Chapter 2

Populations. Concurrency. Distributed Algorithms

We pursue our panorama of different models of continuous dynamical systems, by now focusing on models that exhibit several agents, *à priori* in a huge number, and with potentially some competition or concurrency between agents.

We aims at showing that continuous abstractions are natural to talk about populations of individuals, even if those, or their behaviors, are intrinsically discrete (or even stochastic).

Furthermore, we aims at proving that population models, as well as models from (evolutionary) game theory are pertinent and natural to understand today's distributed algorithmic, in particular when the number of involved agents is big, or when the topology is not known except by statistic or probabilistic arguments.

2.1 Bioinformatics

2.1.1 Genetic Regulation Networks

Would be this only by the sequencing of human genome, and the consequent deluge of biological data, or the recent understanding of many biological pathways, it is vital that computer science tools and mathematical tools adapt to face to all these data.

The domain of what is called bioinformatics (sometimes also called computational biology) is born recently, and aims, for example at identifying the genes that are coded by the DNA, or at understanding the primary, secondary or tertiary structure of proteins, or at understanding the ways biological systems operate.

Actually, about this last point, one can say that a new domain of biology, called system biology, is born [Kitano, 2001]. Whereas traditional biology examines genes or proteins in an isolated way, system biology studies in a simultaneous way the interactions between many agents, trying to explain globally the operations involved in biological systems [Kitano, 2001]. Genetic regulation networks, or more generally, interaction biological networks aim at modelling and understanding the interactions between the involved agents in biochemical or biological mechanisms. These agents can be DNA, RNA, proteins, or molecules in biochemical phenomena.

2.1.2 Distributed Algorithms

As this is argued for example in [Alur et al., 2001], genetic circuits and biomolecular networks share many characteristics with distributed computer systems or with today's embedded systems.

Indeed, at the intracellular level, the previous agents communicate in some ways, by influencing one each other, as the nodes of a distributed system communicate. At the intercellular level, some communications also exist.

The description of the state of a system at a given time corresponds to giving characteristics such as the concentrations of each of the species, or the probability to be in a given state. The dynamic is then often described in terms of these quantities by some differential equations [Gibson and Mjolsness, 2000], [de Jong, 2002]. By the presence of non-linearities in most of the considered systems, these dynamics are often subject to discrete transitions between continuous phases. In some way, these transitions can be seen as changes of phases of agents, similar to the changes of states of agents in distributed systems, according to the states of their close neighbors. The considered models are hence close to today's models for distributed systems, in particular to hybrid systems. As this is mentioned for example in [Alur et al., 2001], [Lincoln and Tiwari, 2004], [Batt et al., 2005], this opens the way with the use of formal methods for the proof of properties on these systems.

However, compared to classical problems in formal verification, where the programs of involved agents are fixed, and where one tries to understand the truth of a given property, here the problem is often to understand or to build the programs of the involved agents. In that point of view, formal methods can help, by drawing aside the bad models. This is why symbolic approaches turn out to be very pertinent for computer-aided system analysis: see for example [Batt et al., 2006]. Actually, one faces to problems closer to synthesis of systems than to verification of systems, if one wants to use the terminology of verification of continuous and hybrid systems.

2.1.3 Levels of Descriptions

Concerning the models that are used to describe the dynamics of systems in computational biology, following the classification in [Gibson and Mjolsness, 2000], one can distinguish the following three levels.

1. Discrete models
2. Continuous models
3. Stochastic models

This classification is coarse, and there exist some hybrid intermediate approaches between each of the previous classes [Gibson and Mjolsness, 2000], [de Jong, 2002]. A very nice and clear survey of existing models that are used for modelling genetic regulation networks can be found in [de Jong, 2002].

Going from a class to the next corresponds to an increasing understanding and an increasing level of modelling of the system. Intrinsically systems are stochastic. However, the stochastic approach is often hard to study, and the continuous approach is preferred when stochastic approach is intractable. Discrete approach is considered when the continuous approach is still intractable [Gibson and Mjolsness, 2000], [de Jong, 2002].

Let's focus now on the continuous approach in the sense of previous classification: systems are described by differential equations that describe the evolution of quantities such as concentrations.

Before that, just observe that the stochastic approach also corresponds to differential equations: the stochastic approach often corresponds (see [Gibson and Mjolsness, 2000], [de Jong, 2002]) to using the technique of the master equation [Gillespie, 1977], that is to say to talking of probability laws, and then in describing the evolution of these probability laws by differential equations.

2.1.4 Continuous Models

If we put aside the (difficult) question on how our friends biologists or biochemists succeed to determine the network of logical interactions between all these agents, and the highly non-trivial problem of the estimation of all the quantitative parameters involved, to an interaction regulation network can be associated a resulting dynamic over agents in a rather simple way.

For example, from the interaction network of Figure 2.1 (example taken from [de Jong, 2002], adapted from [Goodwin, 1963], [Goodwin, 1965]), one can associate a dynamic given by

$$\left\{ \begin{array}{lcl} x_1' & = & k_1 r(x_3) - \gamma_1 x_1 \\ x_2' & = & k_2 x_1 - \gamma_2 x_2 \\ x_3' & = & k_3 x_2 - \gamma_3 x_3 \end{array} \right. \tag{2.1}$$

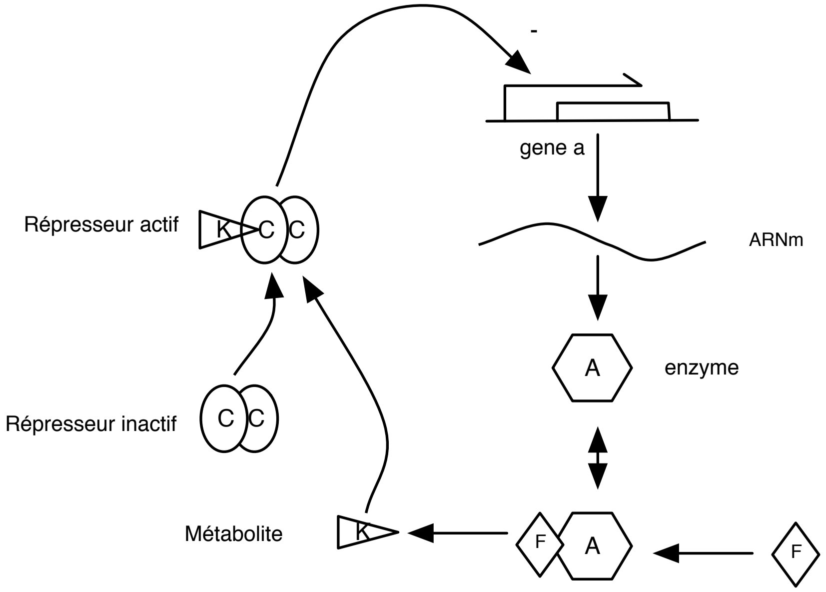

Figure 2.1: An example of a genetic regulation network that involves a terminal product inhibition. Example taken from[de Jong, 2002], adapted from [Goodwin, 1963], [Goodwin, 1965].

where x_1, x_2, x_3 represent respectively the mARN concentrations a, the protein A and metabolite K. Here $k_1, k_2, k_3, \gamma_1, \gamma_2, \gamma_3$ are some constants, $r : \mathbb{R} \to \mathbb{R}$ is a decreasing non-linear function with values between 0 and 1.

The transformation from the description of the network to the dynamic is not the true problematic part. In some way, the description of the network is a "program" of a (distributed) dynamical system, and the "compilation" towards a dynamical system is not the hard part.

One true difficulty is to describe the network, i.e. to determine the "program". This is the major problem for our friends biologists. For computer scientists, one major problem is to do something from the obtained dynamical system. Indeed, when the number of agents becomes huge, the obtained dynamical system becomes intractable. Indeed, the number of variables is simply too big, and the dynamic too complex, the number of indeterminates on involved constants is too high, to really be able to do any valuable simulation, or qualitative or quantitative analysis.

This is one of the reasons behind all the approaches to simplify dynamics, such as for example the replacement of non-linear functions by piecewise linear functions [Lincoln and Tiwari, 2004], or replacement of dynamics by piecewise affine dynamics [Ghosh and Tomlin, 2001], possibly symbolic [Batt et al., 2006], or the use of simpler models than continuous time dynamics [de Jong, 2002], ...

This is fascinating to see how these models of genetic regulation networks correspond to logical descriptions of systems intrinsically continuous, and that, maybe one day, it will be possible to go from the step of understanding the functions to their use for the programming of biological functions.

2.2 Population Models

Population models of biology of populations aim at describing their dynamics. We do a quick review of some of the models. This review is based mainly on [Murray, 2002], with some considerations from [Hirsch et al., 2003] for the case of a unique species.

2.2.1 One Species Models

Before talking of several species in interactions, we think this is important to understand how a population with a unique species naturally growths.

The simplest model consists in supposing that the birth rate of population (i.e. x'/x) is constant. One gets the equation

$$x' = x\lambda. \tag{2.2}$$

This is the model proposed by Malthus in 1798. However, as soon as $\lambda > 0$, the population growths without limits, which is not very reasonable.

It is then more natural to suppose that the environment has a limited capacity N, and that when the number of individuals is higher than this limit capacity, mortality overrides birthrate.

This is the continuous version of the *logistic model* of Verhulst in 1838 (recall that its discrete version has been discussed in previous chapter).

$$x' = \lambda x(1 - x/N). \tag{2.3}$$

In this model, the population converges asymptotically towards N. Refer to [Murray, 2002] for experimental proofs of the validity of this model in relations with true data coming from biology.

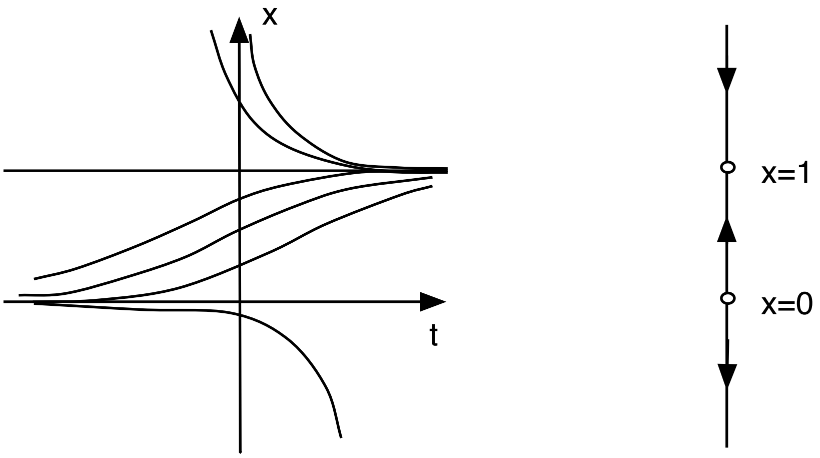

Figure 2.2: On left: some trajectories of $x' = x(1-x)$. On right: phase line.

2.2.2 Constant Harvesting

Fix $N = 1$. Suppose that an adversary (say a hunter) kills individuals with a rate h. We obtain the equation

$$x' = x(1 - x) - h. \tag{2.4}$$

As observed in [Hirsch et al., 2003], a bifurcation phenomenon arises: for $0 < h < 1/4$, if the population is under some threshold value, then it is not sufficiently numerous to reproduce. As soon as the population is above this threshold, its converges toward some limit population. For $1/4 \le h$, the population disappears ultimately.

Hence, we observe an abrupt change, at $h = 1/4$, between a dynamic where the population subsists, to a dynamic with a complete ultimate disappearance of the population. This type of phenomenon is often

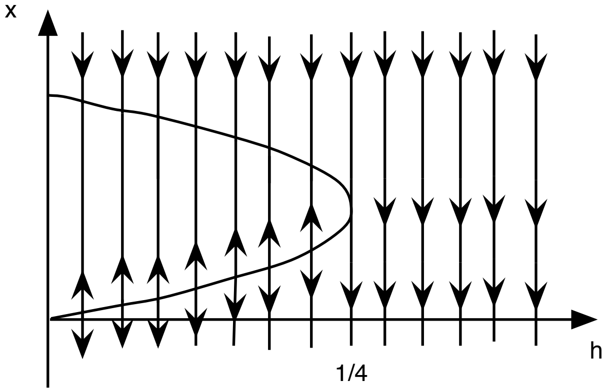

Figure 2.3: Bifurcation diagram of $x' = x(1-x) - h$.

discussed for modelling the sudden disappearance of species, like the ecologic catastrophe that happened when the Nil perche was introduced in lake Victoria in 1960th [Murray, 2002]. Refer to [Murray, 2002] for more examples that could be explained by similar phenomena.

2.2.3 Periodic Harvesting

Let's continue the modelling exercise, as in [Hirsch et al., 2003]. It is reasonable to suppose that hunting is a seasonal process. The model becomes for example

$$x' = x(1-x) - h(1 + \sin(2\pi t)). \tag{2.5}$$

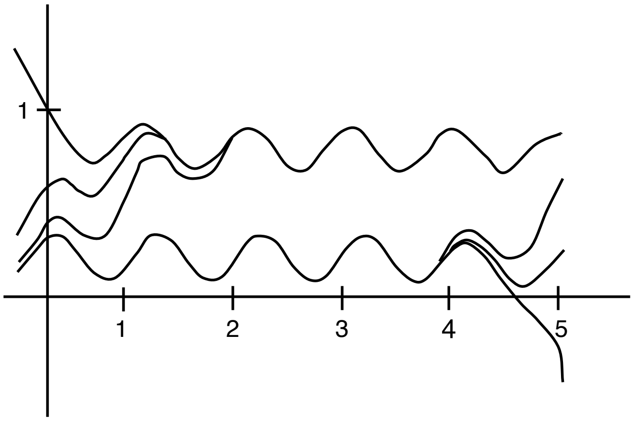

Figure 2.4: Some solutions of $x' = 5x(1-x) - 0.8(1 + \sin(2\pi t))$.

The system still has a bifurcation phenomenon with a sudden disappearance of the population for h bigger than some threshold. Below this threshold, there is apparition of closed periodic trajectories, with an attracting periodic trajectory, and a repulsive periodic trajectory [Hirsch et al., 2003]: see Figure 2.4.

In other words, the presence of seasons introduces oscillations.

2.2.4 Spatial Phenomena

Suppose now that we want to model the fact that the population is not homogeneous, and moves naturally.

Following classical arguments, recalled for example in the excellent book [Murray, 2002], this corresponds to replacing the previous ordinary differential equation $u' = f(u)$ by a partial derivative equation with a diffusion term, so that it becomes

$$\frac{\delta u}{\delta t} = f(u) + D\frac{\delta^2 u}{\delta u^2}. \tag{2.6}$$

The system now has some solutions that correspond to waves that propagate: the population is moving at some speed. Refer to [Murray, 2002], [Murray, 2003] for very interesting discussions on spatial phenomena, with their experimental counterparts on true biological data.

2.2.5 Models with several species

Let's forget now spatial or seasonal effects, and hunters, and suppose that there are two types of species: a population of predator, denoted by y, and a population of prey, denoted by x. Suppose that the population of prey is the only available food for predators. Suppose that in the absence of predators, i.e. when $y = 0$, preys increase accordingly to $x' = ax$. When predators are present, the prey population decreases with a rate that is proportional to the number of preys that they meet. The simplest model for that is a decrease of the birth rate with a factor $-by$. Let's make the opposite hypotheses for predators. In the absence of preys, they decrease with a rate $-c$. In presence of preys, they increase with a rate that is proportional to the number of predator-prey.

We obtain the following system.

$$\left\{\begin{array}{lcl} x' & = & x(a - by) \\ y' & = & y(-cx + dx). \end{array}\right. \tag{2.7}$$

This is precisely the model proposed by Volterra [Volterra, 1931], already discussed. The fact that it corresponds to Lotka model [Lotka, 1920], which is motivated by chemistry, is not surprising: discussing about a population of individuals, or molecules, is rather equivalent at the level of models. We will come back to this issue.

The system has closed trajectories, which correspond to level curves of some function H.

More generally, the population dynamics in the general case of two populations, give rise to dynamics of type

$$\left\{\begin{array}{lcl} x' & = & M(x, y)x \\ y' & = & N(x, y)y. \end{array}\right. \tag{2.8}$$

where M and N are functions of the two variables.

We let the reader guess the case of n populations in interactions. For example, the n-dimensional Lotka-Volterra equation is a system of equations of type

$$x_i' = x_i(b_i - \sum_j a_{i,j} N_j). \tag{2.9}$$

2.3 Models from Virology

2.3.1 In biology

The models of diffusions of epidemics are also particular population models.

For example, the model SIR divides the population of individuals in three types of individuals. The individual S, that are susceptible to be infected, I for infected, and R for cured. It is called SIR since a susceptible individual can become infected, and an infected can become cured. The model $SIRS$ is on the same principle, except that we add the fact that a cured can become again susceptible; And so on for the different acronyms of models like SIS, where a susceptible can become cured, and a cured susceptible.

An overview of all the models can be found in the survey [Hethcote, 2000], or in monograph [Murray, 2002].

If we suppose that we are facing to an illness like malaria, where people that are cured cannot become again ill, we are in the case of the SIR model, that is:

$$\left\{\begin{array}{lcl} S' & = & -\beta SI \\ I' & = & \beta SI - \nu I \\ R' & = & \nu I. \end{array}\right. \tag{2.10}$$

where S, I, R denote the number of individuals in each population, with $S' + I' + R' = 0$, i.e. a constant population.

This way of formalizing with acronyms S, I, R is sometimes called the Kermack-McKendrick model. The model SIR was historically proposed to explain the diffusion of pledge in London in 1665-1666, in Bombay in 1906, and of cholera in London in 1865.

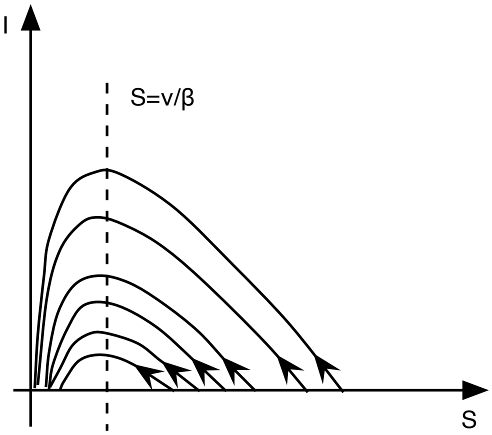

Figure 2.5: Phase portrait of SIR model

A simple mathematical analysis of the model shows that the important parameter is the *epidemic threshold* denoted by $R = \beta S/\gamma$. If this threshold is strictly less than 1, the infection will ultimately disappear. If it is strictly greater than 1, the infection will propagate without limits.

If we consider illnesses where one can become ill again, one obtains the $SIRS$ model, whose dynamic is similar by introducing new constants modelling the rate of passages from state R to state S: see [Murray, 2002], [Hethcote, 2000].

More complex illnesses, such as AIDS, where there are states like seropositivity, lead to more complex models, with a higher number of classes, but still according to same principles: see [Murray, 2002], [Hethcote, 2000].

2.3.2 In Computer Science

In computer science, the modelling of diffusions of computer viruses is essentially based on same models, with the difference that this is necessary to take into consideration the dynamic resulting from protection again the infection, or the superinfection phenomena (saturation of the system by the virus itself).

Among first models, there is [Kephart and White, 1991], that illustrates well that to talk about propagation of viruses, as in classical population models, the principle is to abstract individuals to talk about proportions, or number of individuals.

This article, that takes into account the topology of the network by considering it probabilistic according to a certain law, shows clearly that assuming an underlying graph probabilistic, leads to abstracting the interactions into probabilities that an individual of a given type infects an individual of another type. Finally, the article proposes a model that corresponds to the SIS model in biology, when graphs are of constant mean degree. Simulations are done to understand what happens with families of graphs closer to that that are observed in practice over networks.

The model SIS is rather criticized in computer virology, since once infected, and an antivirus administrated, a program cannot become again ill of the same virus. The model SIR above is so more pertinent. The article [Staniford et al., 2002] thus models the diffusion of red code I by a similar continuous time logistic model.

The model is then extended in [Zou et al., 2002] to introduce a two factors model, to model the fact that the infection rate β decreases in reality with time, since the number of counter-measures, or of protections, has the consequence that people have the tendency to protect with time. This corresponds somehow to consider a model $SIRQ$, where Q denotes a new class of individuals, with a dynamic of type

$$\left\{\begin{array}{lcl} S' & = & -\beta(t)SI - \mu SJ \\ I' & = & \beta(t)SI - \nu I \\ R' & = & \nu I \\ Q' & = & \mu SJ \\ \beta(t) & = & \beta_0(1 - I/(S+I+R+Q))^{\alpha}. \end{array}\right. \tag{2.11}$$

Taking into account the hierarchical structure of the Internet into autonomous systems (AS), the authors of [Serazzi and Zanero, 2003] propose a compartmented model. If we denote by a_i the proportion of hosts that are infected in AS number i, and by N_i the number of hosts that are susceptible in this AS, we obtain a dynamic of type

$$a_i' = (\sum_{j=1}^{n} N_j a_j)(1 - a_i)K/N, \tag{2.12}$$

where K is a constant.

Actually, in this model, if we define $a = (\sum_{j=1}^{n} N_j a_j)/N$, we obtain

$$\left\{\begin{array}{lcl} a' & = & aK(1-a) \\ a_i' & = & a(1-a_i)K, \end{array}\right. \tag{2.13}$$

and so a follows a logistic model. Each a_i has an independent dynamic that converges towards limit value 1 of a.

The model is then extended in [Serazzi and Zanero, 2003] to take into account the saturation of the bandwith of the network, when the contamination occurs.

Let's cite some alternatives. The article [Zou et al., 2003] studies a probabilistic model of diffusion of viruses by emails, based on laws that are observed in practice on the topology of underlying networks, on the time to read an email, and so on.... The study is more based on simulations than on analytic methods, because of the complexity of the model. The article [Chen et al., 2003] presents a discrete model of the diffusion of worms. Here, once again, the study is more relevant to simulations than to analytic methods. Furthermore, several hypotheses under the model, in particular the interest of discretizing are discussed [Serazzi and Zanero, 2003].

2.3.3 Spatial Phenomena

In all previous models, the modelling of spatial phenomena leads to add a diffusion term to each of the equation dynamics, following classical arguments: see [Murray, 2002], [Murray, 2003]. This is true in all the previous models: population models from biology, or from virology, even if, as far as we know, this has not yet been done explicitly in the domain of computer virology.

Still here, this is very instructive to observe the resulting phenomena of infection waves that propagate, both in models and in experimental data: see [Murray, 2002], [Murray, 2003] for biology. Notice that volume [Murray, 2003] is completely devoted to spatial phenomena in biology.

2.4 Models from Game Theory

2.4.1 Introduction to Game Theory

Game theory is one of the ways to model the situations of competitions. It aims at predicting in which situation(s) will be a set of rational players in a situation of concurrence. Refer to the classical reference [Osbourne and Rubinstein, 1994] for a presentation in English.

Some of its ideas come from 18th century, but the main developments of the theory started in years 1920th with the work of Emile Borel and John von Neuman. A decisive event was the publication of the book [von Neumann and Morgenstern, 1944] in 1944 by John von Neumann and Oskar Morgenstern, that fixed the terminology and the presentation of problems that are still in use today. At that time, the attention was mainly about null-sum games, with two players. Some attention was one cooperative games, but non-cooperative games were mostly ignored.

In the 1950th, John Nash proved in [Nash, 1950] that any game has a mixed equilibrium situation, called *Nash equilibrium*, in which no rational player has a unilateral interest to deviate. This major concept for games has from then started to be a central point of all its developments. The attention now is mainly on non-cooperative games.

Game theory begun in these years to be used in economy and in political sciences, but this is only in the 1970th that it really raises a revolution in economy. It has also numerous applications in biology, sociology and in psychology.

The economists are among the largest users of its concepts. The recognition of its interest for the economy was attested by the attribution of several Nobel prices: Nobel price of economy in 1994 to John Nash, John Harsanyi and Reinhard Selten; and in 2005 to Thomas Schelling and Robert Aumann. Another proof of its recognition was the enrolment of famous game theorists for the conception of the auctions for attributing the frequencies of the electromagnetic spectrum in USA in the 1990th: see [McMillan, 1994].

Algorithmic game theory can be distinguished from this mathematical theory, whose constructions are sometimes only existential, by the taking into account of algorithmic and complexity aspects. Its domains of applications are wide: for example, it aims at understanding the complexity of problems related to the computation of equilibria [McKelvey and McLennan, 1996], the loose of performance that results from selfish behaviors in distributed algorithms: see [Papadimitriou, 2001], [Anshelevich et al., 2003] and [Correa et al., 2004], the conception of incitative mechanisms in algorithmic [Nisan and Ronen, 1999], problems related to the pricing of services in some protocols [Feigenbaum et al., 2001], problems related to auction theory [Briest et al., 2005].

2.4.2 Basic Concepts of Game Theory

Let's present the simplest concepts from Game Theory. We focus on non-cooperative games, with complete information, in extensive form.

The simplest game is a two player games, called I and II, with a finite set of options, called *pure strategies*, $Strat(I)$ and $Strat(II)$. Denote by $a_{i,j}$ (respectively: $b_{i,j}$) the score (or if it is a random variable its expected value) for player I (resp. II) when I uses strategy $i \in Strat(I)$ and II strategy $j \in Strat(II)$.

The scores are given by $n \times m$ matrices A and B, where n and m are the cardinality of $Strat(I)$ and $Strat(II)$.

Example 1 (Prisonner dilemma) *The case where A and B are the following matrices*

$$A = \begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix}, B = \begin{pmatrix} 3 & 5 \\ 0 & 1 \end{pmatrix}$$

is called the game of prisoners*, or* prisoner dilemma*. We denote by C (for cooperation) the first pure strategy, and by D (for defection) the second pure strategy of each player.*

A *mixed strategy* of player I, which consists in using $i \in Strat(I)$ with probability x_i, will be denoted by vector $\mathbf{x} = (x_1, \dots, x_n)^T$. One must have $\sum_{i=1}^{n} x_i = 1$, i.e. $\mathbf{x} \in S_n$, where S_n is the unit simplex of $\mathbb{R}^n$, generated by the vectors $\mathbf{e}_i$ of the unit standard basis of $\mathbb{R}^n$.

In a similar way, a mixed strategy for II corresponds to $\mathbf{y} = (y_1, \dots, y_m)^T$ with $\mathbf{y} \in S_m$.

If player II uses mixed strategy $\mathbf{x}$, and player II mixed strategy $\mathbf{y}$, then the first has mean score $\mathbf{x}^T A\mathbf{y}$ and the second $\mathbf{x}^T B\mathbf{y}$.

A strategy $\mathbf{x} \in S_n$ is said to be a best response to strategy $\mathbf{y} \in S_m$, denoted by $\mathbf{x} \in BR(\mathbf{y})$ if

$$\mathbf{z}^T A\mathbf{y} \leq \mathbf{x}^T Ay \tag{2.14}$$

for all strategy $\mathbf{z} \in S_n$.

A pair $(\mathbf{x}, \mathbf{y})$ is a *mixed Nash equilibrium* if $\mathbf{x} \in BR(\mathbf{y})$ and $\mathbf{y} \in BR(\mathbf{x})$. Nash theorem [Nash, 1950] claims, by a fixed point argument, that such an equilibrium always exist. However, it is not necessarily unique.

In other words, two strategies $(\mathbf{x}, \mathbf{y})$ form a Nash equilibrium if in that state none of the player has a unilateral interest to deviate from it.

A Nash equilibrium in pure strategies, that is to say with $\mathbf{x} \in \{0,1\}^n$, $\mathbf{y} \in \{0,1\}^m$, does not always exist.

Example 2 *On the example of the games of prisoners, $BR(\mathbf{y}) = (0,1)$ for all $\mathbf{y}$, and $BR(\mathbf{x}) = (0,1)$ for all $\mathbf{x}$. So (D,D) is the unique Nash equilibrium, and it is pure. In it, each player has score 1. The paradox is that if they had played (C,C) (cooperation) they would have had score 3, that is more. The social optimal (C,C), is different from the equilibrium that is reached by rational players (D,D), since in any other state, each players fears that the adversary plays C.*

This is important to understand that game theory is not at all a theory that aims at talking about dynamism, but a theory that aims at talking about equilibrium in presence of rational players. Actually, one strongly assumes that each game is played once and only once, and the theory predicts what should be played. This is true that this is quite diverting to interpret what means a mixed strategy, that is to say probabilities, when one plays only once: see the interesting discussion in the beginning of book [Osbourne and Rubinstein, 1994].

To introduce some dynamism in game theory, there are two main approaches. The first consists in repeating games. The second in using models from evolutionary game theory.

2.4.3 Repeated Games

Repeating k times a game, is equivalent to extend the space of choices into $Strat(I)^k$ and $Strat(II)^k$: player I (respectively II) chooses its action $\mathbf{x}(t) \in Strat(I)$, (resp. $\mathbf{y}(t) \in Strat(II)$) at time t for $t = 1, 2, \cdots, k$. Hence, this is equivalent to a two-players game with respectively n^k and m^k choices for players.

Following [Binmore, 1999], to avoid confusions, we will call *actions* the choices $\mathbf{x}(t), \mathbf{y}(t)$ of each player at a given time, and *strategies* the sequences $X = \mathbf{x}(1), \cdots, \mathbf{x}(k)$ and $Y = \mathbf{y}(1), \cdots, \mathbf{y}(k)$, that is to say the strategies for the global game.

If the game is repeated an infinite number of times, a strategy becomes a function from integers to the set of actions, and the game is still equivalent to a two-players game[1].

The score of a player is usually evaluated with a discount rate $\delta > 0$ (that can be chosen to be 1, if k is finite; choose $\delta < 1$ if k is infinite). Concretely, the score of player I (respectively II) with strategies X,Y is counted as

$$Gain(X, Y) = \sum_{i=1}^{k} \delta^{i-1} gain(i), \tag{2.15}$$

---

[1] but whose matrices are infinite.

where $gain(t)$ is the score of the player at time t, that is to say $\mathbf{x}(t)^T A\mathbf{y}(t)$ for player I (resp. $\mathbf{x}(t)^T B\mathbf{y}(t)$ for player II).

Repeating a finite and fixed number of times a game does not change anything to expected behaviors. For example, a simple backward reasoning shows that each player has still interest to play D at time $n, n-1, \cdots, 1$ in the repeated prisoner lemma, if the game is played a number of times k, that is fixed and known by each players: see [Binmore, 1999].

This is more interesting to repeat the game an infinite number of times. Now, the strategy of systematic cooperation (that is to say $\mathbf{x}(t) = \mathbf{y}(t) = (1, 0)^T = C$ for all t) becomes a Nash equilibrium for the repeated game: in other words, cooperation becomes interesting for each player.

Observe that the score that is given by equation 2.15 for the repeated game has another simple interpretation: if one supposes that the game is repeated a number of times which is random, with at each step a probability δ that the game continues at least one more step, and a probability $1 - \delta$ that this is the last step, then the probability that the ith step is played is δ^i. In expectation, the score of each player is hence given by equation 2.15: see [Binmore, 1999].

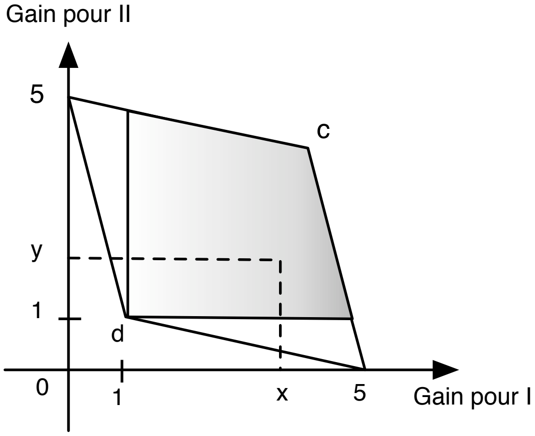

Figure 2.6: For any point (x, y) of the greyed zone, there is a pair of strategies X, Y that guaranties the score x for player I and score y for player II, and such that (X, Y) is a Nash equilibrium for the repeated game. In particular, for the point c that corresponds to systematic cooperation, and for the point d that corresponds to systematic defection.

The problem is that the systematic cooperation strategy is not the only Nash equilibrium. The systematic defection strategy is still a Nash equilibrium of the repeated game. Actually, one shows that for any point (x, y) of the greyed zone on Figure 2.6, there is a pair of strategies (X, Y) that guaranties score x for player I and score y for player II, and such that (X, Y) is a Nash equilibrium for the repeated game: see [Binmore, 1999]. In other words, the number of Nash equilibrium for the repeated games is infinite, and even uncountable.

2.4.4 Behaviors

In practice, player I (respectively II) has to solve the following problem at each time t: given the history of the game up to now, that is to say $X_{t-1} = \mathbf{x}(1), \cdots, \mathbf{x}(t-1)$ and $Y_{t-1} = \mathbf{y}(1), \cdots, \mathbf{y}(t-1)$ what should I play at time t? That is to say how to choose $\mathbf{x}(t) \in Strat(I)$? (resp. $\mathbf{y}(t) \in Strat(II)$?)

This is natural to suppose that the answer of each of the players is given by some behavior rules: $\mathbf{x}(t) = f(X_{t-1}, Y_{t-1})$, $\mathbf{y}(t) = g(X_{t-1}, Y_{t-1})$ where f (respectively g) is a function that determines the behavior of I (resp. II).

The question of the best behavior rule to use for the prisoner lemma gave birth to an important literature. In particular, after the book [Axelrod, 1984], that describes the results of tournaments of behavior rules for the iterated prisoner lemma, and that argues that there exists a best behavior rule called $TIT-FOR-TAT$. The behavior $TIT-FOR-TAT$ consists in cooperating at the first step, and then do the same thing as the adversary at following times. Refer to [Beaufils, 2000], [Labbani, 2003] for critical studies of these results. A lot of other behaviors, most of them with very picturesque names have been proposed and studied: see for example [Axelrod, 1984], [Beaufils, 2000], [Labbani, 2003].

2.5 Dynamics Coming from Games

2.5.1 Games on a Graph

An example of behavior is $PAVLOV$.

Example 3 (Behavior $PAVLOV$) *The $PAVLOV$ behavior consists, in the iterated prisoner lemma, in fixing a threshold, say 3, and at time t, replaying the previous pure action if the last score is above this threshold, and changing the action otherwise.*

Concretely, if we denote $+$ for C, and $-$ for D, one checks easily that this corresponds to rules

$$\left\{\begin{array}{lll} ++ & \rightarrow & ++ \\ +- & \rightarrow & -- \\ -+ & \rightarrow & -- \\ -- & \rightarrow & ++, \end{array}\right. \tag{2.16}$$

where the left hand side of each rule denotes $\mathbf{x}(t-1)\mathbf{y}(t-1)$, and the right hand side the corresponding result for $\mathbf{x}(t)\mathbf{y}(t)$.

$PAVLOV$ behavior is Markovian: a behavior f is *Markovian*, if $f(X_{t-1}, Y_{t-1})$ depends only on $\mathbf{x}(t-1), \mathbf{y}(t-1)$.

From a set of such rules, this is easy to obtain a distributed dynamic. For example, let's follow [Dyer et al., 2002].

Suppose that we have a connected graph $G = (V, E)$, with N vertices. The vertices correspond to players. An instantaneous configuration of the system is given by an element of $\{+,-\}^N$, that is to say by the state $+$ of $-$ of each vertex. Hence, there are 2^N configurations.

At each time t, one chooses randomly and uniformly one edge (i, j) of the graph. At this moment, players i and j play the prisoner dilemma with the $PAVLOV$ behavior, that is to say the rules of the equation 2.16 are applied.

What is the final state reached by the system?

The underlying model is a huge Markov chain with 2^N states. The state $E^* = \{+\}^N$ is absorbing. If the graph G does not have any isolated vertex, this is the unique absorbing state, and there exists a sequence of transformations that transforms any state E into this state E^*. As a consequence, from well-known classical results for Markov chains, whatever the initial configuration is, with probability 1, the system will be in state E^*: see [Brémaud, 2001]. The system is *self-stabilizing*.

Several results about the convergence time towards this stable state can be found in [Dyer et al., 2002], and [Fribourg et al., 2004], for rings, and complete graphs.

What is interesting in this example is that it shows how to go from a game, and a type of behaviors to a distributed dynamic on a graph. Clearly this is easy to associate a similar dynamic to any[2] Markovian behavior on a symmetric game.

An interesting question is the following:

Question 1 *Which distributed dynamics correspond to a game?*

---

[2] But not necessarily Pavlovian. Actually, the behavior $PAVLOV$, as described here, is not ambiguous only on 2 by 2 matrices.

We think the question interesting not only for this way of associating a distributed dynamic to a game, but for the many natural ways to associate a dynamic to a game.

There are indeed other means to associate a dynamic to a game. We now review some of them.

2.5.2 Spatial Games

Let's first discuss some particular distributed games on the grid that are well-known and often cited.

The articles [May et al., 1995] and [Nowak and May, 1992] discuss the iterated prisoner lemma on a bidimensional grid. At each time, each player plays with its immediate neighbors at the prisoner dilemma. The score of a player is the sum of the scores of its games with its neighbors. At the beginning of the next generation, each player chooses the action of its neighbor with the best score among itself and its neighbors.

These articles focus on the experimental evolution of cooperation in this game, and on phenomena like spatial patterns, or chaos, that arise in this game.

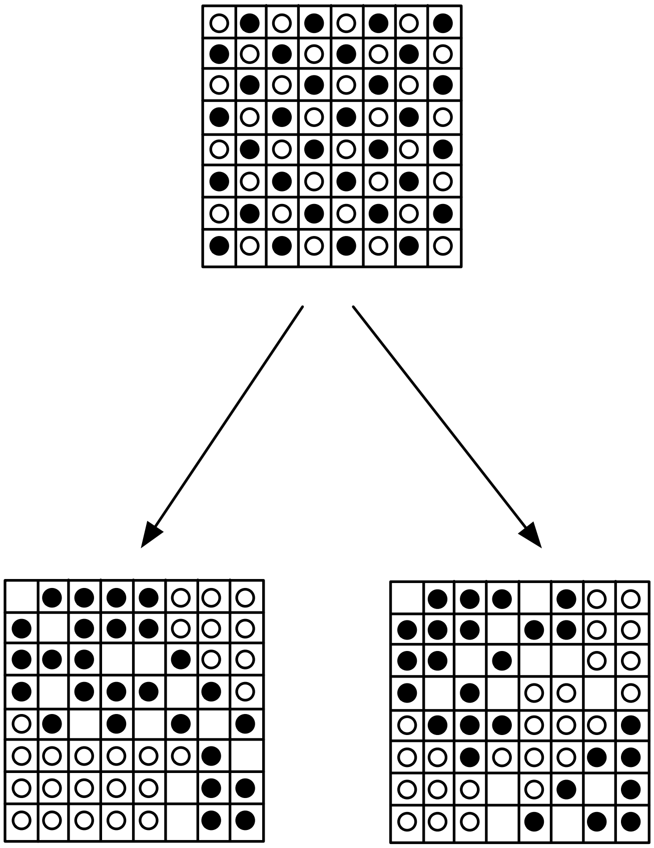

Figure 2.7: Schelling's game

Other spatial games, with a clear pedagogical value, are the games of Schelling, presented by their author in [Schelling, 1978], [Schelling, 1971a]. We will follow here the presentation from [Binmore, 1999] of the solitaire's game of Schelling.

It aims at being a (simplistic) model of racial segregation, and the game is player on a chessboard. Each box of the chessboard represents a house. Each house can contain at most one pawn. Each pawn is either black or white. Each blank pawn wants that at least half of his neighbors are white. Each black pawn wants that at least one third of his neighbors are black. At each generation, a black or white pawn that has his constraint not satisfied moves to the closest point where its constraint is. The order in which the unhappy pawns are moved is not important, and we can suppose it random. The process continues until all pawns have their constraint satisfied.

If one starts with configurations as the one on the top of Figure 2.7, by simulation, one ends with configurations that are typically as in the bottom of Figure 2.7, that is to say with a strong racial segregation. In other words, the individual choices of individuals, that are not necessarily the manifestation of a manifest racial segregation, lead to global situations with strong racial segregations. The social equilibrium that results from individual preferences is not the one that each would globally prefer.

Refer to articles [Schelling, 1971a], [Schelling, 1978], [Schelling, 1969], [Schelling, 1971b] for the study of many variants on different parameters, and to the important literature that cite these articles for discussions on the social implications or philosophical implications of such phenomena.

2.5.3 Shortsighted Dynamic

In the general case, to every 2-player repeated game, one can associate the *shortsighted* behavior. It consists in the fact that each player makes systematically the hypothesis that the opposite player will replay at time t the same thing as he played at time $t-1$. As a consequence, this behavior consists in choosing systematically at time t the (or a) best response to the action of the opposite player at time $t-1$:

$$f(X_{t-1}, Y_{t-1}) \in BR(\mathbf{y}(t-1)).$$

Take, like [Binmore, 1999], the example of the Cournot duopoly game. The Cournot duopoly game is a well-known economical model of the competition of two producers of a same good. In this model, the production of a unit article of this good costs c. One makes the hypothesis that the total demand is of the form $q = q_1 + q_2 = M - p$, where p is the sold price, and q_1 and q_2 the number of produced articles by each of the firms.

The problem of firm I (respectively II) is to fix q_1 (resp. q_2) in order to maximize its profit $(p-c)q_1$ (resp. $(p-c)q_2$). One shows easily (see [Binmore, 1999]), that the best response to q_2 is to choose

$$q_1 = 1/2(M - c - q_2), \tag{2.17}$$

and that the best response to q_1 is to choose

$$q_2 = 1/2(M - c - q_1), \tag{2.18}$$

so that the unique Nash equilibrium corresponds to the intersection of the two lines defined by equations (2.17) and (2.18).

The shortsighted dynamic for the two players then gives on this game

$$\left\{ \begin{array}{rcl} q_1(t) & = & 1/2(M - c - q_2(t-1)) \\ q_2(t) & = & 1/2(M - c - q_1(t-1)). \end{array} \right.$$

This is easy to show that whatever the initial point is, such a dynamic converges towards the Nash equilibrium. The collective dynamic converges towards the rational equilibrium.

Unfortunately, as shown in [Binmore, 1999], this is not always the case: inverting for example the curves of reactions of players in this game, leads to a dynamical system that converges for some initial cases to some of the Nash equilibria, but for others, that doesn't converge at all (it can oscillate).

2.5.4 Fictious Player Dynamic

The shortsighted behavior can be considered as very too basic. A more reasonable behavior seems to be the following: to predict what will play the opposite player at time t, let's use the statistic of what he did at time $1, 2, \cdots, t-1$: if he played action i n_i times, let's estimate that he will play action i with probability $x_i = n_i/(t-1)$ at time t. This is what is called the *fictious player dynamic*.

To simplify things, let's follow [Binmore, 1999], and suppose that $n = m = 2$, and that the matrices are given by

$$A = \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}.$$

If at time $1, 2, \cdots, t-1$, player 2 used n_i times action number i, player I will estimate that player II will play at time t action i with probability $y_i(t) = n_i/(t-1)$. Player II will evaluate probability $x_i(t)$ that player I play action i in a symmetric way.

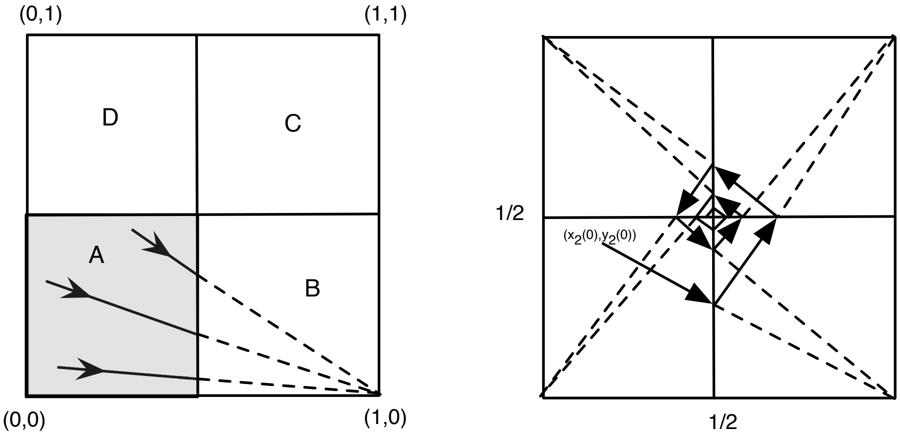

Figure 2.8: Convergence towards a mixed equilibrium.

To study the dynamic, as this is shown in [Binmore, 1999], one just needs to go from discrete time to continuous time.

A simple analysis (see [Binmore, 1999]) shows that as long as $(x_2(t), y_2(t))$ stays in zone A of the left part of Figure 2.8, player I will use its second pure strategy, and player II its first pure strategy as a best response to what he or she expects from the opposite player.

The dynamic $(x_2(t), y_2(t))$ will stay in this zone up to time $t + \tau$ for $\tau > 0$ sufficiently small. Since we know the choice of player II between time t and time $t + \tau$, one can hence evaluate $y_2(t + \tau)$ as

$$y_2(t+\tau) = \frac{t y_2(t)}{t+\tau}. \tag{2.19}$$

This can be written as

$$\frac{y_2(t+\tau) - y_2(t)}{\tau} = -y_2(t).$$

Bet letting τ converge to 0, we obtain

$$y_2'(t) = \frac{y_2(t)}{t}.$$

In a similar way, we obtain

$$x_2'(t) = \frac{1 - x_2(t)}{t}.$$

The points that satisfy these two equations are on a straight line that starts from $(x_2(t), y_2(t))$ and that joins point $(1, 0)$. A similar study on zones B,C, and D of the left part of Figure 2.8 shows that the dynamic must be the one depicted on the right part of Figure 2.8. It converges towards the mixed Nash equilibrium of the game. Once again, the collective dynamic converges towards the rational equilibrium.

Unfortunately, once again, this is not the case for all the games: one can easily consider games where trajectories do not converge, or with limit cycles [Binmore, 1999].

2.5.5 Dynamic by Social Libration

Still following [Binmore, 1999], consider the same game, but suppose that we have a huge population of players. We have two types of players, boys and girls. The boys are player I, while girls are player II. Each player, whatever type it may have, chooses to play a given pure strategy I or II forever. We suppose that at any time, some players disappear: during any period T, the probability that a player leaves the population is λT.

Every player that disappears is replaced by a player whose action is the best response to the global statistical situation of the opposite gender: if in the current population a fraction $y_i(t)$ (respectively $x_i(t)$) of girls (resp. boys) uses action i, then the new boys (resp. girls) that are injected will play the (pure) action which is the best response to the mixed strategy $(y_1(t), y_2(t))$ (resp. $(x_1(t), x_2(t))$). This is indeed the action that will give him (resp. her) the best score if he (resp. she) meets a player of opposite gender at random, and then plays the 2 by 2 game with her (resp. him).

One shows (see [Binmore, 1999]) that we must have in zone A

$$x_2'(t) = -\lambda x_2(t) + \lambda,$$

and

$$y_2'(t) = -\lambda y_2(t).$$

The dynamic is different from the one of previous section, nevertheless the points that satisfy these two equations are still on the straight line that joins point $(x_2(t), y_2(t))$ and point $(1, 0)$.

A similar study on zones B,C, and D of the left part of Figure 2.8 shows that the trajectories are identical to those of the previous section (but not the speed on these trajectories): see [Binmore, 1999].

2.6 Evolutionary Game Theory

Evolutionary game theory is another way to associate dynamics to games.

2.6.1 Biological Motivation

Evolutionary game theory is born from the book from Maynard Smith [Maynard-Smith, 1981]. This book presents some applications of game theory to biology.

To illustrate how, to a game, can be associated a biological dynamic, let's take the fictive example of a population of individuals from [Binmore, 1999]. Binmore chooses to call these individuals *dodos*.

The day of a dodo lasts a fraction τ of a year. There are n types of dodos: the dodos that play action 1, the dodos that play action 2, ..., and the dodos that play action n. Babies of a dodo of type i are always of type i.

We are interested in the proportion $x_i(t)$ of dodos that play action i. We have of course $\sum_{i=1}^{n} x_i(t) = 1$.

At the end of each day, the dodos fight pairwise. The outcome of a fight has some influence on the fecundity of involved participants. One reads on a matrix A $n \times n$ at entry $a_{i,j}$ the birth rate of a given dodo, if it is of type i and if he fights again an individual of type j: his expected number of babies at next day is given by $\tau a_{i,j}$.

How many babies a dodo of type i can expect to have at next day? The answer is

$$\sum_{j=1}^{n} x_j(t) \tau a_{i,j} = (A\mathbf{x})_i \tau.$$

Indeed, since the pairing for the fights on the evening between dodos are chosen at random and uniformly among the population, in expectation its birth rate is given by previous expression.

The number of dodos of type i in the next morning is hence

$$N x_i(t)(1 + (A\mathbf{x})_i \tau).$$

Mortality being not a function of the type of dodos, the next day, the fraction of dodos of type i will be given by

$$x_i(t+\tau) = \frac{N x_i(t)(1+(A\mathbf{x})_i\tau)}{N x_1(t)(1+(A\mathbf{x})_1\tau) + \cdots + N x_n(t)(1+(A\mathbf{x})_n\tau)}.$$

Hence

$$x_i(t+\tau) = \frac{x_i(t)(1+(A\mathbf{x})_i\tau)}{1+\mathbf{x}^T A\mathbf{x}\tau}.$$

where $\mathbf{x}^T A\mathbf{x}\tau$ can be interpreted as the expected number of birth for a dodo in a day.

This can be rewritten as

$$\frac{x_i(t+\tau) - x_i(t)}{\tau} = x_i(t)\frac{(A\mathbf{x})_i - \mathbf{x}^T A\mathbf{x}}{1+\mathbf{x}^T A\mathbf{x}\tau}.$$

By taking limit when τ goes to 0, we obtain

$$x_i' = x_i((A\mathbf{x})_i - \mathbf{x}^t A\mathbf{x}).$$

This is what is called a *replicator equation.* Such an equation models the fact that individuals whose score (fitness) given by matrix A is above mean score (fitness) have tendency to reproduce, whereas those that have a score under the mean score have tendency to disappear.

2.6.2 Objectives of this Theory

Of course the model about dodos is relatively ad hoc, but many situations and models give rise to same dynamics: refer to [Maynard-Smith, 1981] for discussions about that.

Evolutionary game theory aims at studying the behaviors of such dynamics in function of matrix A.

It has its own notions of equilibria, motivated by the stability of underlying dynamical systems, such as the notion of evolutionary stable equilibrium. An evolutionary stable equilibrium is a particular Nash equilibrium. It makes it possible to link the notions of equilibria for the game given by A to the notions of stability for the corresponding dynamical system.

Actually, it does not only consider replicator dynamics, but other dynamics such as imitation dynamics, best response dynamics, and so on. . . , with in all dynamics the idea that individuals with highest score reproduce faster than others.

Refer to [Maynard-Smith, 1981], [Weibull, 1995], [Hofbauer and Sigmund, 2003] for a presentation of its main results.

2.7 Distributed Algorithmic

2.7.1 Classical Models and Machines

We now come to our motivations related to distributed algorithmic.

There are several models in distributed algorithmic, according to the considered type of physical machines and associated constraints (for example the type of communications that are used, for e.g. by messages, by shared variables, . . . , the existence or not of a shared memory, how conflicts between reading and writing operations are solved, . . .).

However, in an abstract way, all these models consist in saying that the set of agents corresponds to a set of particular communicating automata[3] [Lynch, 1997], [Tel, 1994].

In general, a very strong hypothesis in all these models is that a fixed topology is assumed. There is an underlying graph such that communications between agents can only happen between neighbors in the graph.

---

[3] or sometimes Turing machines

In practice, this is true that communications can only happen between physically connected neighbors, or, as in the so-called ad hoc networks, the graph is dynamic, but it changes little locally.

In the evaluation of the complexity of solutions, the number of involved processors is often considered as an important parameter. Indeed, in practice, the number of physically existing processors is limited, and buying new processors has some cost.

Refer to [Lynch, 1997], [Tel, 1994] for a presentation of classical distributed algorithmic, and of its methods, and to [Balcàzar et al., 1990] for a presentation of models seen from the angle of complexity.

2.7.2 Towards a New Distributed Algorithmic

However, we think that this is necessary to call into question some of the classical hypotheses.

First, the size of some networks is now such, that this seems necessary not to talk anymore about individuals, but of statistics. Indeed, for example, when one describes the state of a network such as the Internet, this is not pertinent to describe the state of a given individual, but rather the proportion of individuals in each state. Another example: When one describes as in [Adar and Huberman, 2000] the possible behaviors of free-riding of individuals in a peer-to-peer application for downloading files, the only pertinent way to do it is to talk about proportions of agents with a given behavior.

We think this is time to reason about statistics, and proportions, that is to say with models such as the previous ones that are used in biology, chemistry, physics,... to describe systems with a huge number of agents.

Indeed, several developments of distributed algorithmic, which have been recently considered, call strongly into question some of the classical hypotheses: See for example the discussion in [Angluin et al., 2004] for sensor networks. Indeed, one considers that today's cost of production of a sensor is so low, that the number of involved sensors in an application is not so important. In addition, in some applications, the sensors can happen not to control which are their neighbors: for example, when Paris underground society RATP equips its clients of sensors [Alberganti, 2006], the programs on these sensors have no mean to predict which other sensors they meet.

By a reasoning similar to what is done by the author of [Kephart and White, 1991] when talking about computer viruses, where one doesn't control *à priori* the contacts of viruses with uninfected applications, or by the classical reasoning done in biology, physic to establish all previous dynamic models such as population models, this is equivalent to consider that there is no topology, or at best a statistic on the topology.

2.7.3 From Description to Programming

The models from previous sections make it possible to describe populations in interactions. Of course, that makes it possible to study the dynamic of systems such as the Internet, or the dynamic of users of a given application when the number of users is huge.

However, we think that a true stake is to go from these models as models for describing populations to models used for programming.

For example, as the biologists use their diffusion models for epidemics to anticipate the propagation waves and determine the most economic zones to be treated [Murray, 2003], it seems possible to fight against the diffusion of a computer virus by the anticipated local treatment of some zones, or by the diffusion of some counter-viruses that would use the same weaknesses of the system to propagate.

Question 2 *Can we use the previous dynamical models to "program" systems in concurrence, and not only to describe them?*

We will try to explicit this question through examples below.

If this is possible, this legitimates the following question.

Question 3 *What is the power of models from bioinformatics, population biology, virology, or of evolutionary game theory of previous sections?*

Each of them models systems with some concurrence, or competition. For many of them, this does not seem reasonable to consider that one can control all the individuals, but only a fraction of them. The models from game theory, or from evolutionary game theory then take all their interests, since one wants to program a global behavior from local behaviors in presence of some competition.

2.7.4 Populations Protocols from Angluin et al

This type of questions leads again to understand the power, that is to say what can be computed, by each of the models.

To present a type of models that goes in the sense of our discussion, as well as associated problems, we present the recent *population protocol model* of [Angluin et al., 2004], that comes from algorithmic for sensor networks.

In this model, a protocol consists in giving a finite set of internal states $Q = \{1, 2, \ldots, k\}$, and a transition rule $\delta : Q \times Q \to Q \times Q$. For $\delta(p, q) = (p', q')$, write $\delta_1(p, q) = p'$, $\delta_2(p, q) = q'$.

A configuration of a system at a given time is given by the internal states of each of the n individuals.

We suppose that the individuals are completely indiscernible. It follows that the state of a system can be described by the number n_i of individuals in state i, for $1 \leq i \leq k$, better than by the state of each individual.

At each discrete round, a unique individual i is put in relation with some other individual j: at the end of this meeting, the individual i is in state $\delta_1(q_i, q_j)$, and individual j is in state $\delta_2(q_i, q_j)$.

We suppose that we cannot control the interactions, and that there is some fairness: if in a configuration C one can go to configuration C' in one step (denoted by $C \to C'$) then in any derivation $C_0 C_1 \cdots$, with $C_i \to C_{i+1}$ for all i, if C appears infinitely often, then C' also.

One wants to consider population protocols as predicate recognizers $\psi : \mathbb{N}^m \to \{0, 1\}$.

To do so, fix a subset $Q^+ \subset Q$, and say that an tuple $(n_1, \ldots, n_m) \in \mathbb{N}^m$, for $m \leq k$, is accepted (respectively rejected) by the protocol, if starting from any configuration with n_i individuals in state i, eventually all the individuals will be in some internal state that belongs to Q^+ (resp. its complement), and this stays true at any time afterward.

One says that the protocol recognizes $\psi : \mathbb{N}^m \to \{0, 1\}$ if for all tuple $(n_1, \ldots, n_m)$, it is accepted when $\psi(n_1, \ldots, n_m) = 1$ and it is rejected when $\psi(n_1, \ldots, n_m) = 0$.

We have the following very nice result (recall that the sets that are definable in Presburger arithmetic coincide with the semi-linear sets over the integers).

Theorem 1 ([Angluin et al., 2006b])

- *Any predicate $\psi : \mathbb{N}^m \to \{0, 1\}$ that can be defined in Presburger arithmetic can be computed by a population protocol.*
- *Conversely, any predicate $\psi : \mathbb{N}^m \to \{0, 1\}$ that is computed by a population protocol can be defined in Presburger arithmetic.*

For example, since this is definable in Presburger arithmetic, there is a protocol to decide if more than 5% of agents are in internal state 1.

This theorems shows, if needed, that these models are really different from classical models, such as cellular automata. Indeed, cellular automata are very close to Turing machines: it is very easy to simulate a Turing machine with a cellular automaton. By opposition, the link between Turing machines and this type of models is far from being so direct.

Refer to [Angluin et al., 2005a], [Angluin et al., 2005b], [Angluin et al., 2006a] for recent and other results about this model, and some variants.

2.7.5 Populations Protocols

Previous model is motivated by sensor networks. But, even if it can decide problems talking about proportions, such as "decide if there is more than 5% of agents in internal state 1", the model does not deal with proportions, but with number of individuals in each state.

If the number of individual is high, this is natural not to talk about numbers, but about proportions or statistics.

For example, consider the following protocol: we have a population of n agents. Each agent is either in the state $+$, or in state $-$. A configuration corresponds to a point of $S = \{+, -\}^n$.

We suppose that time is discrete. At each discrete time, (all or a fixed fraction of) the agents interact in pairs, according to the following rules:

$$\begin{array}{lcl} ++ & \rightarrow & 1/2+, 1/2- \\ +- & \rightarrow & + \\ -+ & \rightarrow & + \\ -- & \rightarrow & 1/2+, 1/2- \end{array}$$

One must interpret the second rule in the following way: if an individual is of type $+$, and it interacts with an individual of type $-$, then it becomes of type $+$. One must interpret the first rule in the following way: if an individual is of type $+$ and interacts with an individual of type $+$, he becomes of type $+$ with probability $1/2$, and $-$ with probability $1/2$. And so on for the interpretation of other rules.

We suppose that the pairings are chosen at random uniformly.

Experimentally, the proportion of $+$ in the population converges towards $\sqrt{2}/2$, when the number of individuals increases.

Formally, this could be expected, since, if p denotes the proportion of $+$, with probability p an individual meets a $+$, and $1-p$ a $-$. Now, the first and fourth rule destroy in mean $1/2+$ each, whereas the second and third rules create one $+$ each. By doing the sum, one can write that in expectation, the number of $+$ that are created at each time is

$$1/2p^2 + 2p(1-p) + 1/2(1-p)^2 = 1/2 + p - p^2.$$

Now, at equilibrium, there must be conservation, and so it must be equal to p. Hence

$$p^2 = 1/2.$$

$$p = \sqrt{2}/2.$$

Question 4 *The previous system converges towards $\sqrt{2}/2$. Which numbers are computable by such protocols?*

Of course, by using pairwise pairing, rational probabilities, and a finite number of internal states for each agent.

We conjecture that this corresponds to algebraic numbers of $[0, 1]$, but at this moment, we don't even have the formal proof that previous system is converging when n goes to infinity. When n is supposed infinite (the number of agents is infinite) the previous study is valid and proves that $\sqrt{2}/2$ is computable with an infinite number of agents.

Question 5 *Can we characterize the computable predicates in this model?*

Of course, one says that a predicate is computable if it captures the relation between the initial proportion and the limit proportion.

2.7.6 Why Focusing on Such Protocols?

We already mentioned several reasons why considering such type of systems seems interesting.

To go in the sense of this discussion, and to validate the interest of all these models in this section for distributed algorithmic, here is a nice exercise.

Question 6 *Among all differential equations that describe a dynamic in this chapter, which ones correspond to a dynamic of a distributed algorithm? And which distributed algorithm(s)?*

Distributed algorithms should be understood as population protocols, or simple-to-explain variants of it. We do not consider that time must necessarily be discrete. It can be continuous. For example, the previous system of rules that computes $\sqrt{2}/2$ corresponds to the description of a (huge) discrete-time Markov chain. But, by using pairings that would occur according to a Poisson process, one would also compute $\sqrt{2}/2$. We would now describe a continuous-time Markov chain with same limit. Probably this might be more natural to solve this exercise.

People from distributed algorithmic will probably argue that the hypothesis of abstracting topology in previous models is probably too strong.

We take these few lines to insist. As soon as the global system is such that the local proportions of configurations in such or such state does not depend on the point that it considered, that is to say there is no spatial phenomena, in the sense of previous population models, even if there is a topology, it can often be abstracted in statistical terms.

Observe that if there are spatial phenomena, then it is still possible to use models of dynamics of populations, with its methods, but this would correspond to partial differential equations.

Bibliography

[Adar and Huberman, 2000] Adar, E. and Huberman, B. A. (2000). Free riding on gnutella. *First Monday*, 5(10). *http://firstmonday.org/issues/issue5_10/adar/index.html*.

[Alberganti, 2006] Alberganti, M. (2006). Mille milliards de mouchards. *Le Monde*. Edition du 2 juin 2006.

[Alur et al., 2001] Alur, R., Belta, C., and Ivancic, F. (2001). Hybrid modeling and simulation of biomolecular networks. In Benedetto, M. D. D. and Sangiovanni-Vincentelli, A. L., editors, *Hybrid Systems: Computation and Control, 4th International Workshop, HSCC 2001, Rome, Italy, March 28-30, 2001, Proceedings*, volume 2034 of *Lecture Notes in Computer Science*, pages 19–32. Springer.

[Angluin et al., 2005a] Angluin, D., Aspnes, J., Chan, M., Fischer, M. J., Jiang, H., and Peralta, R. (2005a). Stably computable properties of network graphs. In Prasanna, V. K., Iyengar, S., Spirakis, P., and Welsh, M., editors, *Distributed Computing in Sensor Systems: First IEEE International Conference, DCOSS 2005, Marina del Rey, CA, USE, June/July, 2005, Proceedings*, volume 3560 of *Lecture Notes in Computer Science*, pages 63–74. Springer-Verlag.

[Angluin et al., 2004] Angluin, D., Aspnes, J., Diamadi, Z., Fischer, M. J., and Peralta, R. (2004). Computation in networks of passively mobile finite-state sensors. In *Twenty-Third ACM Symposium on Principles of Distributed Computing*, pages 290–299. ACM Press.

[Angluin et al., 2006a] Angluin, D., Aspnes, J., and Eisenstat, D. (2006a). Fast computation by population protocols with a leader. In *20th International Symposium on Distributed Computing (DISC'2006)*, Lecture Notes in Computer Science. Springer. To appear.

[Angluin et al., 2006b] Angluin, D., Aspnes, J., and Eisenstat, D. (2006b). Stably computable predicates are semilinear. In *PODC '06: Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*, pages 292–299, New York, NY, USA. ACM Press.

[Angluin et al., 2005b] Angluin, D., Aspnes, J., Fischer, M. J., and Jiang, H. (2005b). Self-stabilizing population protocols. In *Ninth International Conference on Principles of Distributed Systems (OPODIS'2005)*, Lecture Notes in Computer Science, pages 79–90. Springer. To appear.

[Anshelevich et al., 2003] Anshelevich, E., Dasgupta, A., Tardos, E., and Wexler, T. (2003). Near-optimal network design with selfish agents. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 511–520. ACM Press.

[Axelrod, 1984] Axelrod, R. M. (1984). *The Evolution of Cooperation*. Basic Books.

[Balcàzar et al., 1990] Balcàzar, J., Dìaz, J., and Gabarrò, J. (1990). *Structural Complexity II*, volume 22 of *EATCS Monographs on Theoretical Computer Science*. Springer.

[Batt et al., 2006] Batt, G., Casey, R., de Jong, H., Geiselmann, J., Gouzé, J.-L., Page, M., Ropers, D., Sari, T., and Schneider, D. (2006). Qualitative analysis of the dynamics of genetic regulatory networks using piecewise-linear models. In A. Maass, S. Martinez, E. P., editor, *Mathematical and Computational Methods in Biology*. Hermann.

[Batt et al., 2005] Batt, G., Ropers, D., de Jong, H., Geiselmann, J., Mateescu, R., Page, M., and Schneider, D. (2005). Validation of qualitative models of genetic regulatory networks by model checking: analysis of the nutritional stress response in escherichia coli. *Bioinformatics*, 25:19–28.

[Beaufils, 2000] Beaufils, B. (2000). *Modèles et simulations informatiques des problèmes de coopération entre agents.* PhD thesis, Université de Lille I.

[Binmore, 1999] Binmore, K. (1999). *Jeux et Théorie des jeux.* DeBoeck Universié, Paris-Bruxelles. Traduit du livre "Fun and Games: a text on game theory" par Francis Bismans et Eulalia Damaso.

[Bower and Bolouri, 2000] Bower, J. and Bolouri, H. (2000). *Computational Modeling of Genetic and Biochemical Networks.* The MIT Press.

[Brémaud, 2001] Brémaud, P. (2001). *Markov Chains, Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer-Verlag, New York.

[Briest et al., 2005] Briest, P., Krysta, P., and Vöcking, B. (2005). Approximation techniques for utilitarian mechanism design. In Gabow, H. N. and Fagin, R., editors, *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC), Baltimore, MD, USA, May 22-24, 2005*, pages 39–48. ACM Press.

[Chen et al., 2003] Chen, Z., Gao, L., and Kwiat, K. (2003). Modeling the spread of active worms. In *Prooceedings of 22nd Annual IEEE Conference on Computer Communications (INFOCOM'2003).* IEEE Computer Society Press.

[Correa et al., 2004] Correa, J. R., Schulz, A. S., and Moses, N. E. S. (2004). Selfish routing in capacitated networks. *Mathematical Operational Research*, 29(4):961–976.

[de Jong, 2002] de Jong, H. (2002). Modeling and simulation of genetic regulatory systems: A literature review. *Journal of Computational Biology*, 9(1):67–103.

[Dyer et al., 2002] Dyer, M. E., Goldberg, L. A., Greenhill, C. S., Istrate, G., and Jerrum, M. (2002). Convergence of the iterated prisoner's dilemma game. *Combinatorics, Probability & Computing*, 11(2).

[Feigenbaum et al., 2001] Feigenbaum, J., Papadimitriou, C. H., and Shenker, S. (2001). Sharing the cost of multicast transmissions. *Journal of Computer and System Sciences*, 63(1):21–41.

[Fribourg et al., 2004] Fribourg, L., Messika, S., and Picaronny, C. (2004). Coupling and self-stabilization. In Guerraoui, R., editor, *Distributed Computing, 18th International Conference, DISC 2004, Amsterdam, The Netherlands, October 4-7, 2004, Proceedings*, volume 3274 of *Lecture Notes in Computer Science*, pages 201–215. Springer.

[Ghosh and Tomlin, 2001] Ghosh, R. and Tomlin, C. (2001). Lateral inhibition through delta-notch signaling: A piecewise affine hybrid model. In Benedetto, M. D. D. and Sangiovanni-Vincentelli, A. L., editors, *Hybrid Systems: Computation and Control, 4th International Workshop, HSCC 2001, Rome, Italy, March 28-30, 2001, Proceedings*, volume 2034 of *Lecture Notes in Computer Science*, pages 232–246. Springer.

[Gibson and Mjolsness, 2000] Gibson, M. and Mjolsness, E. (2000). *Computational Modeling of Genetic and Biochemical Networks [Bower and Bolouri, 2000]*, chapter Modeling the activity of single genes. The MIT Press.

[Gillespie, 1977] Gillespie, D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361.

[Goodwin, 1963] Goodwin, B. (1963). *Temporal Organization in Cells.* Academic Press.

[Goodwin, 1965] Goodwin, B. (1965). Oscillatory behavior in enzymatic control processes. In Weber, G., editor, *Advances in Enzyme Regulation*, pages 425–438. Pergamon Press, Oxford.

[Hethcote, 2000] Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, 42(4):599–653.

[Hirsch et al., 2003] Hirsch, M. W., Smale, S., and Devaney, R. (2003). *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Elsevier Academic Press.

[Hofbauer and Sigmund, 2003] Hofbauer, J. and Sigmund, K. (2003). Evolutionary game dynamics. *Bulletin of the American Mathematical Society*, 4:479–519.

[Kephart and White, 1991] Kephart, J. O. and White, S. R. (1991). Directed-graph epidemiological models of computer viruses. In *Proceedings of the 1991 IEEE Computer Society Symposium on Research in Security and Privacy (SSP '91)*, pages 343–361, Washington - Brussels - Tokyo. IEEE Computer Society Press.

[Kitano, 2001] Kitano, H., editor (2001). *Foundations of system biology*. MIT Press.

[Labbani, 2003] Labbani, O. (2003). Comparaison des théories des jeux pour l'étude du comportement d'agents. Master's thesis, Université de Lille I.

[Lincoln and Tiwari, 2004] Lincoln, P. and Tiwari, A. (2004). Symbolic systems biology: Hybrid modeling and analysis of biological networks. In Alur, R. and Pappas, G. J., editors, *Hybrid Systems: Computation and Control, 7th International Workshop, HSCC 2004, Philadelphia, PA, USA, March 25-27, 2004, Proceedings*, volume 2993 of *Lecture Notes in Computer Science*, pages 660–672. Springer.

[Lotka, 1920] Lotka, A. (1920). Analytical note on certain rhytmic relations in organic systems. *Proceedings of the National Academy of Science, USA*, 6:410–415.

[Lynch, 1997] Lynch, N. (1997). *Distributed Algorithms*. Morgan Kaufmann Publishers, Inc.

[May et al., 1995] May, R. M., Bohoeffer, S., and Nowak, M. A. (1995). Spatial games and evolution of cooperation. In Morán, F., Moreno, A., Merelo, J. J., and Chacón, P., editors, *Proceedings of the Third European Conference on Artificial Life : Advances in Artificial Life*, volume 929 of *Lecture Notes in Artificial Intelligence*, pages 749–759, Berlin. Springer Verlag.

[Maynard-Smith, 1981] Maynard-Smith, J. (1981). *Evolution and the Theory of Games*. Cambridge University Press, Cambridge.

[McKelvey and McLennan, 1996] McKelvey, R. and McLennan, A. (1996). Computation of equilibria in finite games. In *Handbook of Computational Economics*. Elsevier.

[McMillan, 1994] McMillan, J. (1994). Selling spectrum rights. *Journal of Economic Perspectives*, pages 145–162.

[Murray, 2002] Murray, J. D. (2002). Mathematical biology. i: An introduction. In *Biomathematics*, volume 17, pages xiv + 767. Springer Verlag, third edition.

[Murray, 2003] Murray, J. D. (2003). Mathematical biology. ii: Spatial models and biomedical applications. In *Biomathematics*, volume 18, pages xiv + 767. Springer Verlag, third edition.

[Nash, 1950] Nash, J. F. (1950). Equilibrium points in n-person games. *Proc. of the National Academy of Sciences*, 36:48–49.

[Nisan and Ronen, 1999] Nisan, N. and Ronen, A. (1999). Algorithmic mechanism design (extended abstract). In *Proceedings of the thirty-first annual ACM symposium on Theory of computing*, pages 129–140. ACM Press.

[Nowak and May, 1992] Nowak, M. A. and May, R. M. (1992). Evolutionary games and spatial chaos. *Nature*, 359(6398):826–829.

[Osbourne and Rubinstein, 1994] Osbourne and Rubinstein (1994). *A Course in Game Theory*. MIT Press.

[Papadimitriou, 2001] Papadimitriou, C. (2001). Algorithms, games, and the Internet. In ACM, editor, *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing: Hersonissos, Crete, Greece, July 6–8, 2001*, pages 749–753, New York, NY, USA. ACM Press.

[Schelling, 1969] Schelling, T. C. (1969). Models of segregation. *American Economic Review, Papers and Proceeings*, 59:488–493.

[Schelling, 1971a] Schelling, T. C. (1971a). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186.

[Schelling, 1971b] Schelling, T. C. (1971b). On the ecology of micromotives. *The public interest*, 25:61–98.

[Schelling, 1978] Schelling, T. C. (1978). *Micromotives and Macrobehavior*. Norton, NewYork.

[Serazzi and Zanero, 2003] Serazzi, G. and Zanero, S. (2003). Computer virus propagation models. In Calzarossa, M. and Gelenbe, E., editors, *MASCOTS Tutorials*, volume 2965 of *Lecture Notes in Computer Science*, pages 26–50. Springer.

[Staniford et al., 2002] Staniford, S., Paxson, V., and Weaver, N. (2002). How to own the internet in your spare time. In Boneh, D., editor, *Proceedings of the 11th USENIX Security Symposium, San Francisco, CA, USA, August 5-9, 2002*, pages 149–167. USENIX.

[Tel, 1994] Tel, G. (1994). *Introduction to Distributed Algortihms*. Cambridge University Press.

[Volterra, 1931] Volterra, V. (1931). *Leçons sur la théorie mathématique de la lutte pour la vie*. Gauthier-Villars, Paris.

[von Neumann and Morgenstern, 1944] von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, New Jersey, first edition.

[Weibull, 1995] Weibull, J. W. (1995). *Evolutionary Game Theory*. The MIT Press.

[Zou et al., 2002] Zou, C. C., Gong, W., and Towsley, D. F. (2002). Code red worm propagation modeling and analysis. In Atluri, V., editor, *Proceedings of the 9th ACM Conference on Computer and Communications Security, CCS 2002, Washingtion, DC, USA, November 18-22, 2002*, pages 138–147. ACM Press.

[Zou et al., 2003] Zou, C. C., Towsley, D., and Gong, W. (2003). Email virus propagation modeling and analysis. Technical report, University of Massachusetts, Amherst. Technical report TR-CSE-03-04.