was used for solving the one-stage-ahead problem. The choice for switch over the one-stage-ahead of the control variable is calculated by:

$$\mathbf{u}(k) = \arg\max_{\mathbf{u}(k)} \left\{ f\left(\frac{Q(t_k)}{t_k}, \mathbf{u}\right) + \widetilde{B}^6\left(\frac{Q(t_k)}{t_k}, \mathbf{u}(k)\right) \right\} \quad (8)$$

Where: u is vector of control variables, k is the optimization stages, B is Bellman equation.

Following this procedure, a program on MATLAB 8.0 has been developed and the optimal profile of the control variable has been obtained.

With this procedure a program on MATLAB 8.0 is developed and it the optimal profile of the optimal control variables are obtained. The time procedure is in 1100 s. The profile is shown on Figure 2.

The concentration of the biomass before and after optimization is shown in Figure 3.

The program for optimal control is developed using NDP method. The results show generally an

increase of 13.37 % amount of biomass production after the dynamic optimization. This result is shown in Figure 9.

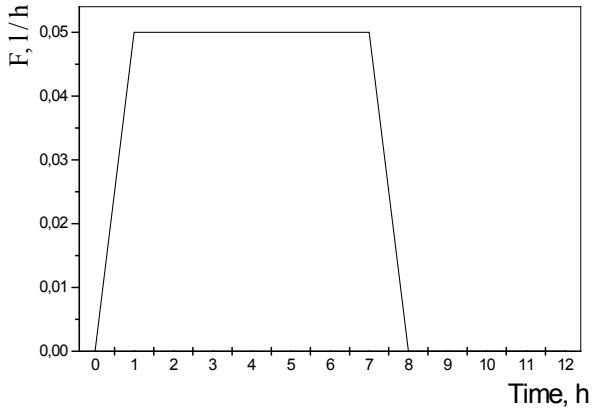

Figure 2 : Optimal profile of the feeding rate received with and NDP.

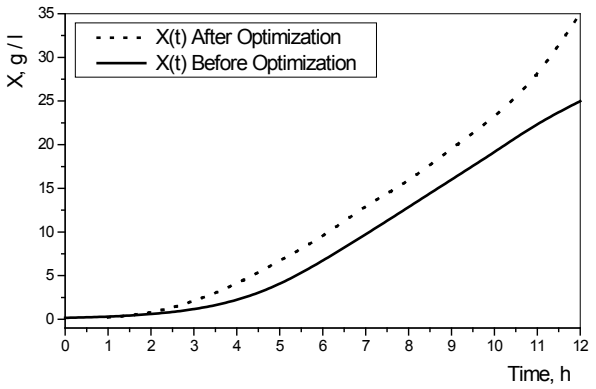

Figure 3 : Concentration of the biomass before and after optimization

IV. Conclusion

A technique based on Neuro-Dynamic Programming approach has been developed to achieve an optimal feed rate profile for biomass production a fed-batch bioprocess.

The approach proposed such as one methodic for alleviation of "curse of dimensionally" of dynamic programming. The results show that quality biomass is risen at the end of the process. Using of the method shows that it to be able for application in on-line optimal problems.

Realization of such an optimal control approach combined with advanced control techniques (artificial neural networks with classical optimization method) in practice can conduct to the value and elaboration time reduction in the laboratory fed-batch bioprocesses, but not only that, as well in elaboration time reduction technique for optimal control.

The producing of lactose with using of cheese whey, which is a waste product at the production of white brine cheese leads to the receiving one of close cycle and an ecological clean and wasteless technology.

V. Acknowledgement

This work was published with financial support of the "Human Resources Development" Operational Programme, "Science–Business" Scheme Grant No.BG051PO001-3. 3 - 05/0001.

References Références Referencias

1. Bertsekas D., &. Tsitsiklis. (1996)Neuro-Dynamic Programming. Massachusetts: Athena Scientific Belmont.
2. Driessens K., & S. Dzeroski. (2004) Integrating Guidance intro Relational Reinforcement Learning. Machine Learning, 57, 217-304.
3. Sutton, R. S. (1998) Learning to Predict by the Methods of Temporal Differences, Machine Learning, 3, 9-44.
4. Peroni C. V., & N.S. Kaisare & J.H. Lee. (2009) Optimal control of a fed-batch bioreactor using simulation-based approximate dynamic programming, IEEE Transactions on Control Systems Technology, 13, 786-790.
5. Ilkova T. & M. Petrov. (2011) L-lysine Neuro-dynamic Control. Journal of Medical Research, 11, 55-60.
6. Nono Y. J., & D. G. Libouga, & D. Ngongang, & J. P. Ramet & M. Parmentier. (2006) Feasibility of Cheese Production and whey Valorization in the Adamawa province of Cameroon. African Journal of Biotechnology, 5 (6), 517-522.
7. Anastassiadis S. (2007) L-Lysine fermentation. Recent Patents on Biotechnology, 1, 11-24.
8. Vlachos D. G., & A. B. Mhadeshwar & N.S. Kaisare. (2006) Hierarchical Multiscale Model-based Design of Experiments, Catalysts and Reactors for fuel Processing. Comput. Chem. Eng, 30, 1712-1724.
9. Krishnamoorthy K., & M. Pachter, & S. Darbha, & P. Chandler (2011) Approximate Dynamic Programming with State Aggregation Applied to UAV Perimeter patrol. Int J of Robust and Nonlinear Control, 21, 1396-1409.
10. Lewis F. L., & L. Derong L. (2013) Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, Wiley-IEEE Press.

This page is intentionally left blank