New Trends in Applied Mathematics

Dedicated to Professor Cor Baayen

A. Bensoussan
University Paris-Dauphine and INRIA

Applied mathematics has become an extremely important and useful discipline in the context of development of powerful computers. On the one hand, mathematics (in a broad sense) is the most efficient approach to model reality, especially complex reality. Moreover, it provides the best possibilities of reasoning. With cheap powerful computers, mathematics becomes implementable and unavoidable in designing, producing, deciding ...

On the other hand, mathematics has evolved considerably to extend its applicability to real problems. This is why applied mathematics is so alive and fast progressing. Needless to say, the connection between applied mathematics and information technology is an extremely fruitful approach to new ideas and a basic source of research topics. This is a line to which Professor Cor Baayen has always dedicated his efforts. He has greatly contributed to closing the gap between mathematics and computer science. To give an exhaustive presentation of all directions of applied mathematics in a short talk is of course out of reach, and beyond the possibilities of one speaker. So the purpose of this lecture is more to outline some significant features, among many others.

1 Scientific computing

The traditional applications of mathematics arise in Physics, Mechanics, ... Powerful computing means and supercomputers have permitted :

- to study completely new areas of physical sciences.
- to consider new numerical techniques
- to investigate new approaches.

1.1 New Areas of physical sciences

It would be particularly unrealistic to be exhaustive here. Nevertheless, among important developments in several fields, we emphasize the *Numerical Simulation of Reactive flow.* It applies indeed to *combustion, aeronomy, partially*

ionized plasmas, aerodynamics, gas dynamic lasers, astrophysics, general multiphase and magneto-hydrodynamic flows, ...

The model takes into account the *coupling* between *fluid dynamics* and *chemical reactions*, and thus opens the door to a large family of complex problems.

The traditional model of an *homogeneous, viscous, incompressible flow* with *no chemical reactions* and *no external forces* consists of Navier Stokes equations :

$$\begin{aligned} \rho_0(\frac{\partial u}{\partial t} + (u \cdot D)u) - \mu\Delta u + Dp &= 0 \\ \text{div}u &= 0 \end{aligned}$$

If the fluid has a constant *specific heat* c and there are no external heat sources, then the temperature of the fluid is the solution of :

$$\rho_0 c(\frac{\partial T}{\partial t} + u \cdot DT) - \lambda\Delta T == 2tr\varepsilon^2(u)$$

where $\varepsilon = \frac{1}{2}(Du + (Du)^T)$ is the *velocity tensor.* The internal energy density is cT.

In general, all variables are coupled and appear as the solution of a complex system of P.D.E.

The main unknown are the mass density ρ, the velocity of the flow u, the number densities n^i of the individual chemical species and the total energy density E.

The system of equations is the following :

$$\begin{aligned} \frac{\partial \rho}{\partial t} + D \cdot (\rho u) &= 0 \\ \frac{\partial(\rho u)}{\partial t} + D \cdot (\rho u u) + D \cdot \sigma &= \sum_i \rho^i a^i \\ \frac{\partial n^i}{\partial t} + D \cdot (n^i(u + u^i)) &= Q^i - L^i n^i \\ \frac{\partial E}{\partial t} + D \cdot (Eu) + D \cdot (u \cdot \sigma) + D \cdot (q + q_r) &= \sum_i (u + u^i) \cdot m^i a^i \end{aligned}$$

where σ is the *pressure tensor*, q the *heat* flux, q_r the radiative heat flux, a^i represent external forces, and Q^i, L^i represent the chemical production rates and losses of species i, u^i is the diffusion velocity of species i. They are highly nonlinear expressions of the unknowns, including the temperature T.

In view of the complexity, a *modular* approach is useful. Each physical process is calculated accurately and calibrated separately .

The physical properties should be incorporated in the numerical algorithms and a mathematical analysis of the behaviour of the algorithms should be performed. For more details, see [18].

1.2 Numerical methods

We shall illustrate the general idea of *decoupling* the difficulties in the case of Navier Stokes equations:

$$\begin{aligned} &\frac{\partial u}{\partial t} + (u \cdot D)u - \mu\Delta u + Dp = f \\ &\text{div } u = 0 \\ &u(x,0) = u_o(x) \qquad\qquad \text{div } u_0 = 0 \\ &u = g \text{ on } \Gamma \qquad\qquad \int_\Gamma \nu \cdot g \, d\Gamma = 0 \end{aligned}$$

The two main difficulties are non linearities and incompressibility condition. *Operator splitting* will realize the decoupling.

Let θ be a parameter in $(0, \frac{1}{2})$ and α, β with $\alpha + \beta = 1$.

Knowing u^n, we compute $\{u^{n+\theta}, p^{n+\theta}\}, u^{n+1-\theta}$ and $\{u^{n+1}, p^{n+1}\}$ by the iteration :

$$\begin{aligned} &\frac{u^{n+\theta} - u^n}{\theta\Delta t} - \alpha\mu\Delta u^{n+\theta} + Dp^{n+\theta} = \beta\mu\Delta u^n \\ &-(u^n \cdot D)u^n + f^{n+\theta} \\ &\text{div } u^{n+\theta} = 0 \\ &u^{n+\theta} = g^{n+\theta} \qquad \text{on } \Gamma \end{aligned} \tag{1}$$

$$\begin{aligned} &\frac{u^{n+1-\theta} - u^{n+\theta}}{(1-2\theta)\Delta t} - \beta\mu\Delta u^{n+1-\theta} + (u^{n+1-\theta} \cdot D)u^{n+1-\theta} = \alpha\mu\Delta u^{n+\theta} \\ &-Dp^{n+\theta} + f^{n+1-\theta} \\ &u^{n+1-\theta} = g^{n+1-\theta} \qquad \text{on } \Gamma \end{aligned} \tag{2}$$

$$\begin{aligned} &\frac{u^{n+1} - u^{n+1-\theta}}{\theta\Delta t} - \alpha\mu\Delta u^{n+1} + Dp^{n+1} = \beta\mu\Delta u^{n+1-\theta} \\ &-(u^{n+1-\theta} \cdot D)u^{n+1-\theta} + f^{n+1} \\ &\text{div } u^{n+1} = 0 \\ &u^{n+1} = g^{n+1} \qquad \text{on } \Gamma \end{aligned} \tag{3}$$

(2) is nonlinear and solved by a *least square technique* , and *conjugate gradient* minimization. (1) and (3) are linear and can be reformulated as variational problems for the pressure p.

Various possibilities of finite element approximation, multigrid methods and domain decomposition can then be used at the discretization stage.

Efficient software packages result in the combination of all these techniques. For more details, see [10].

1.3 New approaches

We present two new directions :

1.3.1 Wavelets

An alternative to Fourier analysis has been developed in recent years, with applications to signal and image processing, sound analysis and numerical analysis. It has foundations in quantum field theory, statistical mechanics and pure mathematics (geometry of Banach spaces). This is the *Wavelet analysis.*

It combines advantages of the Haar system and of the trigonometrical system. The Haar system is defined by :

$$\psi(x) = \begin{cases} 1, & 0 \leq x < \frac{1}{2} \\ -1 & \frac{1}{2} \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$

$$\psi_{m,n}(x) = 2^{-\frac{m}{2}} \psi(2^{-m}x - n), \qquad m, n \in Z.$$

The $\psi_{m,n}$ form an orthonormal basis of $L^2(R)$, (and even L^p) but not for Sobolev spaces (unlike trigonometric series for periodic Sobolev spaces). On the other hand, the $\psi_{m,n}$ have good localization properties unlike trigonometric functions (the reverse being true for their Fourier transforms).

A wavelet system is defined by a function $\psi(x)$ and

$$\psi_{m,n}(x) = 2^{-\frac{m}{2}} \psi(2^{-m}x - n)$$

with the property

$$\begin{aligned} L^2(R) &= \oplus_{m \in Z} W_m \\ W_m &= \overline{\text{span } \{\psi_{m,n}\}}, \qquad \text{orthogonal spaces} \end{aligned}$$

$\{\psi_{m,n}, n \in Z\}$ is an orthonormal basis for W_m.

Y. Meyer has constructed a wavelet system with ψ, C^∞ with rapide decay (faster that any power). Later one has constructed a wavelet system with ψ, C^k with exponential decay, and finally I. Daubechies has shown the existence of wavelet systems with compact support and arbitrary regularity. They will be very useful for all kinds of applications.

They are obtained from sequences h_n, with compact support, satisfying *additional assumptions* by the following procedure :

$$\phi(x) = \lim_{k \to \infty} \eta_k(x)$$

with

$$\begin{aligned} \eta_k(x) &= \sqrt{2} \sum_n h_n \eta_{k-1}(2x - n) \\ \eta_0 &= \mathbb{1}_{[} -\frac{1}{2}, \frac{1}{2}[\end{aligned}$$

then

$$\psi(x) = \sqrt{2}\sum_n (-1)^n h_{-n+1}\phi(2x-n)$$

The *most* compact support corresponds to the two possible choices :

$$h_0 = \frac{1\mp\sqrt{3}}{4\sqrt{2}}, \quad h_1 = \frac{3\mp\sqrt{3}}{4\sqrt{2}}, \quad h_2 = \frac{3\pm\sqrt{3}}{4\sqrt{2}}, \quad h_3 = \frac{1\pm\sqrt{3}}{4\sqrt{2}},$$

For more details, see [15], [9].

1.3.2 Cellular Automata

The availability of massively parallel computers, has motivated the use of cellular automata on large lattices for obtaining solutions to P.D.E., in particular the incompressible Navier Stokes equations. A lot of work is necessary to justify this approach.

We describe here a model due to B.M. BOGHOSIAN, C.D. LEVERMORE,[5]. See also U. FRISCH, B. HASSLACHER, Y. POMEAU, [13].

Consider Burgers' equation :

$$\frac{\partial u}{\partial t} + c\frac{\partial}{\partial x}(u - \frac{u^2}{2}) = \nu\frac{\partial^2 u}{\partial x^2}$$

$$\begin{array}{lllll} \text{Replacing} & \dfrac{\partial u}{\partial t} & \text{by} & \dfrac{u(x,t+\Delta t)-u(x,t)}{\Delta t} \\ \text{Replacing} & \dfrac{\partial}{\partial x} & \text{by} & \dfrac{1}{2}\dfrac{u(x+\Delta x,t)-u(x-\Delta x,t)}{\Delta x} \\ \text{Replacing} & \dfrac{\partial^2 u}{\partial x^2} & \text{by} & \dfrac{1}{\Delta x^2}(u(x+\Delta x,t)+u(x-\Delta x,t)-2u(x,t)) \end{array}$$

and choosing $\Delta t, \Delta x$ such that $2\nu = \dfrac{\Delta x^2}{\Delta t}$, we obtain the discretization scheme

$$\begin{aligned} u(x,t+\Delta t) &= \frac{1-\frac{c}{2\nu}\Delta x}{2}u(x+\Delta x,t) + \frac{1+\frac{c}{2\nu}\Delta x}{2}u(x-\Delta x,t) \\ &+ \frac{c\Delta x}{8\nu}(u^2(x+\Delta x,t) - u^2(x-\Delta x,t)) \end{aligned}$$

This can be simulated "approximately" by the stochastic process :

$$\begin{aligned} \xi_1(x+\Delta x,t+\Delta t) &= \frac{1+w(x,t)}{2}(\xi_1(x,t)+\xi_2(x,t)) \\ &- w(x,t)\xi_1(x,t)\xi_2(x,t) \\ \xi_2(x-\Delta x,t+\Delta t) &= \frac{1-w(x,t)}{2}(\xi_1(x,t)+\xi_2(x,t)) \\ &+ w(x,t)\xi_1(x,t)\xi_2(x,t) \end{aligned}$$

where ξ_1, ξ_2 take the values $0, 1, w$ is random and takes the values -1 or 1. The random variables are independent and :

$$Ew = \frac{c}{2\nu}\Delta x$$

It can be proved that :

$$u(x,t) \sim E(\xi_1 + \xi_2)$$

The process ξ_1, ξ_2 is a cellular automata which can be simulated on a *massively parallel computer.*

Reseach on similar types of stochastic processes is important in the context of solving nonlinear P.D.E. on massively parallel machines.

2 Control, identification, estimation.

The applications of these techniques are extremely diversified and come from physical sciences as well as from economic or even social sciences.

We describe some :

- new areas of applications
- new algorithms
- new approaches .

2.1 New areas of application

2.1.1 Environmental studies. The program "Global Change"

In view of the growing importance of environmental issues, a worldwide program of research has been developing in recent years, under the name of "Global Change". It connects specialists of Climate Dynamics, Oceanography, Planetary Physics,... It seems that this direction is a source of important mathematical problems, of somewhat new nature.

The basic problem deals with the *prediction* of physical quantities, solutions of a set of nonlinear evolution P.D.E., with *unknown* parameters and unknown initial state. Nonlinearity creates an important sensitivity with respect to initial data and unknown quantities, resulting in a lack of predictability beyond some length of time. A fundamental question is to identify the *important regimes* of the physical variables, those which contain the main futures of interest and are *persistent.* There are several ways to give a mathematical meaning to this question. The interesting feature is that they result in a *mixture of statistical and dynamical methods.* A lot of work is needed in that direction, even for simple nonlinear systems.

The point of view of dynamical systems is to obtain the stationary solutions of the nonlinear P.D.E. (or system of P.D.E.) and the long-time behaviour of solutions. This is the theory of *attractors.*

A complementary statistical theory has been developed, for which we describe only two ideas, that of *persistent anomalies* and that of *EOF analysis* (Empirical orthogonal functions).

Consider a vector representing physical variables (typically a flow) which is computable through a model, which is not in general completely known (this is an important difficulty, which we leave aside). We represent it by $\psi_k(t), k = 1 \ldots N$ where k may represent a point x_k on a grid, or a component if the solution is obtained by an expansion.

We set $< \psi_k >=$ average of $\psi_k(t)$ over some record of data.

The instantaneous *anomaly* is defined by :

$$\widetilde{\psi}_k(t) = \psi_k(t) - < \psi_k >$$

The *pattern correlation* between an anomaly at time t and at a later time $t + \tau$ is defined by :

$$p(t, \tau) = \frac{\sum_k \widetilde{\psi}_k(t)\widetilde{\psi}_k(t+\tau) - (\sum_k \widetilde{\psi}_k(t))(\sum_k \widetilde{\psi}_k(t+\tau))}{\sigma(t)\sigma(t+\tau)}$$

where

$$\sigma(t)^2 = \sum_k \widetilde{\psi}_k(t)^2 - (\sum_k \widetilde{\psi}_k(t))^2.$$

We say that an anomaly $\psi_k(t_0)$ *persists* from $t = t_0$, to $t = t_0 + J\tau$, if :

$$p(t_j, \tau) \geq p_0, \text{ where } t_j = t_0 + j\tau, j = 0 \ldots J - 1$$

and p_0 represents the persistence criterion. What is *expected* is that the anomalies which satisfy a reasonable persistence criterion fall into a *small number of easily identifiable patterns*, related to the attractors of dynamical system.

The EOF analysis goes as follows. Let :

$$\Gamma_{k\ell} = < \widetilde{\psi}_k \widetilde{\psi}_\ell >$$

Consider the eigenvalues of the matrix $\Gamma \quad \lambda^1 \ldots \lambda^N$, ranked in decreasing order and $e^1 \ldots e^N$ are the corresponding eigenvectors, called the 1st EOF, the 2nd EOF, ...

Next expand the vector $\widetilde{\psi}(t) = (\widetilde{\psi}_k(t))$ on the basis $e^1 \cdots e^N$, hence :

$$\widetilde{\psi}(t) = \sum_{i=1}^{N} \alpha_i(t) e^i$$

then one can easily check that :

$$< \alpha_i \alpha_j >= \lambda^i \delta_{ij}.$$

The coefficients $\alpha_i(t)$ are called the *principal components*. The EOF are interpreted as directions of variability of the anomaly, λ^i representing the part of the variance related to EOF e^i (the total variance being $\lambda^1 + \ldots + \lambda^N$). The important *conjecture* is that the main OEF are related to the patterns associated to persistent anomalies.

In [16] theses connections are exhibited experimentally on some models.

Is there a general theory for these phenomenon, at least for some class of nonlinear dynamical systems ? This is an open question, which has a crucial importance for the understanding of the variability of atmospheric dynamics.

2.1.2 Computer vision

2.1.2.1 The segmentation problem An image can be represented by a function $g(x)$ measuring the strength of the light signal striking a plane at point x. Such a function is expected to have discontinuities reflecting edges of objects, and shadows. Outside such lines the function g is expected to behave more smoothly.

Having this in mind, one defines a segmentation of a region Ω, as a set of open connected subsets $\Omega_i, i = 1 \cdots n$, each one with a piecewise smooth boundary and Γ is the union of the parts of the boundaries of the Ω_i inside Ω.

An approximation of g is a function u which is differentiable on $\Omega - \Gamma$. One defines a cost function :

$$J(u, \Gamma) = \mu \int_\Omega (u-g)^2 dx + \int_{\Omega-\Gamma} |Du|^2 dx + \nu|\Gamma|$$

The *segmentation problem* consists in minimizing the functional J over the pair (u, Γ). Note that if $\nu = 0, inf J = 0$.

This is a new class of problems in the calculus of variations, introduced in [17].

It has attracted a lot of interest and some progress has been made, concerning existence, and approximation.

It is interesting to consider the one dimensional problem, in which $\Omega = (0,1), \Gamma = \{a_1; \cdots a_N$, with $0 < a_1 < a_2 \cdots < a_N < 1\}$ and $|\Gamma| = N$. One has :

$$J(u, a_1, \cdots, a_N) = \mu \int_0^1 (u-g)^2 dx + \sum_{i=0}^{N} \int_{a_i}^{a_{i+1}} u'^2 dx + \nu N$$

and we have defined $a_0 = 0, a_{N+1} = 1$.

Since we do not impose continuity at points a_i, we may write preferably :

$$J(u_1, \cdots, u_N; a_1, \cdots, a_N) = \mu \sum_{i=0}^{N} \int_{a_i}^{a_{i+1}} [(u_i - g)^2 + u'^2] dx + \nu N$$

There is a probabilistic interpretation of J. Consider in the segment (a_i, a_{i+1}) a process x_i such that :

$$x_i(t) = x_i(a_i) + w_i(t), \qquad t \in [a_i, a_{i+1}[$$

where $x_i(a_i)$ is not random, and $w_i(t)$ is a standard Wiener process. We observe on (a_i, a_{i+1}) the process $y_i(t)$ with :

$$dy_i(t) = x_i(t)dt + db_i \qquad y_i(a_i) = 0$$

where b_i is Wiener process independent from w_i.

The "a priori probability" of the trajectory $x_i(t)$ to coincide with a given function $u_i(t)$ which is $H^1(a_i, a_i + 1)$ is :

$$\exp -\frac{1}{2} \int_{a_i}^{a_{i+1}} [(u_i'^2 + u_i^2)dt - 2u_i dy_i]$$

For details see [20].

Considering independent processes in each interval, we obtain :

$$\exp -\frac{1}{2} \sum_{i=0}^{N} \int_{a_i}^{a_{i+1}} [(u_i'^2 + u_i^2)dt - 2u_i dy_i]$$

and the maximization of this probability results in minimizing J, up to the correspondence $dy_i \to g$ on (a_i, a_{i+1}). It would be extremely interesting to treat the 2 dimensional problem, which is the real one, by similar probabilistic methods. It is an open problem.

2.1.2.2 Axiomatic derivation of image processing models We describe here a new approach to image processing due to L. Alvarez, F. Guichard, P.L. Lions and J.M. Morel [4]. Consider the signal $g(x)$ representing the image. We look at it at a scale t, measuring roughly speaking the size of details of the image (small t means fine scale, while large t means coarse scale). An analysis at scale t is a transformation $T_t g$. A multiscale analysis is thus a family, parametrized by $t \geq 0$, of nonlinear operators (or filters).

Of course, some conditions have to be made on the operator T_t, in order to fulfill physical requirements of the filter. These restrictions or axioms are such that the function $u(x,t) = (T_t g)(x)$ appears as the solution of a fully nonlinear, parabolic, possibly degenerate second order equation

$$\begin{aligned} &\frac{\partial u}{\partial t} = F(Du, D^2u) \\ &u(x,0) = g(x). \end{aligned} \tag{4}$$

In fact, the choice of the function F is equivalent to the choice of the family T_t. Among physical requirements, one has the following main one

$$F(p, A) \leq F(p, B), \forall p, A \leq B$$

which is in fact the condition which suffices to give a meaning to (4) in viscosity sense.

Examples :

- The Gaussian pyramid. It corresponds simply to the heat equation

$$F(p, A) = \frac{1}{2} trA$$

- Quasilinear filters :

$$F(p, A) = a(|p|)trA + a'(|p|)\frac{(Ap, p)}{|p|}$$

where

$$a \geq 0, \qquad a(|p|) + a'(|p|)|p| \geq 0$$

- Morphological filters :

$$F(p) = \inf_{q \in S} \quad p.q$$

where S is a compact set of R^2.

- Curvature operators

$$F(p, A) = |p| G(\frac{1}{|p|}(trA - \frac{Ap \cdot p}{|p|^2}))$$

with possible $G(s) = s$ or $|s|^{\alpha-1}s$ (in particular $\alpha = \frac{1}{3}$).

2.1.2.3 Mobile Robotics Consider the problem of a mobile robot which tries to recover its environment, during its motion (the environment is assumed to be static). The robot is equipped with a camera, which takes images between time intervals. One way of approaching the problem is to extract tokens from the images in the sequence, match them from image to image and recover the motion and the structure of the environment.

Naturally, the tokens we compute in the images should be closely related to objects in the scene, if we want the matches to be meaningful. They are in general surface markings, shadows, depth discontinuities.

Let us explain the general ideas in the case of a point M, which is the object to be recognized by the mobile robot (see Figure 1). So M is the real point, C_1, C_2 represent the motion of the camera (installed on the robot), m_1, m_2 the images of M. The motion is decomposed into a rotation R with a rotation axis going through C_1, and a translation $t = C_1C_2$.

If we consider a coordinate system attached to the camera, then we can measure C_1m_1 and C_2m_2 with the local coordinate system. The coordinates with respect to a common coordinate system, that related to C_1 are C_1m_1

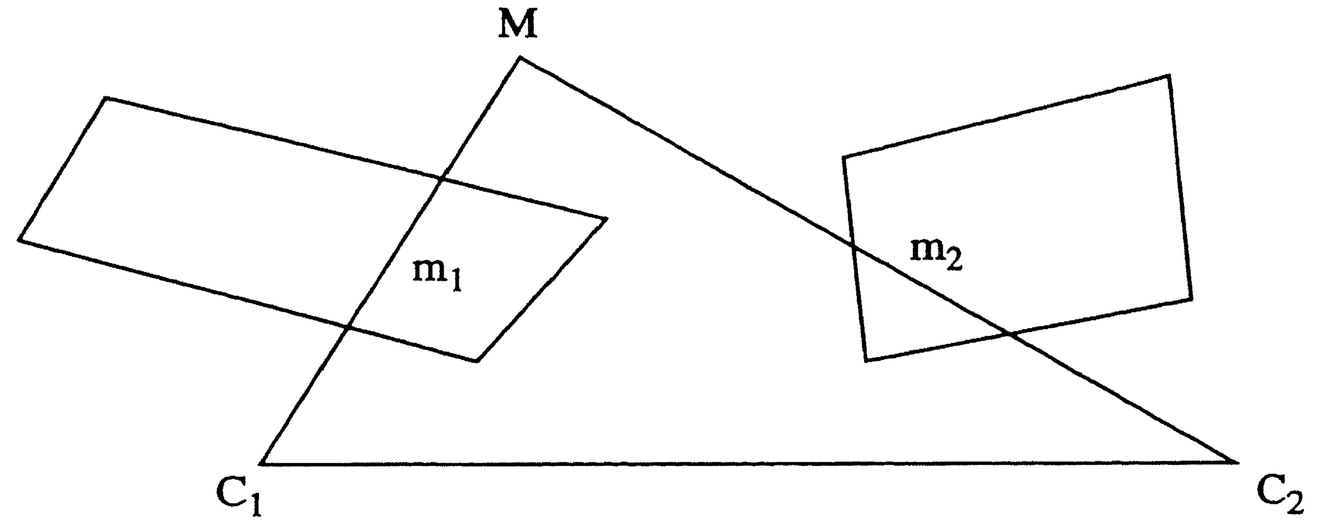

FIGURE 1.

and $\mathcal{R}C_1m_1$. Then one expresses the planarity constraint, namely that C_1m_1, C_2m_2 and t are coplanar ; it amounts to :

$$C_1m_1 \cdot (t \wedge \mathcal{R}C_2m_2) = 0.$$

The vector t has coordinate t_x, t_y, t_z but from the linearity, we can assume that $\|t\| = 1$, hence 2 parameters are enough. The matrix $\mathcal{R}$ depends of 3 parameters which characterize the unit rotation axis (2 parameters) and the rotation angle.

Conceptually, what is important is to recognize that the previous relations amounts to :

$$f(x, a) = 0$$

where a is a vector of parameters $\in \mathbb{R}^n$, and x is a vector of measurement $\in \mathbb{R}^n$ and f is a nonlinear relation.

Each successive image leads to a relation :

$$f(x_k, a) = 0$$

However the observation is not exact and rather described by the model

$$z_k = x_k + \nu_k$$

where ν_k is a white noise of covariance Γ. Considering that

$$a_{k+1} = a_k = a$$

we are in the framework on nonlinear filtering if we can express x_k as a function of a_k. It is of course natural to linearize around a given estimate of a, and to

use extended Kalman filtering. Once t, R is obtained, one can recover M by expressing the relations :

$$\lambda C_1 m_1 = t + \mu R C_2 m_2$$

where λ, μ are unknown scalars. In this relation again t, R are known random variables, as well as $C_1 m_1, C_2 m_2$. Thus we are in a situation similar to the above and can use again a Kalman filter.

These techniques have been extensively used in the context of mobile robotics by O. FAUGERAS and his team, see for instance [11].

2.2 New algorithms

2.2.1 Parallel algorithms

The development of multiprocessors has generated a substantial interest in the obtaining of *parallel algorithms*. A thorough analysis is needed, since surprises can arise in comparison with the sequential approach.

Take for instance Jacobi and Gauss Seidel iterations for obtaining a fixed point of :

$$x = f(x) \qquad x \in \mathbb{R}^n$$

A Jacobi iteration is the following :

$$x_i^{k+1} = f_i(x^k), \qquad i = 1 \cdots n$$

and a Gauss Seidel is :

$$x_i^{k+1} = f_i(x_1^{k+1}, \cdots, x_{i-1}^{k+1}, x_i^k, \cdots, x_n^k)$$

The advantage of Gauss Seidel iteration is that it converges more frequently that Jacobi, and sequentially it performs much better (the convergence rate of Gauss Seidel iteration is better).

Parallel implementation will change the situation considerably.

Consider the case when there are n processors, and the sequence x^k such that :

$$x^{k+1} = f(x^k)$$

denoted by $x^{k,J}$ (Jacobi sequence) converges towards the fixed point. Suppose also f monotone, i.e. $f(x) \leq f(y)\ \forall x, y$ with $x \leq y$. Then take a sequence $x^{k,U}$ defined by :

$$\begin{aligned} x_i^{k+1,U} &= f_i(x^{k,U}), & \forall i \in U_k \\ x_i^{k+1,U} &= x_i^{k,U}, & \forall i \notin U_k \end{aligned}$$

U_k is a subset of $\{1, \cdots, n\}$.

One can prove (T.N. TSITSIKLIS) that if one starts with the same initial value x^0 and $f(x^0) \leq x^0$ or $x^0 \leq f(x^0)$, then :

$$x^* \leq x^{k,J} \leq x^{k,U}, \qquad \forall k$$

where x^* is the limit fixed point. Hence Jacobi iteration performs better than any parallel version of Gauss Seidel iteration. When they are less than n processors available, or the assumption of monotonicity is not satisfied, no general statement can be made (see [4])

2.2.2 Simulated Annealing and global optimization

This type of algorithm has been developed in the recent years in order to obtain a *global minimum* for a function $U(x)$, over $x \in B, B$ compact, in the case when U is smooth. It is clear that such a problem occurs in many applications. Simulated annealing has first been used in the context of image processing.

The algorithm consists in a discrete version of the following stochastic differential equation:

$$dx_t = -DU(x_t)dt + c_t\sigma(x_t)dw_t, \quad x(0) = x$$

where the following assumptions are made

- U is C^2 from B to $[0, \infty)$ and

  $$Min_{x \in B} U(x) = 0, \qquad DU(x) \cdot x > 0, \quad \forall x \in B - \overset{\circ}{B}_1$$

  where B is a ball in $\mathbb{R}^n$, centered at the origin, and B_1 is an other ball, also centered at the origin and strictly included in B.
- σ is Lipschitz continuous from B to $[0, 1]$, with $\sigma = 1$, for $x \in B_1, \sigma = 0$ for $x \in \partial B, \sigma > 0$ on $\overset{\circ}{B}$
- $c_t = \dfrac{c}{Logt}$, for t large, $c > 0$.
- w_t standard Wiener process in $\mathbb{R}^n$
- $\pi^\varepsilon(x) = \dfrac{1}{Z^\varepsilon}(exp - \dfrac{2U(x)}{\varepsilon^2})\mathbb{1}_B$ with $\int \pi^\varepsilon(x)dx = 1$ converges weakly to a probability π as $\varepsilon \to 0$.

Note that π is a probability concentrated on the set of global minima of U(.).
Then the following result can be proved :

$$Ef(x_t) \to \pi(f)$$

$\forall f$ bounded, continuous, as $t \to \infty$, uniformly for x (the initial value) in B.
(For more details see [7]).

2.3 New approaches

Let us just mention the developments related to H_∞ theory and which permit to obtain protection of dynamic systems from disturbances via *feedback control* . We just mention some recent results concerning linear systems.

Let us consider the linear system

$$\begin{aligned} \dot{x} &= Ax + Bu + Dw, \\ y &= Cx \end{aligned}$$

where w represents a disturbance, and u a control. We consider *feedback* controls, $u = Ky$. The *transfer* matrix $T_K(s)$ is given by

$$T_K(s) = C[sI - (A + BK)]^{-1}D$$

and we consider those K for which $A + BK$ is stable. The H_2 norm is defined by :

$$\|T_K\|_2 = \left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} \text{tr } T_K(-j\omega)^* T_K(j\omega) d\omega\right)^{\frac{1}{2}}$$

and the H_∞ norm is defined by :

$$\|T_K\|_\infty = \sup_{\omega \in R}\left(\text{ tr } T_K(-j\omega)^* T_K(j\omega)\right)^{\frac{1}{2}}$$

which are finite since $A + BK$ is stable.

The problem of H_∞ or H_2 control *consists* in minimizing the above norms with respect to K.

Note that

$$\|T_K\|_\infty = \sup_w\left\{\left(\int_0^\infty |y(t)|^2 dt\right)^{\frac{1}{2}} | \left(\int_0^\infty |w(t)|^2 dt\right)^{\frac{1}{2}} \leq 1\right\}$$

and thus this norm expresses the *sensitivity of the system* with respect to external disturbances.

Among the important results obtained recently, it has been proven that *we can chose* a K such that $\|T_K\|_\infty \leq \gamma, \forall\gamma$ given, *if* there exists ε such that one can solve the Riccati equation

$$PA + A^*P - \frac{1}{\varepsilon}PBB^*P + \frac{1}{\gamma}PDD^*P + \frac{1}{\gamma}CC^* + \varepsilon I = 0$$

In fact $K = -\dfrac{B^*P}{2\varepsilon}$ will serve for this purpose (for more details, see [14]).

3 Discrete systems

3.1 Discrete event systems

New applications strongly related to information technology have created the need to develop a theory of DEDS, *discrete event dynamic systems.* Such applications are production or assembly lines, computer/communication networks, traffic systems, ... A special issue of IEEE, Jan. 1989 is devoted to dynamics of discrete event systems.

Many new mathematical techniques have been developed in this context. We describe here one of them, the use of an algebraic structure, called *dioid*, in the modelling of *timed event graphs.*

Let us just recall the basic definition of a dioid. It is a set $\mathcal{D}$ provided with two inner operations $\oplus$ and $\otimes$ (addition and multiplication) such that

- they are both associative

- addition is commutative

- multiplication is right distributive with respect to addition

- there exists a null and identity elements

$$\exists \varepsilon \in \mathcal{D} \ : \ \forall a \in \mathcal{D}, \qquad a \oplus \varepsilon = a$$
$$\exists e \in \mathcal{D} \ : \ \forall a \in \mathcal{D}, \qquad a \otimes e = e \otimes a = a$$

- the null element is absorbing

$$\forall a \in \mathcal{D}, \quad a \otimes \varepsilon = \varepsilon \otimes a = \varepsilon$$

- the addition is idem potent

$$\forall a \in \mathcal{D}, \quad a \oplus a = a.$$

When addition is commutative, the dioid is called commutative. As an example take $\mathcal{D} = Z \cup \{-\infty\} \cup \{+\infty\}$ and

$$\oplus = \max, \qquad \otimes = +$$
$$\varepsilon = -\infty, \quad e = 0$$

(note that we impose the rule $(-\infty) \otimes (+\infty) = (-\infty)$).

We can also consider

$$\oplus = \min, \qquad \otimes = +$$
$$\varepsilon = +\infty, \qquad e = 0$$

(in which case $(-\infty) \otimes (+\infty) = +\infty$).

A dioid is a structure somewhere between linear algebra and lattices.
One can define a partial order relation

$$a \geq b \Leftrightarrow a = a \oplus b$$

and a *pseudo left inverse* denoted $a \backslash c$ which is the greatest subsolution of

$$a \otimes x = c.$$

Starting with these premises affine equations can be solved, as well as matrices defined and a matrix calculus is available. Matrix equations can also be solved. Let us see briefly how these concepts apply to timed event graphs.

Times event graphs are a special kind of Petri nets. They are directed graphs with two types of edges, *places* and *transitions*

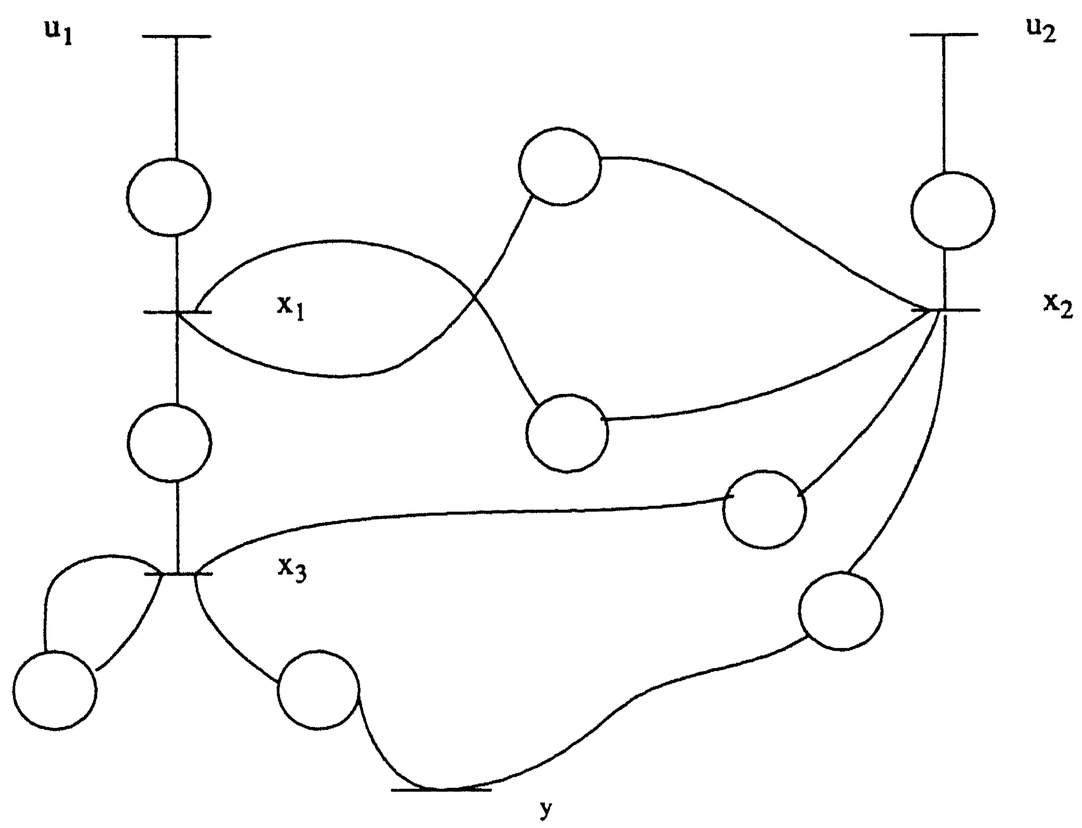

FIGURE 2.

In Figure 2,, the transitions are $u_1, u_2, x_1, x_2, x_3, y$ and the places are denoted by $x_1|u_1$, $x_2|u_2$, $x_3|x_1$, $x_3|x_2$, $x_3|x_3$, $y|x_3$, $y|x_2$, $x_1|x_2$, $x_2|x_1$.

There is a single transition upstream and downstream, at each place.

In places, there are tokens or not. Tokens are created or consumed when transitions are fired, more precisely when a transition t is fired one token is consumed at each place which precedes t and one is created at each place which succeeds it.

Let us assume that transitions are immediate, but a token must stay at a place an amount of time called the holding time, which depends on the place. The following symbols are used

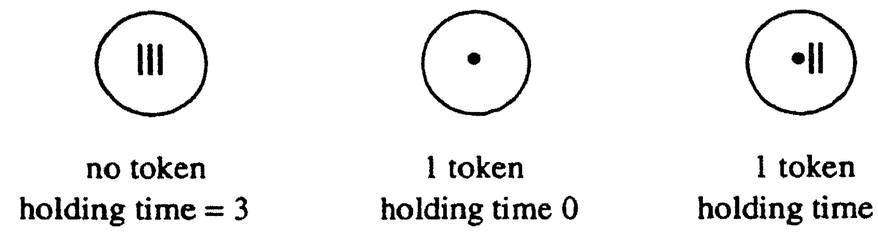

FIGURE 3.

For instance consider the places which precede x_1, we complete the information as follows

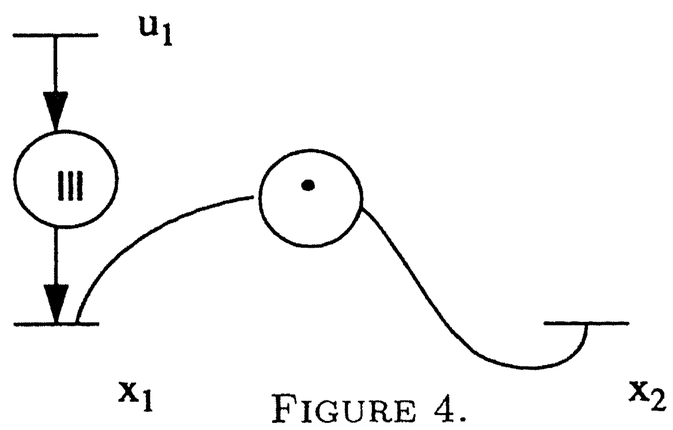

FIGURE 4.

Let for a transition x, x_n be date at which transition x has been fired for the n^{th} time. We can write the relation

$$(x_1)_n = \max[(x_2)_{n-1}, (u_1)_n + 3]$$

and of course similar relations for other transitions.

If the dates take values in $Z \cup \{+\infty\} \cup \{-\infty\}$, then we can work with the dioid considered above $\mathcal{D}$, with the operations $\oplus = \max$, $\otimes = +$.

The preceding relation writes

$$(x_1)_n = (x_2)_{n-1} \oplus 3(u_1)_n$$

where $3(u_1)_n = 3 \otimes (u_1)_n$ to simplify the notation.

One of the objectives of research in these directions is to obtain a theory similar to that of linear dynamic systems. In particular a theory of *stability* is being developed. This is important to obtaining an evaluation of performances for the real system which is modelled by the event graph. (See [8]).

3.2 Hybrid systems

An hybrid system is a system whose state contains continuously as well as finitely valued variables. When the discrete variables take a given set of values, the continuous variables behave as the solutions of differential equations. Transitions between the possible sets of values of the discrete variables are obtained through the action of a monitor (a controller). The action of the controller may be instantaneous or require some delay. The objective is in general to keep the continuous variables within a given range. Decisions are taken as feedbacks.

An hybrid system will be characterized by a given feedback, and the problem is to prove that this feedback rule fulfills the goal.

EXAMPLE 1 Suppose we want to control the temperature of a room through a thermostat, which can turn instantaneously a heater on and off. The temperature is the continuous variable $x(t)$, $\nu(t) = 1$ or 0 according whether the heater is on or off is the discrete variable. We have :

$$\begin{aligned} \dot{x} &= -Kx \quad \text{if} \quad \nu = 0 \\ \dot{x} &= K(h - x) \quad \text{if} \quad \nu = 1 \end{aligned}$$

If $d(t)$ is the decision taken by the thermostat, $d(t) = 1$ or 0 and we have:

$$\nu(t + 0) = d(t)$$

We want to maintain $x(t)$ between m and M. Then we take

$$\begin{aligned} d(t) = 1 \quad &\text{if} \quad x(t) = m \quad \text{and} \quad \nu(t) = 0 \\ d(t) = 0 \quad &\text{if} \quad x(t) = M \quad \text{and} \quad \nu(t) = 1 \end{aligned}$$

and $d(t) = \nu(t)$ otherwise . Such a feedback fullfills the objective.

EXAMPLE 2 Suppose we control the water level in a tank through a monitor which can turn a pump on and off. The water level is $x(t)$, and we set $\nu(t) = 1$ if the pump is on, and $\nu(t) = 0$ if it is off. We have

$$\begin{aligned} \dot{x} &= -2 \quad \text{if} \quad \nu(t) = 0 \\ \dot{x} &= 1 \quad \text{if} \quad \nu(t) = 1 \end{aligned}$$

Let $d(t)$ be the decision taken by the monitor, $d(t) = 1$ or 0 and suppose there is a delay of 2 before the decision is executed then :

$$\nu(t) = d(t - 2)$$

We wish to keep the water level between 1 and 12. We then consider the feedback

$$d(t) = 1 \quad \text{if} \quad x(t) = 5 \text{ and } \nu(t) = 0$$
$$d(t) = 0 \quad \text{if} \quad x(t) = 10 \text{ and } \nu(t) = 1$$

and $d(t) = \nu(t)$ otherwise such a feedback fulfills the desired behaviour. In general, proving that a specific feedback satisfies a given objective of the continuous variables is not easy. Results on decidability of such a problem are available for a particular class of Hybrid systems (cf. R. Alur et al. [1]).

4 New areas of information technology

Let us mention only some recent mathematical problems motivated by I.T. (again this is by no means exhaustive).

4.1 Artificial intelligence

Since artificial intelligence needs to deal with *qualitative* aspects, more than with quantitative aspects (or in connection with them), this has motivated the development of *qualitative simulation* (or *qualitative physics*) in particular at Xerox Parc. Note that the economists needed much before similar techniques, in the context of the theory of *comparative economics* (P.A. SAMUELSON).

Our presentation here relies on some recent work of J.P. AUBIN.

We pose the problem of the *qualitative evolution* of solutions to a differential equation

$$\dot{x} = f(x_t) \qquad x \in \mathbb{R}^n$$

and more precisely to the *qualitative evolution* of a set of functionals

$$V_1(x_t), \ldots, V_m(x_t)$$

which are of importance (energy, entropy, indicators, ...).

The qualitative behavior is expressed by the evolution of the functions sign $(\frac{d}{dt}V_j(x_t))$ with values in $\mathcal{R}^m = \{-1, 0, +1\}^m$.

This is the problem of interest. But we want to obtain this evolution, without solving the equation, since some independence should be obtained with respect to the *initial condition.*

Since sign $(\frac{d}{dt}V_j(x_t)) =$ sign $(DV_j(x_t)f(x_t))$ it is convenient to introduce in the closed subspace K of $\mathbb{R}^n$, where lives x_t, the *qualitative cells*

$$K_a = \{x \in K | \text{ sign } (DV_j(x)f(x)) = a_j\}$$

where $a \in \mathcal{R}^m$, and their closure (*large qualitative cells*)

$$\bar{K}_a = \{x \in K | \text{ sign } (DV_j(x)f(x)) = a_j \text{ or } 0\}.$$

Let $\mathcal{D}(f,V)$ be the subset of qualitative states a such that K_a is not empty. Let also denote by $x(t;x_0)$ the solution of the differential equation corresponding to an initial date x_0. One is interested in the study of *transitions between qualitative cells.*

If $b \in \mathcal{D}(f,V)$, we say that $c \in \mathcal{D}(f,V)$ is a *successor* of b, if $\forall x_0 \in \bar{K}_b \cap \bar{K}_c$, there exists $\tau > 0$, such that $x(t;x_0) \in K_c$, for all $t \in]0,\tau[$.

A qualitative state a is a *qualitative equilibrium*, if it is its own successor. It is said to be a *qualitative repellor* if $\forall x_0 \in \bar{K}_a$, there exists $t > 0$ such that $x(t;x_0) \notin \bar{K}_a$.

The theory developed by J.P. Aubin permits to characterize the map of successors, the qualitative equilibria, and the qualitative repellors.

It has been applied to the so-called "replicator systems", a prototype of which is the differential system ([2])

$$\dot{x}_i = x_i(\alpha_i - \sum_{j=1}^{n} \alpha_j x_j)$$

4.2 Neural networks

The basic neural network can be viewed as an undirected graph with n nodes, to which are attached a pair (W,θ) where

W is an $n \times n$ symmetric matrix, W_{ij} is the weight attached to the edge (i,j), $W_{ii} = 0$

θ is an n vector, θ_i is the threshold attached to the node i.

Nodes are called *neurons*. Each neuron has two possible states $(1,-1)$. Let v be the state of the neural network, v_i being the state of neuron i.

Let

$$E_i(v) = -\sum_{j=1}^{n} W_{ij} v_j + \theta_i$$

then the following calculation is performed by the network

$$\begin{aligned} v_i^{k+1} &= \text{ sign } (E_i(v^k)), \quad \text{for } i \in S^k \\ v_i^{k+1} &= v_i^k \text{ for } i \notin S^k \end{aligned}$$

where S^k is a subset of the neurons.

For instance if

$$k = hn + j \quad j = 0 \ldots n-1$$

and $S^k = \{j+1\}$, the *network operates in serial mode.*

Note that in our notation

$$\text{sign } (a) = \begin{cases} 1 & \text{if} \quad a \geq 0 \\ -1 & \text{if} \quad a < 0 \end{cases}$$

A *stable* state is a state such that

$$v^{k+1} = v^k = v.$$

A basic theorem of HOPFIELD is that if the network operates in serial mode, then it will converge to a *stable state*.

The applicability of neural networks in practice arises from the possibility of interpreting the stable states. For instance, in pattern recognition, the stable states are known patterns, and for a given input pattern, the network will converge to the known pattern which is the closest to the input. It is clear that the neural network realizes the following search problem

$$\min E(v) = -\frac{1}{2}\sum_{ij} W_{ij}v_iv_j + \sum_i \theta_i v_i \quad v_i = \{-1, +1\}$$

and attains a *local* minimum.

One can clearly consider many variants of the above problem. For instance consider the following model in continuous time

$$\begin{aligned} v_i(t) &= g(u_i(t)) \\ \frac{du_i}{dt} &= -E_i(v(t)) \end{aligned}$$

where g is an increasing function from R to $[0,1]$ and $E_i(v) = \dfrac{\partial}{\partial v_i}E(v)$, $E(v)$ energy function (for instance the above). It will converge towards a local minimum of $E(v)$. It can be realized as an analog integrated circuit.

In the spirit of *simulated annealing*, considered above, one can try to attain a *global* minimum of the Energy function, by considering a stochastic version of the preceding model. This has been done by E. WONG.

Consider the model

$$\begin{aligned} v_i(t) &= g(u_i(t)) \\ du_i &= -E_i(v(t))dt + \sqrt{\frac{2T}{g'(u_i(t))}}dw_i \end{aligned}$$

where T is a constant, and w_i are independent standard Wiener processes. The stationary probability density of the process $v(t)$ is

$$p(v) = \frac{1}{Z}\exp -\frac{1}{T}E(v)$$

where Z is the normalization factor.

The simulated annealing adaptation of the preceding algorithm (for instance take $T(t) \to 0$ as $t \to \infty$) remains to be done. For more details, see [6] and [19]).

4.3 Analytic analysis of algorithms

Computer science leads quite frequently to combinatorial algorithms. A quite interesting approach of P. Flajolet [12] has shown how generating functions and complex analysis provide a way to treat these problems. In particular, formal languages, tree enumerations, comparison based searching and sorting, digital structures, hashing and occupancy have been interesting applications.

A class of combinatorial structures is a pair of a finite or denumerable set $\mathcal{A}$, whose elements are called the atoms.

Each atom $\alpha \in \mathcal{A}$ will have a size $|\alpha|$. We can perform the following operations :

The product relation $\mathcal{C} = \mathcal{A} \times \mathcal{B}$:

$$\mathcal{C} = \{\gamma \in \mathcal{C} | \gamma = (\alpha, \beta), \alpha \in \mathcal{A}, \beta \in \mathcal{B}\} \qquad \text{with} \qquad |\gamma| = |\alpha| + |\beta|$$

The union relation $\mathcal{C} = \mathcal{A} + \mathcal{B}$ $\quad$ $\mathcal{C} = \mathcal{A} \cup \mathcal{B}$ where $\mathcal{A} + \mathcal{B}$ are disjoint.

The sequence $\mathcal{C} = \mathcal{A}^*$

$$\mathcal{C} = \{\varepsilon\} + \mathcal{A} + \mathcal{A} \times \mathcal{A} + \mathcal{A} \times \mathcal{A} \times \mathcal{A} + \cdots$$

where $|\varepsilon| = 0$. The set construction $\mathcal{C} = \mu(\mathcal{A})$, is the collection of all subsets of $\mathcal{A}$:

$$\mathcal{C} = \{\{\alpha_1, \cdots, \alpha_k, \cdots\}| \qquad \alpha_1, \cdots, \alpha_k, \cdots \text{ in} \mathcal{A}, \alpha_1, \cdots \alpha_k \cdots \text{ different}\}.$$

The multi set construction $\mathcal{C} = M(\mathcal{A})$ allows repetitions.

The cycle construction $= \mathcal{C}(\mathcal{A})$ is the set whose elements are (non empty) cycles of elements of $\mathcal{A}$.

Let A_n be the number of elements of $\mathcal{A}$, whose elements are (non empty) cycles of elements of $\mathcal{A}$. Let A_n be the number of elements of $\mathcal{A}$, whose size is n, then the interesting problem is to calculate the C_n corresponding to the more complex structure $\mathcal{C}$. This is where the generating functions are useful. Define

$$A(z) = \sum_n A_n z^n = \sum_{\alpha \in \mathcal{A}} z^{(\alpha)}$$

and

$$C(z) = \sum_n C_n z^n = \sum_{\gamma \in \mathcal{C}} z^{(\gamma)}$$

It is possible to express $C(z)$ in function of $A(z)$. For instance, for $\mathcal{C} = \mathcal{A} \times \mathcal{B}$ one has :

$$C(z) = \sum_{(\alpha,\beta) \in \mathcal{A} \times \mathcal{B}} z^{|\alpha|+|\beta|} = A(z)B(z)$$

For $\mathcal{C} = \mathcal{A} + \mathcal{B}$

$$C(z) = \sum_{\alpha \in \mathcal{A}} z^(\alpha) + \sum_{\beta \in \mathcal{B}} z^{|\beta|} = A(z) + B(z)$$

For $\mathcal{C} = \mathcal{A}^*$

$$\begin{aligned} C(z) &= 1 + A(z) + A(z)^2 + \cdots \\ &= \frac{1}{1 - A(z)} \end{aligned}$$

For $\mathcal{C} = \mu(\mathcal{A})$, we note that

$$\mu(\mathcal{A}) = \Pi_{\alpha \in \mathcal{A}}(\{\varepsilon\} + \{\alpha\})$$

hence

$$\begin{aligned} C(z) &= \Pi_{\alpha \in \mathcal{A}}(1 + z^{|\alpha|}) = \Pi_n (1 + z^n)^{A_n} \\ &= exp(A(z) - \frac{A(z^2)}{2} + \frac{A(z^3)}{3} + \cdots) \end{aligned}$$

For $\mathcal{C} = M(\mathcal{A})$ we have

$$M(\mathcal{A}) = \Pi_{\alpha \in \mathcal{A}} \{\alpha\} *$$

hence

$$\begin{aligned} C(z) &= \Pi_{\alpha \in \mathcal{A}} \frac{1}{1 - z^{|\alpha|}} = \Pi_n (1 - z^n)^{A_n} \\ &= exp(A(z) - \frac{A(z^2)}{2} + \frac{A(z^3)}{3} + \cdots) \end{aligned}$$

Consider further $\mathcal{C} = M_2(\mathcal{A})$, the collection of subsets of $\mathcal{A}$ with cardinality 2, with possible repetition. Then

$$\begin{aligned} C(z) &= \sum_{|\alpha_1| > |\alpha_2|} z^{|\alpha_1| + |\alpha_2|} + \sum_{\alpha} z^{2|\alpha|} \\ &= \frac{1}{2} \sum_{|\alpha_1| \neq |\alpha_2|} z^{|\alpha_1| + |\alpha_2|} + \sum_{\alpha} z^{2|\alpha|} \\ &= \frac{1}{2} \sum_{\alpha_1, \alpha_2} z^{|\alpha_1| + |\alpha_2|} + \frac{1}{2} \sum_{\alpha} z^{2|\alpha|} \end{aligned}$$

hence

$$C(z) = \frac{1}{2}(A(z))^2 + \frac{1}{2} A(z^2).$$

From the previous structures, it is possible to construct further complex structures, which will lead to functional equations. For instance, consider in Figure 5 the structure of binary trees (the size of a binary tree is the number of leaves)

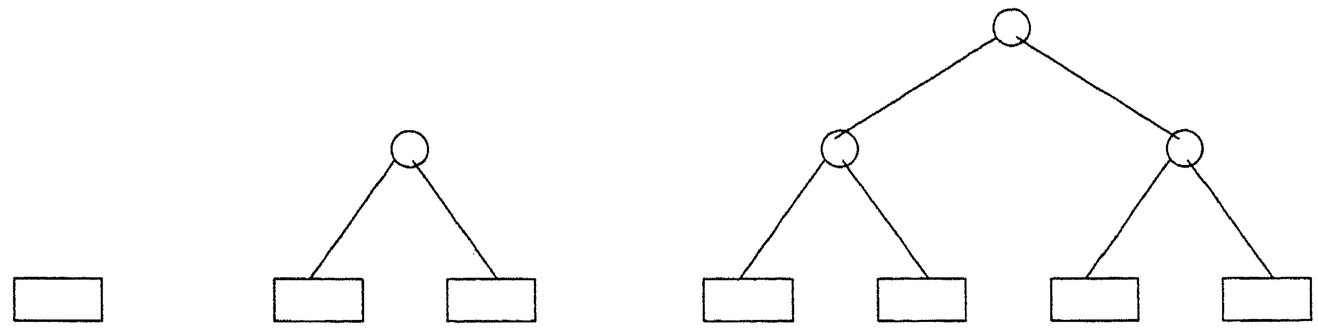

FIGURE 5.

Let $\mathcal{A} = \{\text{leaf}\}$, then

$$\mathcal{B} = \mathcal{A} + \mathcal{B} \times \mathcal{B}$$

hence

$$B(z) = z + (B(z))^2$$

which yields a unique formal power series solution

$$B(z) = \frac{1}{2}(1 - \sqrt{1 - 4z})$$

Similarly consider trees with multiples branches (at least 2), each branch having at least 2 leaves, one has

$$\mathcal{B} = \mathcal{A} + \mathcal{B} \times \mathcal{B} + \mathcal{B} \times \mathcal{B} \times \mathcal{B} + \cdots$$

hence

$$B(z) = z + \frac{(B(z))^2}{1 - B(z)}$$

which obtains

$$B(z) = \frac{1}{4}(1 + z - \sqrt{1 - 6z + z^2}).$$

Formulas like (1) allow among other things to study the asymtotic behavior of B_n. This is governed by the singularities of the generating function $B(z)$, according to a famous theorem of Darboux.
Suppose we consider the class of mathematical expressions involving constants, the vaiable x, e^x and additions or products of similar type of expressions. We can visualize the set of such expressions by :

$$\mathcal{E} = \{c\} \cup \{x\} \cup \left\{ \begin{array}{cc} & + \\ / & \quad \backslash \\ \varepsilon & \quad \varepsilon \end{array} \right\} \cup \left\{ \begin{array}{cc} & \times \\ / & \quad \backslash \\ \varepsilon & \quad \varepsilon \end{array} \right\} \cup \left\{ \begin{array}{c} \exp \\ | \\ \varepsilon \end{array} \right\}$$

This permits to represent an element of ε as a tree, for instance the expression $x + e^{e^x + x}$ is represented by Figure 6. The size of an expression will be the

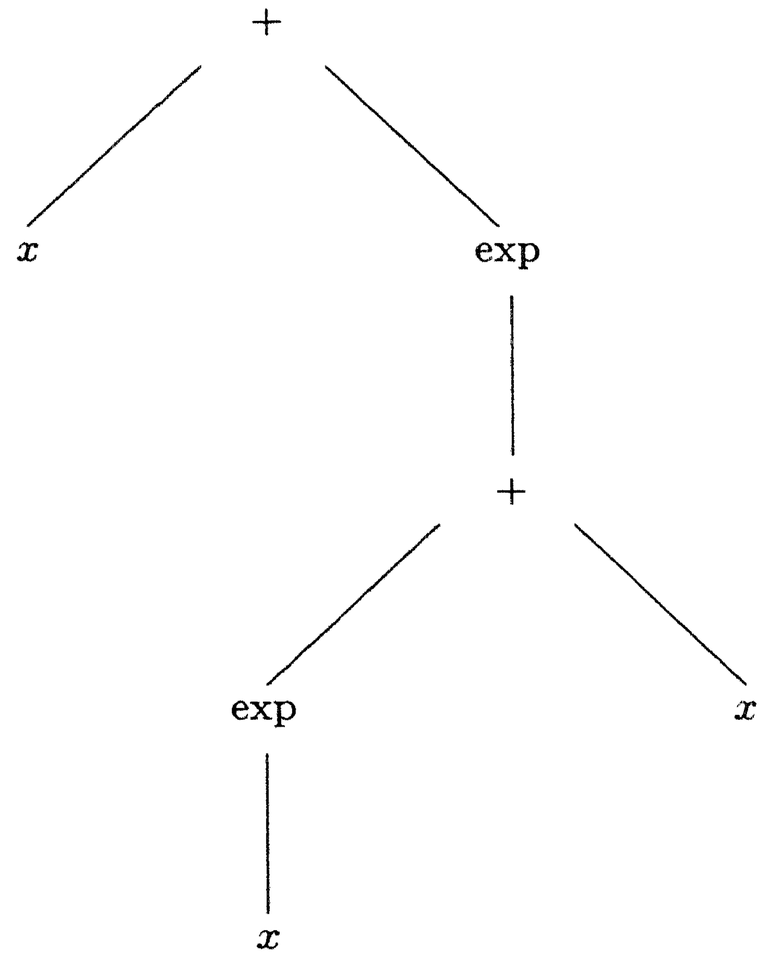

FIGURE 6.

number of nodes of the tree. The above tree has size 7.
Let $E(z)$ to be the generating function corresponding to $\mathcal{E}$, then we have the functional equation

$$E(z) = 2z + 2zE(z)^2 + zE(z)$$

Let $[z^n]E(z)$ to be the coefficient of z^n in the formal series $E(z)$, it represents the number of expressions of size n.
Among possible uses of this machinery, one can compute the complexity of formal differentiation. One can estimate the asymptotic average size of derivatives. Many more applications can be given.

REFERENCES

1. R. ALUR et al, *The algorithmic Analysis of Hybrid Systems*, Theoretical Computer Science, (Jan. 1995)
2. J.P. AUBIN, *Mathematical Methods of Artificial Intelligence*, to be published.

3. L. ALVAREZ, F. GUICHARD, P.L. LIONS, J.M. MOREL, *Axioms and Fundamental Equations of Image Processing*, Arch. Rational Mech. Anal. (1993).
4. D.P. BERTSEKAS, J.N. TSITSIKLIS, *Parallel and Distributed Computation: Numerical Methods*, Prentice Hall, Englewood Cliffs, N.J. 1989.
5. B.M. BOGHOSIAN, C.D. LEVERMORE, *Complex Systems* 1 (1987).
6. J. BRUCK, J. SANZ, *A study on neural networks*, International Journal of intelligent systems, vol. 3, 59-75, (1988).
7. CHIANG T.S., HWANG C.R., SHEU S.J., *Diffusion for global optimization in* $\mathbb{R}^n$, SIAM Control, 25, pp. 737-752, 1987.
8. G. COHEN, P. MOLLER, J.P. QUADRAT, M. VIOT, *Algebraic tools for the performance evaluation of discrete event system*, Proceedings IEEE, special issue of dynamics of discrete event systems, Jan. 1989.
9. I. DAUBECHIES, *Orthonormal Bases of Compactly supported Wavelets*, CPAM, 1988.
10. E. DEAN, R. GLOWINSKI, C.H. LI : *Supercomputer solutions of P.D.E. problems in computational fluid dynamics and in control*, University of Minnesota, Supercomputer Institute.
11. O. FAUGERAS, *A few steps towards artificial 3D Vision*, INRIA, Technical Report series, Fev. 88, N°790.
12. P. FLAJOLET, *Analytic Analysis of Algorithms* Lecture Notes in Computer Science, Vol. 623, Springer Verlag, (1992).
13. U. FRISCH, B. HASSLACHER, Y. POMEAU, *Lattice Gas Automata for the Navier Stokes Equation*, Physical Review Letters, 1986.
14. P.P. KHARGONEKAR, I.R. PETERSEN, M. ROTEA,H_∞ *Optimal Control with State Feedback* IEEE Trans. Automatic Control, 1988.
15. Y.MEYER, *Wavelets and Operators*, Book to appear.
16. K.C. MO and M. GHIL, *Statistics and Dynamics of Persistent Anomalies*, Journal of the Atmospheric Sciences, March 1987, .
17. D. MUNFORD, J. SHAH, *Optimal Approximation by piecewise smooth functions and associated variational problems*, Communications on Pure and Applied Mathematics, 1988.
18. E.S. ORAN, J.P. BORIS,.*Numerical Simulation of Reactive Flow* - Elsevier 1987.
19. E. WONG, *Stochastic neural networks*, ERL, Berkeley, Feb. 89.
20. O. ZEITOUNI, A. DEMBO, *A maximum a Posteriori Estimator for Trajectories of Diffusion Processes*, Stochastics, 1987, Vol. 20.