# Computer aided problem solving via optimization[1]

Ferenc Szigeti*, Gisela De Sarrazin+ and Juan Cardillo*

* Department of Control, Faculty of Engineering
"Los Andes " University. Mérida 5101 - Venezuela
e-mail: szigeti@ing.ula.ve, ijuan@ing.ula.ve

\+ Department of Mathematics, Faculty of Science
"Los Andes " University. Mérida 5101 - Venezuela
e-mail: gcovi@ciens.ula.ve

**Abstract** Assuming a finite set of tools with a rule based implementation, we define a language. The active tools are defined at each word of language. Suppose that a valuation function is given on our language, which indicates whether a sequence of tools (word) solves a certain task. Then we construct a minimal realization, a graded state space representation of the valuation function. Finally, by calculating an "optimal control" problem for a multi-(two) objective function, we can compute that after making a "move" (applying a new active tool) the solution of our task is approaching, moving away, or that the step does not affect our ability to solve our task. Hence, after the implementation of our algorithm, the problem solving is supervised by an intelligent computer, capable of the evaluation and decision making.

**Keywords.** Language, Discrete Events Dynamics System, Minimum principle, Multi-objetive function, Intelligent dialogue.

---

[1]The final version of the present work was finished during the stage of the second author at RISC Institute, thanks to the invitation of Prof. Bruno Buchberger

# 1 Introduction

In this paper we will pose and solve the following problem. Suppose that a certain task, such as the manufacture of goods or an exercise in a problems book can be solved, using a well defined set of tools. The correct solution of that task can be archieved using the given tools sequentially.

We propose medelling the solution of the problem using Discrete Event Dynamic Systems (DEDS).

The tracking and evaluation of the partial results are achieved by an optimal control problem for the DEDS which is associated with the task. This optimal DEDS can be computed by using a minimum principle, similar to Pontryagain's minimum principle.

The "coast" function $\Phi_2$ punishes if the value is large, for example it could be the number of tools required to solve the problem. Hence the objective function is $\Phi = (\Phi_1, \Phi_2)$.

We suppose that the output space $Z_2$ x $Z_+$ is equipped with the lexicographic ordering. Hence, the minimum of $\Phi$ over DEDS process, associated to the task is achieved if the value of $\Phi$ over the final state is $(O, T*)$, where $T*$ is the minimal length of the sequence of tools which solves the task from a given position. If the task does not have solution, then, the values of the objective function $\Phi$ over all processes is of the form $(1, T)$ and the approach of optimal control can not solve the problem.

The communication between the user who ask for help from the computer and the computer is based on the following 3 possibilities: if after a "move" the value of the $\Phi$ is less, then this step is approaching the solution of the task. If it is moving away, then the value of the objective function grows, and if it is indifferent, then the objective function is stationary.

# 2 Languages and class of Discrete Event Dynamical Systems

Let $\Omega = \{\alpha, \beta, \gamma\}$ be a finite set of tools (we will use as synonyms the words: tool, letter and event). The finite sequences (strings) of letters are called words. The empty word, denoted by $\Theta$, does not contain any letter.

The operation, of concatenation between word is defined in the following way: if $u = \alpha\beta\ldots\phi$ v $= \gamma\eta\ldots\kappa$ are any words made up of letters in $\Omega$, the concatenation $u\mathrm{v}$ is defined by $u\mathrm{v} = \alpha\beta \ldots \phi\gamma\eta \ldots \kappa$.

Obviously, $(u\mathrm{v})\mathrm{w} = u(\mathrm{vw})$, $\quad \Theta u = u\Theta = u$.

Then, the set $\Omega^*$, consisting of words of letters in $\Omega$, is a monoide with respect to the concatenation with the neutral element $\Theta$. A subset of $\Omega^*$ is called a language.

The language $L \subset \Omega^*$ is closed if for each word $\omega \in L, \omega = \omega_1\omega_2 \ldots \omega_k$, the words, called prefixes:

$$\mathrm{w}_0 = \Theta, \mathrm{w}_1 = \omega_1, \ldots, \mathrm{w}_i = \omega_1\omega_2 \ldots \omega_i, \ldots, \mathrm{w}_k = \omega_1\omega_2 \ldots \omega_k$$

belongs to the language $L \subset \Omega^*$.

The letter $\alpha \in \Omega$ is acceptable or admissible for the word $u \in L$, ($u$-acceptable) if the concatenation $u\alpha \in L$.

Let $\Omega(u) \in \Omega$, denotes the subset of $u$-acceptable words.

The closed language induces a DEDS

$$f : L \text{ x } \Omega \mapsto \Omega \tag{1}$$

where the application $f$ is partially defined for the pairs $(u, \omega)$ with $\omega \in \Omega(u)$ and

$$f(u, \omega) = u\omega, \quad \omega \in \Omega(u).$$

The application $f$ defines the transition in the language $L \subset \Omega^*$ by using the keys or events $\omega$. More generally, if $X$ is a set, then a Dynamical System is partially defined application:

$$f : \mathrm{X x}\Omega \mapsto \ \mathrm{X}$$

We can now define for each point $x \in \ \mathrm{X}$ the $x$-acceptable words in the following way: $\omega$ is $x$-acceptable if $(x, \omega)$ belongs to the domain of the aplication $f$, hence $f(x, \omega)$ is well defined for the letters $\omega$ which are $x$-acceptable.

If we have another application, $g : \mathrm{X} \mapsto \ \mathrm{Y}$ then the quintuple:

$$S = (\mathrm{X}, \Omega, \mathrm{Y}, f, g)$$

is a DEDS with output observation. If $x_0 \in \ \mathrm{X}$ is the initial point, then it is an initialized system $S_{x_o}$,

$$S_{x_0} = (\mathrm{X}, x_0, \Omega, \mathrm{Y}, f, g)$$

if $L \subset \Omega^*$ is a closed set and $g : L \mapsto$ Y is an app lication then

$$S_\Theta = (L, \Theta, \Omega, \mathrm{Y}, f, g) \tag{2}$$

is an initialized dynamical system with observation, where $f$ is the application in 1.

For $\xi \in$ X, a word $u \in \Omega^*$, $u = \omega_0, \omega_1 \ldots \omega_{k-1}$ is $\xi$-admissible if

$$\omega_0 \in \Omega(\xi), x_0 = \xi, x_1 = f(x_0, \omega_0),$$

and if $x_0, x_1, \ldots, x_i$ are defined in such a way that $\omega_j \in \Omega(x_j), j \leq i < k-1$, then, for $\omega_{i+1}$ we have:

$$x_{i+1} = f(x_i, \omega_i), \omega_{i+1} \in \Omega(x_{i+1}) \tag{3}$$

We Denote the set of $\xi$-admissible words by $A(\xi)$.

If $L \subset \Omega^*$ is a closed language, then all the words $u \in L$ are $\Theta$-admissible. Moreover, if $u \in L, u = \omega_0\omega_1 \ldots \omega_{k-1}$, and

$$x_{i+1} = f(x_i, \omega_i), x_0 = \Theta,$$

Then $x_k = u$ where $f$ is the application in 1.

The global dynamics

$$F : \mathrm{X}\mathrm{x}\Omega^* \mapsto \mathrm{X},$$

induced by the $f$-dynamics is defined for all of the pairs $(\xi, u) \in \mathrm{X}\mathrm{x}\Omega^*$, such that $u \in A(\xi)$, through the equation

$$F(\xi, u) = x_k,$$

where $u = \omega_0\omega_1 \ldots \omega_{k-1}$, and $x_0, x_1 \ldots, x_k$ are the solutions of (3).

Let

$$S_{x_0} = (\mathrm{X}, x_0, \Omega, Y, f, g)$$

be a DEDS with observation. We say that the states $x, y \in$ X are indistinguishable if:

$$A(x) = A(y), g(F(x, u)) = g(F(y, u)),$$

for all admissible words $u \in A(x) = A(y)$. The concept of undistinguishability is an equivalence relation, and the equivalence classes are the new state.

The equivalence class of $x$ is $[x]$, the initial point of the new system is $[x_0]$, and the new state-space is $[\mathrm{X}] = \{[x]/x \in \mathrm{X}\}$.
The new dynamics is defined by

$$f_{\min}([x], u) = [f(x, u)], u \in A(x)$$

and the new observation is:

$$g_{\min}([x]) = g(x).$$

Obviously, $A([x]) = A(x)$. In this way we have constructed an initialized DEDS with observation:

$$[S]_{[x_0]} = ([\mathrm{X}], [x_0], \Omega, Y, f_{\min}, g_{\min})$$

Note that the alphabet $\Omega$, and the set $Y$ of outputs are the same for both systems $S_{x_0}$ and $[S]_{[x_0]}$. Moreover, for each $x \in \mathrm{X}, A(x) = A([x]) \in \Omega$. The induced input and output functions $P : A(x_0) \mapsto Y, P_{\min} : A([x_0]) \mapsto Y$ satisfy the aqualities:

$$P(u) = g(F(x_0, u)) = g_{\min}(F_{\min}([x_0], u)) = P_{\min}(u)$$

These equations show that the two dynamical systems realize the same input-output function.

The DEDS $[S]_{[x_0]}$ has the following notable properties:

1. If $[x], [y] \in [\mathrm{X}], [x] \neq [y]$, then $A([x]) \neq A([y])$, or, if$A([x]) = A([y])$, then there exists $u \in A([x]) = A([y])$, such that:

   $$g_{\min}(F_{\min}([x], u)) \neq g_{\min}(F_{\min}([y], u))$$

   In this case the system is observable, hence, each pair of different elements is distinguishable.

2. If the system $S_{x_0}$ is reachable; in other words, if for each element $x \in \mathrm{X}$ there exists an element $u \in A(x_0)$, such that $x = F(x_0, u)$, then the system $[S]_{[x_0]}$ has the same property.

3. Let $\mathrm{X}_0 = \{[x] : [x] = F([x_0], u); u \in A([x_0])\}$. Then the system restricted to the state space $\mathrm{X}_0$ is the minimal realization to the input output function $P : A(x_0) \mapsto Y$

4. Consider the system (2) for a closed language, then, the corresponding realization $[S]_{[x_0]}$ is a minimal realization of the input output function $g : L \mapsto Y$.

## 3 The graded realization

Let $L \subset \Omega^*$ be a closed language. We will realize the input output function $g : L \mapsto Y$ over the graded state space. The length of a word $u \in \Omega^*$ is the numbers of letters it contains and it is denoted by $< u >$. Suppose $n$ is a non negative integer number, then the $n-$state space $\mathrm{X}_n$ is defined by

$$\mathrm{X}_n = \{[u] : u \in L, < u >= n\},$$

where the indistinguishability of two words is defined by the Nerode's equivalence:

$$\begin{array}{rl} u \approx v, & \Longleftrightarrow A(u) = A(v), g(uw) = g(vw) \\ & \forall w \in A(u) = A(v) \end{array}$$

Therefore the dynamics of the dynamical system over the graded state-space is defined through the partial applications

$$f_n : \mathrm{X}_n \mathrm{x} \Omega \mapsto \mathrm{X}_{n+1},$$

over the set (domain) of all pairs $(x, \omega) \in \mathrm{X}_n \mathrm{x} \Omega$, that satisfy the relation $\omega \in \Omega(x)$, through the equality:

$$f_n(x, \omega) = f_n([u], \omega) = [u\omega].$$

$f_n$ is well defined since $< u\omega >=< u > +1 = n+1$, then $[u\omega] \in \mathrm{X}_{n+1}$.The observation function

$$g_n : \mathrm{X}_n \mapsto Y$$

is defined by the equation

$$g_n(x) = g_n([u]) = g(u).$$

## 4 Optimal control for DEDS

Consider the initialized dynamical system with observation over the graded state-space

$$f_n : \mathrm{X}_n \mathrm{x} \Omega \mapsto \mathrm{X}_{n+1}, g_n : \mathrm{X}_n \mapsto Y \tag{4}$$

Our aim here is to consider the optimization problem

$$\begin{array}{rll} x_{n+1} = & f_n(x_n, \omega_n), & x_0 = \xi \\ & \min \longrightarrow g_N(x_N) & \end{array}$$

where we take the minimum over the set $A(\xi) \subset \Omega^*, \xi \in \mathrm{X}_0$. We suppose that $u = \omega_0\omega_1 \ldots \omega_{N-1}$. Note that if the set of outputs is finite or numerable, then we can suppose that it is immersed in the set of integers $Z$, or in the product $Z^k$. If the state-space are finite, we may suppose that they are subsets of $Z$. The functions (4) can be interpolated at all points of the domain, so that we can assume that the functions are polynomials.

The optimization problem given, is the optimal control problem for the DEDS, see [1],[2]. Comparing this with the classical minimal principle, we see that if we substitute for the gradient the formal algebraic Taylor residue, then the minimal principle can be formulated in the same manner.
If $m(x,u) = x^k u^l$ is a monomial with variables $x, u$ then the Taylor residue is defined by:

$$m(y,v) - m(x,u) = Rm(x,u;y,v)(y-x, v-u) =$$
$$u^l \left[\sum_{i=0}^{k-1} x^i y^{k-i-1}\right](y-x) + y^k \left[\sum_{i=0}^{l-1} u^i v^{l-i-1}\right](v-u),$$

and the Taylor residue of the monomial is the vector

$$Rm(x,u;y,v) = \left\langle u^l \left[\sum_{i=0}^{k-1} x^i y^{k-i-1}\right], y^k \left[\sum_{i=0}^{l-1} u^i v^{l-i-1}\right]. \right\rangle$$

The multivariable case is similar.

In the same way we can define the partial Taylor residue and also the residue of a vectorial function whose residue is analogous to the Jacobian matrix. The adjoint equation is defined with respect to two processes $(x^*, u^*), (x, u)$. Suppose that the process length is the same, $< u^* >=< u >$. Then the adjoint equations definees the dual or adjoint trayectory by

$$p_n = R_x f_n(x_n^*, \omega_n; x_n)^T p_{n+1}, p_N = Rg_N(x_N^*; x_N)$$

where $u^* = \omega_0^*\omega_1^* \ldots \omega_{N-1}^*, u = \omega_0\omega_1 \ldots \omega_{N-1}$. Then, in terms of adjoint trayectory we can prove the formula:

$$g_N(X_N) - g_N(x_N^*) =$$
$$= \sum_{n=0}^{N-1} f_n(x_n^*, \omega_n)^T p_{N+1} - \sum_{n=0}^{N-1} f_n(x_n^*, \omega_n^*)^T p_{N+1}.$$

We can also prove the Pontryagin minimum principle from this formula.

**Theorem.** Suppose that the process $x^*, u^*$ is the minimum of the optimization problem. If the process $(x_{n,v}, u_{n,v})$ corresponds to the control $u_{n,v} = \omega_0^* \dots \omega_{n-1}^* v \omega_{n+1}^* \dots \omega_{N-1}^*$, then for the solution $p^{n,v}$ of the adjoint equation, with respect to the processes $(x^* u^*), (x_{n,v} u_{n,v})$, we have the inequality, known as Pontryagin's minimum principle.

$$f_n(x_n^*)^T p_{n+1}^{n,v} \geq f_n(x_n^*, \omega_n^*)^T p_{n+1}^{n,v}.$$

# 5 Computer aided, supervised problem solving

In the previous sections we have formalized the use of certain mathematical tools, and we have obtained a DEDS over grade state-spaces. Now we introduce an objective function, or in other words, an output function. This is done by defining the function $s_n : \mathrm{X}_n \mapsto \{c, nc\}, n = 0, 1, \dots, N$, in such a way that if the state corresponds to a complete solution, then the value of the function $s$ is $c$, (complete), and if the solution is not complete, then its value is $nc$. We therefore have

$$g_n : \mathrm{X}_n \mapsto Z, g_n(x) = \min_{v \in A(x)} \{< v >: s(F_n(x, v)) = c\}$$

We identify $c$ with 0, and $nc$ with 1.

It follows from the definition of the output functions that the state $x \in \mathrm{X}_N$ is a complete solution if and only if $g_N(x_N) = 0$.

On the other hand the control values of the lengths of words, are nonnegative integers so that, the state

$$x^* \in \mathrm{X}_N, x^* = [u^*] = [\omega_0^* \omega_1^* \dots \omega_{N-1}^*],$$

corresponds to a complete solution if $(x^*, u^*)$ is the solution of the optimization problem

$$g_N(x_N) = 0 \tag{5}$$

Thus, we introduce the following algorithm compute (5):

1. Starting with the problem, and the indispensable mathematical tools to solve the problem (assuming at least one solution), we describe the admissible language for our tools.

2. We describe the minimal realization of the language. If the state-space is not too large, it is not necessary to construct the graded realization. The importance here is to reduce the complexity of the problem.

3. If the calculus of the output function defined above is too complicated or if it is not completed, then we define a vectorial output function or, in this case, an objective multicriteria function

$$g_n(x) = (s(x), < u >), x = [u].$$

To for compare two different values of the objective function we must define a complete ordering relation in $Z^2$. We may do this simply by taking the lexicographic order:

$$(k, l) < (m, n) \Longleftrightarrow \begin{cases} k < m, \\ k = m \ \ \& \ \ l < m \end{cases}$$

If the problem has a solution, then the minimum is reached over all solutions, where the function $g_n^2 : \mathrm{X}_n \mapsto Z^2$ takes the values $(0, < u >)$. The minimum is the solution with minimal length and thus the optimal control problem for multi-criteria also is able to find a solution for the problem.

Moreover, if we look for the optimal solution of the problem (15) with the objective function $g_n^2 : \mathrm{X}_n \mapsto Z^2$ then the value of the second coordinate, or the solution length, is the value $g_k(\xi)$ of the original objective function.

# References


1. J.J. Cardillo-Albarrán: (1994) Optimización de sitemas a eventos discretos usando el princ´ıpio de m´ınimo de Pontryagin, Thesis Msd, Fac. Engineering, ULA, Mérida.

2. J.J. Cardillo-Albarrán and F. Szigeti: (1995) A symbolic algorithm for the optimization of imteger processes over finite state space, accepted for presentation and publication in the International Congress on Industrial and Applied Mathematics, ICIAM 95, Hamburg, Germany, 3-7, July.

3. L.E. Solé and F. Szigeti: (1992) Reachability of a class of coupled discrete event systems, in the Proc. IFAC/IFORS Symposium on Large Scale Systems: Theory and Applications, Beijing, China, 386-389