Innovation, competition and intellectual property

Antonio Fonseca[1]

Sumário: 1.Introduction: the pace of technical change and policies. 2. Technological background: the chip as a paradigm. 3. Intellectual property and innovation dynamics. 4. The nature of the innovation process and policies.

1. Introduction: the pace of technical change and policies

Over the last two decades, protection of intellectual property has undergone intense negotiations. Two related elements which have attracted the attention of negotiators are the nature of high technologies, and the competition in the innovation process. The aim of this work is to discuss these elements and establish the importance of them to intellectual property policy.

Protection of intellectual property, in the second half of this century, was marked by the emergence of two interrelated phenomena: growing technology and new competition. The former is characterised by the ability of the developing technology to stimulate economic activities and operate revolutionary transformations. Similar to developments in information technology, they have taken place in a large range of sectors, including education and training systems, industrial relations, managerial styles, and financial systems, thus creating a huge volume of new services and affecting greatly the social mode of life. This has led to the perception that innovation is good and desirable. Such awareness has resulted in a type of a syndrome, i.e., an attitude towards a technological race and the world-wide belief that lack of technical capability will hamper a country's economic development.

The latter phenomenon is characterised by a natural change in the competitive structure of industry and the emergence of large-scale firms acting under conditions of increasingly imperfect competition. The perception of the changing nature of competitive behaviour over time explains for the dynamics

1 PhD em Direito pela Universidade de Londres (Queen Mary and Westfield College), mestre em Direito pela Universidade de Brasília, Procurador Regional da República e advogado licenciado, Conselheiro do CADE.

of the market order, and provides a justification for intermediary forms of imperfect competition (monopolistic competition and oligopolist market) unknown by the classical economics. While technical progress tends to be affected by those market models, new technologies have brought about a sense of disequilibrium and uncertainties.

Such a sense of disequilibrium is, firstly, of theoretical nature. Although too formal to function as an economic model, perfect competition is still of beneficial regulatory effects at least because of the maximisation of efficiency and welfare it pursues. When considering the increasing conditions of imperfect competition in which innovation activities occur, it can be assumed that the innovation process entails some degree of welfare and efficiency losses. Secondly, an additional type of disequilibrium stems from the fact that all regions cannot benefit evenly from new technologies as technical changes flow *a la* wave. The structural imbalance resulting from the uneven spatial development is likely to impair the ability of some regions or countries to compete. In order to follow the technical race, they face massive disadvantages and uncertainties.

Any investment project contains conceptual risks, expressed in terms of costs, an ability to satisfy demand and to compete with rivals. Such risks vary according to the magnitude of the undertaking and the level of sophistication of the technology. Additionally, companies are concerned about to what extent and how much existing legal regimes of intellectual property can afford proper protection to new technologies.[2]

Traditionally, intellectual property would protect mental creativity which results in both works (expressed in a particular form), and inventions with industrial applications. These do not include scientific ideas or theories which would belong to the public domain. As a determinant feature of the intellectual property right, the mental element places the creativity output in close relation with creators (authors or inventors). Today, the purity of this theoretical background is being challenged as never before. What matters, essentially, in the protection of new technologies, is to secure the return on investments. Hence, intellectual property rights increasingly express an investment relationship banked by firms. The mental element is no longer the most important, nor is the relation between the creator and his creation the most significant. The law leans towards protecting scientific theories and mathematical sentences mainly through increasing protection of trade secrets and know-how.

---

2 The study does not deal with the problem concerning deficiency of IPR protection for new technologies in particular. It is referred to with a view to emphasizing consequential uncertainties.

Concerned with return on investments and facing all sort of uncertainties, including lack or inadequacy of protection of proprietary rights, innovative firms are likely to develop defensive and strategic activities, i.e., trade practices which increase market imperfection and may not be acceptable in the light of competition standards. As an illustration, patents over chemical inventions from the late nineteenth century to the middle of the current century were strategically used as a tool to prevent access to the market and control output. Cross-licensing of patents in the field of electronics, carried out by leading computer firms in a restrictive style, has also taken place over the last two decades. These are only some examples of how profit-maximising firms struggle against uncertainties. In some circumstances, government assistance is a means of support, if not the only way to turn a radical innovation into a successful project.

The indisputable role of the government makes the *laissez-faire* state inadequate to attain technological progress. The state involvement in assisting firms, plainly justified by neo-liberalism, varies from country to country according to domestic traditions or the policy of the government of the day, and takes different forms. Whether the state, as a regulator, customer or underwriter, does conform its action to welfare and efficiency principles is not a matter to set up a priori. The state presence in the innovation process is crucial for pursuing a balance of interests.

In this work it is argued that while intellectual property is designed to promote technical innovation and enhance competition for public favour, the innovation process tends to be carried out in a context of increasingly imperfect competition. As a result, the achievement of the social-bargain policy, underlying the intellectual property, is impaired.

Built up under the influence of the classical economics, intellectual property has been protected under a framework of legal pre-conditions. Such protection is granted to encourage innovative activities, induce disclosure of information, reward inventors and authors, and boost industrial applications. These are assumed to work for the benefit of the society at large, and to found a sense of social bargain assessed and comprehended in the light of welfare and efficiency goals.[3] The purpose in placing the institutional bargaining in a welfare-and-efficiency perspective is to reach an understanding of the combined institutions of law and economics in intellectual property. Such an

3 There is not a definite description of efficiency and welfare goals. It is here assumed that to limit the intellectual property on efficiency grounds requires a fault or competition mischief on the part of the owner. A limitation not based on efficiency grounds can only be justified on public welfare (e.g., expansion of employment, export and tax basis, balance of payment, supply of a product essential to public health or national security, and to correct a distortion of competitiveness or distribution of the industry).

analytical approach provides a dynamic and more precise sense of the elements forming that bargain in connection with the innovation process.

While protecting themselves against the risks inherent in innovative activities, firms move towards concentration. As a result, barriers for entrants and distortions to competition are likely to be created. These potentially work against the purposes underlying intellectual property. On this grounds lies a strong argument for limiting the use of proprietary rights.

The study begins by describing the nature of chip technology, which is taken as a paradigm for a number of reasons. Firstly, protection of chip design is regulated by specific international treaties, establishing a legal regime of intellectual property and with which this study is mostly concerned. Secondly, as a product with multiple applications, chips have a close relationship with computer software, data-basis and artificial intelligence. Developments on chip designing and manufacturing[4] give rise to theoretical-legal concerns.

2. Technological background: the chip as a paradigm

2.1 The semiconductor chip

2.1.1 Chip: from history to business

The huge difference between today's computers (based on large-scale circuit integration) from those in the 1950s (valve computers) can be explained in terms of cost, reliability, user- friendliness, speed of operations performed, and memory size. Succeeding valve computers, transistors represented a development which was limited due to the difficulty of interconnecting them. Advancement allowed logical units, made up of a number of transistors and its associated circuitry and connections to be placed on single semiconductor material (a chip). The development of this technology led to machine miniaturisation, creating third and fourth generation computers. The integrated circuits not only made the formidable change in performance possible, but also determined the overwhelming growth of computer technology.

4 To a degree, integrated circuits are in the core of, or associated with, technologies such as advanced semiconductor devices, artificial intelligence, digital imaging technology, flexible computer-integrated manufacturing, high-density data storage, high-performance computing, and sensor technology. For details, see "Emerging Technologies — A Survey of Technical and Economic Opportunities", US Department of Commerce, 1990. [From now on "1990 DOC Technical Survey"].

2.1.1.1 Some technical definitions and the importance of the chips

The chip is a popular name for an integrated circuit or semiconductor chip, which is an electronic device with electrical functions. These terms are synonyms, differing only in the product manufacturing process. The terms topography, circuit layout, layout-design and mask work, now legally coined, are used interchangeably to indicate the arrangement of the elements representing the three-dimensional structure of the chip. The term chip-design also appears in this study as a synonym for the same representation of that arrangement on which the legal protection relies. Hence, the chip or integrated circuit is the final product or device in solid state distinguished from its layout or design itself.

The complex collection of transistors contained in an integrated circuit corresponds to minuscule patterns of switches which control electric current and perform assigned functions (manipulation of electrical signals) at nearly the speed of light. The transistors determine the chip capacity assessed in terms of computing power, speed, power consuming, reliability, and cost. These require the use of very sensitive types of material.

Well known as a semiconductor system of circuits, the chip is made up of two broad categories of material: good conductors (rich in conducting electricity) and bad conductors (insulators). Due to high technical methods and processes, thousands of circuits are imprinted in a single, thin structure forming a semiconductor compound or substrate of material such as silicon, glass, sapphire, ceramic, magnetic domain, and superconducting material. Several types of chips differ from each other due to the manufacturing methods or process they apply or the functions they perform.

According to their manufacturing methods, chips are bipolar or MOS (metal oxide semiconductor). Power consumption and speed depend on such methods. In their variations, MOS chips are technologically dominant and have wider applications. Linear and digital circuits differ from each other due to the methods of altering electrical signals. Linear circuits process electrical signals over a continuous voltage range, and are suitable for analog computers, radios and TV sets. Digital circuits are suitable for processing information in bits (binary digital), and are largely used in digital computers. Within the digital category, a distinction is made between logic (microprocessor) and memory chips.

Two basic functions of a chip include computing of processing information and storing data (as either input or output already saved for ulterior computations). Although these functions can be performed by a single device, memory chips have the primary function of storing data or programs; they are ROM (read-only-memory), PROM (programable-read-only-memory), and

EPROM (erasable PROM) chips. A microprocessor has complex logic circuits containing the basic elements (forming a central processing unit - CPU) of a conventional computer. For this reason the microprocessor is regarded as a microcomputer on a chip. Both functions (storing data, and making decisions which rely on data) could be integrated on a single VLSI (very large-scale integrated) chip whose use is not confined to computers.

Developments in computer technology would not have been possible without integrated circuits. Their applications, therefore, go beyond the computer industry, to include consumer products, telecommunication equipment, industrial process control, medical and manufacturing equipment, defense systems, and any area which requires significant use of electronics. These growing applications illustrate how crucial the technological progress and competitiveness of the integrated circuit industry is to the economic growth of any nation. Yet, only a minority have been able to enter the chip business.

The design and manufacture of chips requires a considerable amount of investment and a highly trained labour force. The innovative activity takes thousands of hours of research and development, and is a costly business. Designing and marketing an entire family of integrated circuits may take years and million of dollars. Nevertheless, such high costs are alleviated thanks to automation and mass production, so that the price per unit is only a few dollars.

2.1.2 Designing and manufacturing process

Designing, manufacturing and testing a chip involves decisions regarding which techniques to apply, costs and purposes. Advances in methods and in the manufacturing process offer a variety of options which meet specific needs. The appropriateness of the technology[5] depends on the type and the amount of information one wants to include in a single chip or chip system.[6] The scale of integration, the flexibility of the microprocessor (the versatility to update), the advances in CAD,[7] and the purposes -if it is an application-specific integrated circuit (ASIC), or a general-purpose microprocessor - affect the costs. These technical requirements, costs and purposes are intrinsically related. The search for profit and capability are contributing factors. If designing and production have commercial purpose, a microprocessor may be smaller and cheaper in order to reach competitiveness; perhaps, no similar product has been produced before, and hence a generous scale of production is considered. If the

5 In the description of the steps below, the designing and fabrication of an MOS integrated circuit is considered. The MOS technology dominates the IC market, and is largely applied to VLSI circuits. Cf. A. F. Murray & H. M. Reekie, Integrated Circuit Design, pp. 6, 24.

6 A microprocessor system differ from a microprocessor on a single chip in the sense that the former includes a printed circuit board, a few number of chips and discrete components.

7 The computer-aided design (CAD) consists of a variety of hardware and software tools.

focus is on capability, e.g., a microprocessor for military application, power and performance are decisive.

Designing follows several steps. (*a*) Abstract description. A plan of the electrical functions to be performed is prepared. The electrical specifications are described with precision and in detail. A market study previously undertaken perhaps supports the conception of the desired functions. (*b*) Logic diagram. A detailed schematic data describes the circuits symbolically. This is a very important piece of work, which requires talent and experience.[8] (*c*) Layout design. The arrangement of the components and the complex interconnection patterns is defined. The selected geometrical placement of the elements provides a picture of how the chip topology will be implemented. The designer is then able to make input (progressive specification of data) in order to optimise the layout configuration; by manipulating the schematic he makes choices, selecting a particular way of arranging the elements in the semiconductor substrate.[9] He is bound, however, to adhering to a set of technical design rules. These rules represent "constraints" upon the freedom of design, and are dictated by technological considerations. The geometric rules, for instance, address the problem of the transistor size; the electrical rules specify electrical parameters applied according to the manufacturing process; some mandatory features are also imposed, and are supposed to be present in every design.[10] To observe and implement these rules, the designer enjoys the aid of the computer which is regularly utilised.

Although designing can be computer aided (CAD), the simulator capacity of mimicking the circuit and predicting its behaviour is limited; some inaccuracies do exist, and thus the design automation tolerates certain levels of inefficiency. For this reason, the designer's intuition is needed.[11] Furthermore, automation is developed inside large companies. As access to them is rather difficult, it is uncertain how much simulation is applied. Increasing of IC density has rendered the use of simulation nearly indispensable. Moreover, correction[12] of the chip configuration must be made before the design is released for mass production, otherwise modifying the chip is impossible.[13] The simulation patterns are applied to verify the logic design, i.e., to check its internal consistency, help generate alternate architecture, and file additional information regarding the whole IC network. Such data will be useful during

8 Provided the schematic is sufficiently novel, protection may be available under the patent law.

9 This job carries out the considerable work of mind that the *sui generis* law protects. The layout design corresponds to an encoded set of masks - the "mask work" of the American SCPA.

10 For details about the design rules, see Maurray & Reekie, ob. cit., p. 63 *et seq.*

11 Murray & Reekie, ob. cit., p. 101.

12 At every stage corrections are performed, by adding further specifications and improving earlier results.

13 After fabrication, each layer or mask is permanently fixed or embodied in the semiconductor material.

final testing. The more complex the chip architecture is, the more automation is needed, despite the challenges posed by the simulation.[14] A factual consequence, however, deriving from the CAD (computer-aided design) discipline and the strict design rules to which the designer is bound, is that limitations on engineering techniques lead designers to create independently layout circuit which may be substantially similar.[15]

The material is ready for manufacturing[16] when the interconnecting pattern is complete and correct. The integrated circuit is developed by the transfer of the encoded pattern, through an expensive process and by applying a series of operations.[17] The result is a collection of masks,[18] which determine the features of the transistors.

The stages in the manufacturing of a silicon-based integrated circuit are as follows. The masks are produced by photo-reducing the circuit design. The manufacturing process itself starts with the oxidisation of the silicon. At a high temperature, chemical and photographic treatments are applied on the substrate, including repeated addition and removing of materials. The result is a resistant product consisting of a basic metal-oxide semiconductor transistor. The last step is testing each chip still on the wafer.[19] Those which do not perform the desired functions are rejected and thrown away.

---

14 Challenges, for instance, in terms of reliability. See M. Feuer, "VLSI Design Automation: An Introduction", a supplemental article presented to the "Subcommittee on Courts, Civil Liberties, and the Administration of Justice", of the American House of Representatives, H.R. 1028, pp. 380 *et seq*. In a less accurate source, the generalized use of automation seems to create no problems, mainly in designing of gate arrays and standard cells. Special Report from Business Week, May 23, 1983, transcript in Hearings on S 1201, p. 162.

15 The problem of substantial similarity is legally relevant, and, as a technical fact, was referred by M.A. Lechter in his written comments recorded in the hearings of the H.R. 1028, p. 280. Serious incompatibility would exist, however, in applying the copyright test of similarity in the domain of the semiconductor chip.

16 The chip law does not primarily focus on the fabrication or the product, but on the chip design (the intermediate masks) instead. Nevertheless, the manufacture helps to understand some legal definitions. Moreover, the design normally reflects specific manufacturing process; the interrelation between them may be rather significant. "On one hand, designs may have to be substantially modified because of manufacturing limitations while, on the other hand, advances in manufacturing techniques or materials may compel major changes in design parameters." - Cf. [1988] III(4) Monthly Labor Review 27; see Hearings on S 1201, p. 162.

17 Such as *metallisation*: application of a metal which is used for interconnections of the device, and act against the high resistance of other materials; and *insulation* or oxidation: a layer of oxide, an insulating material, is deposited on the wafer (a disk of silicon) in order "to prevent any undesirable short-circuits" producing silicon dioxide. This material is a very good insulator, permitting the application of the masking technique at a high temperature. Murray & Reekie, ob.cit, p. 48/59.

18 They represent the number of layers (10 to 16), precisely aligned or juxtaposed; each of one has less than one micron (one thousandth of a millimeter), and bears the information concerning both the processing technology, and the electronic system embodied in the chip; they together describe the entire topographical dimension of the chip.

19 A wafer is approximately five inches in diameter and 0.025 inches thick, and can yield 100 to 200 chips at one time. The higher the number of sound devices per wafer, the lower the end-cost per output unit.

2.1.3 Reverse engineering and audit trail

Ordinarily, there are two ways of getting access to a given chip: obtaining a pattern either (*a*) in form of a tape,[20] or (*b*) through the reverse engineering process.[21] Both may derive from a normal technology share agreement[22] but the latter - although being a lawful practice - may be a step towards a misappropriation.

2.1.3.1 Defining reverse engineering

Reverse engineering is a process by which one may disassemble the chip into its constituent patterns (masks or layers), using photomicrography. The top layer is photographed, carefully measured (and the related information preserved appropriately) and etched away in order to expose the next pattern, and so forth layer by layer, until the schematic of the whole chip is drawn. The operation is undertaken with a microscope and a camera mounted to take pictures, and the layers are removed by applying a set of chemical baths. When the entire mask set is reconstructed, the embodied principles, techniques or specifications (concepts and ideas), are evaluated for the purposes of studying or teaching. Next, another IC layout may be designed around the protected one, modifying and improving it, both chips (the model and the second one) being functionally equivalent, but visually dissimilar.[23]

It is indisputable that reverse engineering is an appreciable means of technology diffusion. Its accessibility effect is particularly understood within the context of the second-source manufacture (integrated circuits

---

20 The IC layout tape, including the reticle set and working masks, are carefully kept by the company. These intellectual assets - according to the 1991 amendment to the Unfair Competition Prevention Law of Japan - should be part of an inventory in order to be protected. [1991] ICLA 13, Nov.

21 One well-known case of reverse engineering that has been cited was the NEC version to the Intel 8080 microprocessor. The Intel assumed that its chip was served as a model by NEC, which analyzed the 8080 allowed by a private agreement signed in 1976 with the Intel. See Hearings on HR 1028, pp. 39/40.

22 Technology share agreements are commonly made by great corporations. Toshiba, Siemens and IBM have recently joined to create a memory chip which will hold 256 megabits by 1998. The reason for going into alliance is basically the high cost of research: "Toshiba earns US$ 7 billion from chip each year. It will cost US$ 1 billion to develop the 256-megabit chip." [1992] 135 (1831) New Scientist 9.

23 Cf. Hearings on HR 1028 p. 392, Hearings on S 1201, pp. 27/28 and 38. American firms specialised in chip analysis charge a few tens of thousand dollars for assembling service, including topological layouts, and material analysis. The high-price range may oscillate from $10,000 to $30,000, but one may come across advertised chip reports at $980 to $1880, "with volume discounts for additional copies." Cf. M. D. Goldberg, Intellectual Property Rights and Technology - Semiconductor Chip Protection as a Case Study, paper presented at the Conference on Global Dimensions of Intellectual Property Rights in Science and Technology, held on January 8-9, 1992, at the National Academy of Sciences, Washington, DC.

interchangeable with counterparts). For technical and commercial reasons, a firm may want to make a chip equivalent to a competitor's, or a manufacturer to have a second-source of its product in pursuit of adequate supply, market certainty, technical compatibility and cost reduction.[24] Second sourcing, a common practice in the US semiconductor industry, provides the buyers with at least two possible suppliers, protecting them against the risk of excess demand.[25] The equivalent product, normally resulting from a private agreement, would be a competitive version enjoying lawful circulation. Whatever the status of the equivalent product, whether a copy or a legitimate and similar one, an additional issue is the reproduction of the microcode[26] built into a memory chip.[27] As far as the law[28] is concerned, there is a potential conflict between the decompilation of a chip and a computer program.[29] Apart from this aspect, to find out whether or not a second-comer is a copying output is legally relevant. In this respect, the audit trail[30] is of some assistance.

2.1.3.2 Definition of audit trail

The audit trail consists of the overall documented job of trial-and-error performed along the course of the chip design, and include blueprints, computer simulation outcomes, logic circuit diagrams, trial layouts, test data, and time records. These elements are necessarily generated as a result of significant efforts put in the making of an IC design, and may be printed on paper (paper trail) or electronically stored in a computer (electronic trail).[31] The electronic

24 For details, see J. C. Oxman, Intellectual Property Protection and Integrated Circuit Masks, an article reprinted from the Jourimetrics Journal and presented at US Congress as supplemental material, Hearings on HR 1028 p. 388/9.

25 Cf. UN Chip Report, pp. 142/143. See Table 8.4. Second sourcing is also a legal requirement of the US public procurement law. See Luc Soete, "International Diffusion of Technology, Industrial Development and Technological Leapfrogging", [1985] 13(3) World Development 409, at 421 (footnote 36).

26 The microcode is a particular computer program built into a chip as a pattern of tiny transistors, i.e., a piece or portion of electrical circuitry.

27 This is a matter of great importance because the memory chips or RAMs form a category considered as "the vital fuel of the computer industry". Aware of this and by the time the American SCPA was passed, the US Defense Department was worried about the possibility of the US computers, weapons and telecommunication become dependent on foreign memory chips. This concern made sense, because Japan soon after emerged as a leading force in the market of memory chips. Hearings on HR 1028 p. 359.

28 Although lawful under the *sui generis* chip law, reverse engineering is uncertainty in copyright as this applies to computer program.

29 Disassembling a memory chip technically leads to the decompilation of the computer program microcode embodied in the chip.

30 Audit trail is here applied replacing the expression paper trail. The former seems to be more appropriate, because the elements involved in the concept rely increasingly on electronic means rather than on paper.

31 The electronic printing may include accidental errors or traps. This is the case of a small imperfection fixed in the Intel 8086, causing a chip designer to discover by chance the copying made by NEC in the

trail incorporates technical principles, specifications, ideas and concepts manipulated or arranged by the chip designer in the course of the making of an original chip.

A discerning observer should be able to tell whether a chip is a copy or fruit of reverse engineering.[32] The distinction which needs to be made is a matter of "change" or "adaptation"[33] rather than a direct evidence of authorship.[34] It follows, if the audit trail has been produced it does not necessarily mean the IC-design is an original one.[35] The audit trail is significant in the sense that it does provide evidence of systematic tasks and investment, but it is not a test of originality.

As hardware, an integrated circuit is a device very distinct from computer software. Nevertheless, these two technical elements work together in a large number of applications, mainly in computing. The scale of this technical interplay is such that commercial and industrial exploitation of integrated circuits and software considerably affect each other, and the infringement of a microprocessor chip most likely involves infringement of computer software as well. For this reason, an approach to advancements in computer software seems commendable.

2.2 Computer software and artificial intelligence systems

2.2.1 Development of computer software[36]

2.2.1.1 Definition of computer software

As a legal concept, computer software includes the *computer program*, *program description* plus any other related *supporting material* necessary to the whole specification of the computer program itself.[37] The conception and

---

fabrication of an 8086 version. This most famous copying case is part of the high-scale competition between Intel & NEC, the two giants of the electronic industry.

32 L.L. Vadasz, loc. cit., p. 37.

33 The debates carried out at the US Congress suggested that the audit trail was only half important. As a result of a technical routine and in-door activity, the paper trail could hardly be accepted as a proper test of originality, thus, unsuitable to be included in a legislation dealing with intellectual property.

34 In technical sense, significantly different designs may present very subtle mask changes. Hearings on HR 1028 p. 37.

35 The conclusion is a valid one, but it is assumed that to forge an audit trail is nearly impossible.

36 See IEEE Standards Collection, Software Engineering (1993).

37 The WIPO defines these terms as follows: *computer program:* "a set of instructions capable, when incorporated in a machine-readable medium, of causing a machine having information-processing capabilities to indicate, perform or achieve a particular function, task or result"; *program description*: "a complete procedural presentation in verbal, schematic or other form, in sufficient detail to determine a set of instructions constituting a corresponding computer program;" *supporting material:* "any

execution of a software project may involve a considerable amount of intellectual effort and investment. These inputs vary according to the software application which could be for the control of a nuclear reactor or a washing machine. Regardless what the software function is, its development entails the same basic phases roughly associated with those elements relevant to the legal concept. These phases are the specification, designing and programming.

2.2.1.2 Development steps

At the initial stages of the process, there are specifications or statements of requirements provided by the customer. Usually written in natural language, the requirements may consist of a few pages or a number of volumes, and describe what the program is required to do (function or task) within certain conditions or limits (constraints).[38] The language of the statements often contains plenty of imperfections, such as ambiguities, omissions and so on. Thorough analysis of the statements is then carried out in order to resolve such imperfections and reach an agreed specification, formulated in accordance with certain properties and understandable by both customer and developer. Once completed and tested, the specification describes what the system is to do in terms of application; the description is a basic document to develop the system design.

The designing is the second phase. At this stage, procedures or subroutines are arranged and grouped in units. *Program units* are sets of codes and data which define each function or task and their performing order, and are capable of intercommunicating in a logical flow by parameters. The architecture of the operations organises the data in terms of sequential file, expressed in algorithms, i.e., set of steps, and is expected to satisfy the functional specification and constraints. A detailed logical design structure of the operations is then reduced to a form called a *flowchart,*[39] *which expresses how a system, as a series of functions, is to be implemented in computing terms.*

In the third phase, algorithms and program units are written in computer language.[40] The flowchart is now translated into *source code* or source program,

---

material, other than a computer program or a program description, created for aiding the understanding or application of a computer program, for example problem descriptions and user instruction." Draft Treaty (Article 1) and 1977 Model Provisions (Section 1) on the Protection of Computer Software. Computer software and computer program are terms used by academics and practitioners interchangeably.

38 A fragment of a hypothetical statement: if the driver does not put on the seat belt and the engine is started, an alarm will sound intermittently.

39 The flowchart or flow diagram is independent of the coding, and is said to represent the idea behind the computer program (cf. K R Moon, [1991] CLP 158). Apart from the idea/expression dichotomy, the arrangement per se of algorithms, mathematical statements, procedures or subroutines, whatever form of language expressed, would justify copyright protection.

40 The types of language applied include BASIC, COBOL, FORTRAN, PL/1, PROLOGO, LISP, C, and PASCAL.

which describes key statements in mathematical notions. The translation is made through an *interpreter*, instruction by instruction, or a *compiler* which translates the whole diagram in one operation.[41]

In order to be run and commercialised, the source program is translated into object code (code program or machine code), which is a series of instructions to be operated by the computer, and written in a special format. As the translation is carried out aided by the computer, the source program is taken as an input supplied to the translator. As a second program or output, the object code takes a machine-readable form. Its binary notation makes up sequences of zeros and ones,[42] which correspond to equivalent wired commands electronically expressed as "Off" and "ON" switches. These instructions are loaded into the electronic memory and organised into sets of bytes. The physical means used to store the data includes magnetic or optic disc, electro-mechanical switches and semiconductor chips,[43] appropriate devices to market software. The range of tasks in the different phases are performed by teams of professiona!s hired by a corporate entity. The circumstances in which software is conceived and developed almost invariably do not allow the creator (or creators) a close relationship with the product. This feature is more pronounced in systems, such as artificial intelligence, in which the interoperability of hardware and software is more complex.

2.2.2 Artificial intelligence systems, concepts and functions

Can a machine think? This has been an intriguing question of this century,[44] and a challenge which remains in the frontiers of the computer

---

41 Interpreter and compiler are special programmes written specially to accomplish the translation.

42 In computer sense, the binary digit "zero" or "one" is called "bit"; a sequence of eight bits form a "byte" which is treated as a single unit and represents a character (a letter, number or symbol). According to the American Standard Code for Information Interchange - ASCII the most commonly used characters are represented in decimal codes and interchanged into binary codes. For instance, the letters for MARY have the notation M=77, A=65, R=82, Y=89; in binary codes the name in capitals is represented as follows:

0 1 0 1 1 0 1 0
0 1 0 0 0 0 0 1
0 1 0 1 0 0 1 0
0 1 0 1 1 0 0 1

Electronically, zero and one represent a switch with its contact open or "off" and close or "on", and means low and high voltages, or different polarities of magnetization.

43 A microprogram permanently stored in a ROM chip (Read-Only Memory), in microcode instructions, is called *firmware*. In a microcomputer, the processor unit (CPU) consists of one or more of this device used to control and direct the microprocessor's activities.

44 The question was considered for the first time in 1950 by Alan Turing, cf. Palfreman & Swade, The Dream Machine, pp. 137/138, The BBC Books, 1991.

science. A similar question could have been made two centuries ago with respect to the airplane, can a machine fly. For the average laymen both questions (made in the corresponding due era) allow similar curiosity and skepticism. Under the eyes of science, however, those questions differ fundamentally. Man discovered the principles of flight, which enabled the machine to fly. In order to make the machine think, man would supposedly need to discover the principles of intelligent thought.[45] Moreover, these principles are supposed to provide the scientific ground for the development of a machine with the ability to recognise things, adapt to a new environment, learn and create. Such a *rationality disembodiment* project has no precedent in the history of the industrial revolution. In conceptual terms the implications are enormous.[46] The creations of the so called "electronic brains", however, have not gone beyond "idiot savants,"[47] which have resulted in little success achieved only in confined areas. In this respect, frustrations[48] have been debited to the complexity of the real world that artificial intelligence purports to reproduce, and to the still mysterious scientific concept of (human or real) intelligence.

As a technology in development, AI relies on scientific models not fully confirmed, and not yet satisfactorily defined.[49] In attempting to draw some

---

45 The (human) intelligence is something associated with the process of thought, reasoning and learning. Although consisting of neural events confined to the brain, thinking is not identified today with conscious experiences - these are rather limited, but with the learning experience captured through stimuli (inputs) and responses (outputs). This process would originate cognitive structures or "perceptual representations of the world or parts of the world." Many psychologists are concerned with the mental structures irrespective of man being aware of them. In the 20th century there is no unanimity among the scientists about the intelligence phenomenon. For a sound account on this, see Encyclopedia Britannica, vol. 22, pp. 641 *et seq*.

46 The ability to reason distinguishes the human from the rest of the life forms and things. From this phenomenon, the law has been universally developed under the assumption that the man is the unique being capable of having his own will, and so only the human being is bearer of rights and duties, with the exception of the artificial person or legal entity being applied.

47 Computerized machine built to manipulate concepts like human brains has comparatively been "brilliantly gifted in one small area, but outside that area, he is unable to function competently." Palfreman & Swade, ob. cit., p. 154.

48 Based on unexpected outcomes, irreverent. AI definitions have been made, such as "any software system which is sufficiently sophisticated that it doesn't quite work", and meant as "Always Impossible" or "Advanced Implements". WIPO pub. 698(E), pp. 121, 95. The unfavourable comments do not seems to apply to robots, which have a secure future in the manufacturing process, and are progressing quite well in biorobotics. A team of researchers in Montreal, at the Biorobotics Laboratory of McGill University, is building a microrobot called Micro Surgery Robot-1. The MSR-1 is designed to perform eye surgery. The system "creates a three-dimensional robot's eye view of the inside of the eye that the surgeon can see by wearing a virtual reality helmet that has a small screen in front of each eye." [1992] 134(1826) New Scientist 22.

49 In the Symposium on the Intellectual Property Aspects of Artificial Intelligence sponsored by WIPO, held at Stanford University in 1991, the WIPO Director General delivered in his opening statement a preliminary definition as follows: "an expression commonly used to designate those kinds of computer systems that display certain capabilities associated with human intelligence, such as perception, understanding, learning, reasoning and problem-solving." WIPO pub. 698(E), p. 17.

concepts, specialists are prone to centre on technical concepts associated with operations and outputs.[50] This approach avoids both the underlying debate about the nature of intelligence, and the uncertainty concerning actual learning as a possible result from machine tutorial.[51] Nevertheless, available knowledge in the field of computer science (including development in software and hardware) only provides for limited explanations. In addition to scientific doubts and skepticism and as far as the legal interest[52] is concerned, a way of approaching AI systems is to consider their parts, and that software is one of them. This leads to the question of how AI systems differ from conventional software, involving, *inter alia*, aspects related to concepts, function and structure, categories, applications and development.

Some attempts at a definition regard artificial intelligence as a (a) computer system, (b) possessing certain capabilities (c) developed on a human-like basis and (d) addressed to specific goals. As a computer system, artificial intelligence relies on sophisticated sets of software and hardware, which process or manipulate electronic representations, and draw inferences.[53] These patterns of magnetic or electronic current, common in a digital computer, are responsible for the processing of the internal representations of the external world.[54] As a representational system,[55] AI stands beyond its physical basis and is not reduced to a device.

The output the AI systems intend to operate include sound emission, writing, and perception. These capabilities, achieved through manipulation or application of knowledge (cognitive tasks), result from a process of, or equivalent to, learning, reasoning, and self-adjustment. Such a function is

---

50 Apart from the lack of consensus about the definition of human intelligence, what really matters is to know how an artificial system works in order to be accepted as an intelligent one. This treatment tends to cast aside false and exaggerated expectations derived from the expression artificial intelligence. Cf. Dreier, WIPO pub. 698(E), p. 151.

51 As Johnson-Laird pointed out, "neural networks are not so sure; they only seem to learn from failure. When they are wrong you tell them the correct answer and they adjust. When they are right, it is not clear that they are actually learning." In "Main Categories of Artificial Intelligence and Their Intellectual Property Aspects", WIPO pub. 698(E), p. 45.

52 The approach to these interests has been made much more on basis of speculations, because as far as intellectual property is concerned no serious problem has been encountered yet, as it was reported at the 1991 WIPO Symposium on the Intellectual Aspects of Artificial Intelligence, US Stanford University, Doc. 698(E), p. 298.

53 An expert system (a well-developed subdivision of artificial intelligence) has basically three components: knowledge base, inference engine and user interface. The knowledge base contains interrelated information about particular area. The inference is a reasoning process or a means of using that information and so as to render specific goals.

54 CYC, a super knowledge base, is being built since 1984 in Austin, Texas. The project shall take at least ten years, and is intended both to capture the every-day world knowledge and to express common sense. Such an ambitious project has inspired skepticism. The Dream Machine, p. 157 *et seq.*

55 The representation of knowledge is largely developed in a hand-crafted way. A initial 30%-error rate is something expected. The error rate after the training test set is inferior to 7%.

reduced purely to a mechanism of randomisation,[56] processed before and after the system is made. The operation requires the system to understand or interpret input, and gradually infer solutions from stored knowledge (database).[57] Bearing a utilitarian character, the AI systems aim at meeting a human need, rivaling or assisting man, replacing him in the performance of complex tasks,[58] and solving problems efficiently in narrow areas.

Before discussing the next point (machine tutoring), a conceptual line is now drawn between what is called artificial intelligence and intelligence as a human attribute. AI systems may only assume a putative intelligence in the sense that they express imperfect analogy with few faculties of human beings, and there is no need to demonstrate the nature of the artificial representation of the external world. Such a remark, while limiting the expectation created by the Turing query, confines AI systems to truly semi-autonomous and therefore limited creations. In other words, the "intelligence" of the so-called intelligent machines is reduced merely to a particular achievement which may, to a certain extent, recollect or emulate an attribute inherent to a human one.

In order to sufficiently describe the object of protection, the law-maker has to set, among other prerequisites, the minimal level of complexity, technically defined by the speed and the number of inferences, as a pattern of both intellectual and investment inputs.

2.2.3 AI development and machine tutoring

The creation of an AI system involves a large number of specialists, such as programmers and knowledge engineers, as well as professionals from different fields other than computer science. The development encompasses those activities achieved in the production of any conventional software, plus improvement tasks and testing on a more intensive scale. These activities may vary within the AI categories,[59] however all of them are intended to be capable

---

56 Randomisation is a sort of interaction operator/machine. The operator presents a series of codified facts (inputs) and then the machine is asked questions associated with those facts. Incorrect answers (outputs) are supposed to come out. Each mistake requires adjustments, until the right answer is served. This may be a long job, complex and costly, equivalent to a training or tutoring performance, on which the intellectual content or creativity relies.

57 Something similar, but really in a lower scale, applies to "conventional programming disciplines", reduced to routines mathematically serviced step-by-step and as part of the "intellectual creativity of computer programs (i.e.) the creative combination of instructions and statements expressed therein." The AI system "training", however, is distinguished by its "indeterminacy". See S. Miki, "The Creation of Works of Copyright under Japanese Copyright Law Resulting from the Utilization of Artificial Intelligence"; and R. S. Laurie, "The Patentability of Artificial Intelligence Under US Law", both in WIPO pub. 698(E), pp. 294 and 122.

58 Some applications include medical diagnosis, translation, financial analysis, geological search, weather forecast, and recognition of military target.

59 The WIPO has identified three categories of AI systems: the classical expert system, perception system,

of learning. The focus will be, notably and briefly, on the processing of the representation of knowledge and inference, tutoring, and the audit trail formation.

The processing of knowledge relies on two forms of representation: symbolic and non-symbolic;[60] expert systems, for instance, apply the former and neural networks apply the latter. In standard expert systems, the knowledge (substantive and procedural information) is translated into appropriate and formalised rules (representation) and implemented in a data base. In a further stage, the inference engine (software set)[61] is designed and implemented by applying the knowledge to a particular problem area. In neural networks, which simulate human brain functioning,[62] three layers of artificial neurones, equivalent to RAM-memory chips,[63] comprise the system structure: input layer, "hidden" layer and the output layer. These layers form a sort of connectionist system, in which the relation between input and output is given through assigned weights.[64] For the system to work successfully, the skill in choosing the data representations (i.e., the number of neurones attributed to the input, hidden and output layers), the initial weights, and the selection of training facts are crucial. In addition to such required skills, an interesting feature of the neural system is that its intelligence "derives, at least in part, from the way in which the elements are interconnected rather than being entirely the product of programming."[65]

---

and natural language. Other classification includes less explored subclasses, such as neural network and robotics, and exclude perception systems. From the point of view of the US Patent Office regulation, a broad class (364 - electrical computers and data processing systems) lodges Generic AI Inventions (subclass 513) based on expert systems, neural networks and robotics. There are other tens of subclasses connected with a few classes (381, 382, 414), covering a number of AI applications, besides the non-generic (dedicated AI-based inventions tools) covering related subcategories. [WIPO pub. 389(6), pp. 123/4].

60 Machine translations and genetic algorithms, for instance, apply representations based on symbolic framework, as most AI systems do. Non-symbolic representations rely on connectionist or neural frameworks, as is the case of neural network systems regarded today as an embryonic form of artificial intelligence.

61 See J. H. Spoor, "Protecting Expert Systems, in Particular Expert System Knowledge: A Challenge for Lawyers", in WIPO pub. 698(E), p. 77. The inference engine operates and controls the expert system by "selecting rules to use, accessing and executing the selected rules, and determining when a solution has been found." (Technical Appendix to "The Patentability of Artificial Intelligence Under US Law" by R. S. Laurie, p. 141.)

62 The first logical model of artificial neuron (an idea of brain-like machine) was produced in 1943 by McCulloch and Pitts (University of Illinois). Cf. Aleksander & Burnett (1987), Thinking Machines, The Search for Artificial Intelligence, pp. 156, 198, Oxford Univ Press.

63 A bit-organized RAM (random access memory) is imprinted in microchips or silicon neurons, which are repositories and processors of information. They simply work as interacting computers.

64 The network relates the input values to the correct output by means of weights. Before training, the designer arbitrarily sets "the weights from the input layer to the hidden layer and from the hidden layer to the output layer." Cf. Appendix, WIPO pub. 698(E), p. 143.

65 Aleksander & Burnett, ob, cit., p. 197.

The distinct categories of systems described above lead to different procedures of learning simulation.[66] Two examples of these systems, which focus on specific tutoring patterns, are natural language processing (a translation machine), using symbolic representation, which is nearly hand-crafted, i.e., written and encoded explicitly by hand,[67] and a neural framework applying non-symbolic representation.

The natural language system requires: *a*) a grammar to assure the right order of words in a sentence. Since there is no formalised grammars,[68] such as the existing standard codes of computer language, a particular grammar has to be made, which is time consuming and costly; *b*) a lexical system, which is a definition of words (dictionary); *c*) lexical disambiguation, i.e., a set of rules designed to provide contextual meaning. This is fundamental for dealing with the syntax of certain words which play different roles, such as verb, adjective, or noun, according to the context; and *d*) a combining approach and testing, necessary for generalisation of algorithms and instruction of the system with patterns of translation. This activity is a plus in terms of intellectual content, and so remarkably distinctive in the whole system.

The example of a non-symbolic representation is a neural framework designed for the analysis of DNA sequences, with the purposes of recognising "promoter sites" and "splice junctions."[69] Extracted from biological literature, an inaccurate theory is formulated to explain the rules of promoter sites and splice junctions. Following that, these rules are encoded into a neural network, i.e., in a network typology, and an initial set of weights is provided. The next stage consists of training the network. The training consists of strengthening or weakening the connections between the processors of the system; this adjusting of weight patterns with the initial typology, using known examples of DNA sequences, aims at improving the theory.[70] The result is a particular representation of knowledge in biology, useful for the study of promoter sites and splice junctions. The error rates in the recognising of those elements are inferior to 6.5% considered as very good.[71]

---

66 As it is pointed out, "in a traditional expert system, the knowledge engineer specifies rules and search techniques to correlate input and output. In a neural network, the system itself designs and adjusts the weights in order to correctly correlate input and output." In Technical Appendix, WIPO pub. 698(E), p. 143.

67 L. T. McCarty, *loc cit*, p. 34.

68 The Japanese Electronic Dictionary Research Institute is carrying out research to develop an electronic dictionary intended to apply to any type of system. Cf. Makot Nagao, WIPO pub. 696(E), p. 41.

69 Promoter site is a biological element associated with the process of gene transcription. The identification of a promoter site means that a gene discovery is likely to happen. On a DNA sequence, splice junctions are points "in which segments of messenger RNA are spliced out." To be aware of these is important for the biologist. WIPO pub. 696(E), p. 35.

70 For more about training neural net, see Johnson-Laird, at pp. 50/51 WIPO pub. 696(E), and H. Collins, [1992] 134(1826) New Scientist 40.

The AI systems in general, as aforementioned, are developed on a crescendo of trial and error. The errors and rejected output are imprinted in a way that a trail is electronically coined. The way the trainer has carried out the training, the patterns of tests have been applied, the facts and the code that simulated the neural network have been inputted. This suggests that creative efforts and investments have been carried out. All of these hidden aspects make up a sort of "cartographic trick"[72], and thus assisting in the indirect identification of the system.

The background just described reflects an intellectual work suffused with challenging barriers, explained by the study of the nature of the innovative process, and relating to the bargain underlying intellectual property.

3. Intellectual property and innovation dynamics

3.1 The intellectual property bargain and competition

3.1.1 The nature of the intellectual property bargain

The early structuring of intellectual property emerged as a result of the liberal ideas behind perfect competition, and the property as a right. Both were vital fuel for capitalism centred in the notion of a contractual relationship between the owners of means of production and society. The rationale for that relationship was as follows. Without private property "no rational economic calculation would be possible"[73], and competition was conceived as a bargaining process for public favour, hence, rendering an unrestricted competition with the notion of society.[74] The focus on these ideas is only to state briefly the historical background within which intellectual property was developed.[75]

As a legal institution born under the influence of the classical economics and exempted from unwanted monopolies, intellectual property was designed to ensure temporary protection "only to the end of promoting science and the useful arts."[76] Early in the current century, this steadily founded theory was

71 This is a summarised description made by Prof. McCarty of a work done by Mick Noordewier, biologist and computer scientist at the Rutgers University, WIPO pub. 698(E), pp. 35/36.

72 Cf. Johnson-laird, *loc cit.*, pp. 52/53.

73 A Radomysler, Welfare economics and economic policy, p. 81 passing, in "Readings in Welfare Economics", The American Economic Association series, vol. XII, 1969.

74 Mary S. Morgan, "Competing Notions of Competition" in Late Nineteenth-Century American Economics", [1993] 25(4) History of Political Economy, 563, at 570 and 580.

75 Since existing literature deals with liberal ideas abundantly, further investigation on them is unnecessary and beyond the purpose of the chapter.

76 United States, The Constitution of the United States of America — Analysis and Interpretation, p. 317,

vastly absorbed by the law of the industrial countries. The conceptual basis was first developed by the British courts. In common law, judges learned that letters patents could be ruinous to the society by affecting the price of commodities. The courts had, however, at least two reasons for tolerating patents: the encouragement of manufacture in the country, thus furthering trade for the good of the nation; and even if not recognised as lawful monopolies, letters patents would be granted anyway by the Sovereign "as a convenient means of raising revenue."[77] Principles and practices which prevailed in the construction of the British patent regulation were incorporated into the US Constitution and law and conferred a true right to inventors.[78] The theory behind the clause of science-and-technology promotion of the American Constitution is read univocally as being for the benefit of both inventors or authors and society at large. The clause calls for a balance between private and public interests, or a bargain[79] between inventors or authors and society.

3.1.1.1 The social bargain theory

The sense of bargain is that somebody's gain is someone else's loss. This gain-and-loss relation is synallagmatic in the sense that inventors and authors on one hand and society at large on the other are placed in a prospective context of both gains and losses. The framework of this quid pro quo underlying the concept of protection of intellectual property is determined by four social objectives:[80] encouragement of innovative activity, inducement to the disclosure of the invention, reward inventors and authors, and inducement towards industrial application. Here it is suggested that the achievement of these objectives is a combination of social welfare and efficiency ends. To what extent these legal pre-conditions are achieved has always been a matter of contention and concern.

3.1.1.2 Encouragement of R&D and inventive activities

Although arguable, the literature regards the incentive for R&D activities as the main justification for patent protection. Several surveys, nevertheless, have showed that the stimulating effect varies according to industry, size of

---

edited by N. Small & S. Jayson, 1964.

77 Great Britain, Board of Trade, "Patents and Designs Acts, Second Interim Report of the Departmental Committee", p. 3. Cmd. 6789 (1946). Darcy v. Allin or Allen (Noy 173) [1602] 74 E.R. 1131.

78 U.S. Constitution, Art. I, § 8, cl. 8. *Abraham L. Pennock and James Sellers v. Adam Dialogue*, US Supreme Court, January 1829, pp 327-335.

79 For the discussion of the protection of intellectual property as a bargain, see "OTA Background Paper" p. 7; Copyright and Home Copying: Technology Challenges the Law, OTA-CIT, 1989, ch. 3; US Congress/OTA, Intellectual Property in an Age of Electronics and Information, OTA-CIT 302.

80 Cmd 6789, p. 3.

firms and traditions.[81] In this respect empirical data does not always tell the same story, for instance, the general pharmaceutical industry appreciably relies on patent, and to a certain extent large firms have a propensity to patenting, however, the impact of this tendency on R&D, varies from country to country.

3.1.1.3 Inducement to the disclosure of the invention instead of keeping it segregated

It is expected that access to patent information may render improvement around the invention, enabling the creation of a substitute product. The disclosure of technical data, hence, provides everyone, combining talents and resources, with the competitive opportunity of making a broader use of the technology. Although the patent is a valuable source of technical information by avoiding duplicative R&D activity, in practical terms, its informational function depends on the disclosure of the real value of the invention. For many firms, patent applications are only filled when it is no longer possible to keep the invention secret.[82] The patent hence works as an additional framework with which know-how or a trade secret is extended and negotiated. In the areas of software and integrated circuits, which are characterised by large use of secrecy and fast technical obsolescence, the scope of the disclosure, where patenting is possible, may be very limited indeed. In fact, the increasing reliance on secrecy in the information technology sector stands alone as a component of business strategies. This was considered with much concern by the CONTU Report.

Over the initial period of 12 years when copyright was made available for computer programs in the United States, the US Copyright Register received only slightly more than 1% of the number of computer programs, developed each year, for registration.[83] While the figure showed a very low interest of the 300,000 programmers in copyright, it dismissed the belief that protection of computer software under registrable copyright would ease great "public access to innovative programs"[84]. The industry made it clear that it would not give up trade secrecy protection and, additionally, it "would fight hard to assert its undeniable continuing right"[85] to secrecy. Furthermore, technical know-how

---

81 G. For some accounts, see Sipa-Adjah Yankey, International Patents and Technology Transfer to Less Developed Countries, p. 10-24, 1987; and H. Ullrich, The Importance of Industrial Property Law and Other Legal Measures in the Promotion of Technological Innovation, [1989] Industrial Property 103-112.

82 See Edmund W. Kitch, The Nature and Function of the Patent System [1977] 20 Journal of Law and Economics 265-290, at 275-278.

83 According to the CONTU Report p. 34, only 1,205 programs was registered from 1964 to January 1, 1977, 971 of them were registered by IBM and Burroughs. By that time, about 1,000,000 were developed each year.

84 CONTU Final Report, p. 34.

85 Idem, idem.

necessary to explore the invention is not always satisfactorily disclosed; this may occur deliberately or because of incomplete or inaccurate patent specification.

3.1.1.4 Reward for inventors and authors

A rewarding profit available for successful inventions or works, to the extent in which the invention is commercially practicable and the work original, is in itself indisputable and includes the prospect of a reasonable return on investments. Such a prospect, from the theoretical point of view, relies on the competitive head start over rivals created by the temporary monopoly right. The reward, however, as an isolated function, is an incomplete view of the intellectual property which is more than "a system created to guarantee income to creators."[86]

3.1.1.5 Inducement to industrial application

No protection will be worthwhile if the invention, design or use of copyright on hi-tech renders no industrial application. While the output stemming from them makes it possible to meet a human need, resources are put at risk at the owner's expense, by joint application, or by means of licensing in return for royalties.

3.1.1.6 Towards the welfare/efficiency perspective

An approach of intellectual property within the perspective of welfare and efficiency necessarily faces a margin of conceptual insecurity reflecting the inaccuracy of existing theories. Avoiding the disputed aspects involving the meaning of welfare,[87] the economic theory has dealt with it in terms of individual preferences and associated it with both economic and technical efficiency.[88] The higher the efficiency of resource allocation, the higher the

86 US Congress, OTA Background Paper, p. 7.

87 In a popular sense, welfare describes the happiness of human beings. In politics, the term welfare state is associated with social justice, i.e., the state has the duty to provide assistance for those people in need. Philosophically, the exercise of individual preferences as a value linked to personal satisfaction is contested. Through the perception of values, which are associated with a process of justification, people understand the world. Some preferences, as that concerning food, for instance, needs no justification, but others do. For some accounts on these notions and welfare economics, see Robin W Broadway & Neil Bruce, Welfare Economics, Blackwell, 1984; Amartya Sen, "Choice, Welfare and Measurement", Blackwell, 1983; Kenneth J Arrow, General Equilibrium (Collected Papers), Blackwell, 1983; American Economic Association, "Readings in Welfare Economics", (papers selected by K J Arrow & T Scitovsky), 1969.

88 Efficiency "relates to the most effective manner of utilizing scarce resources." There is an increase in *allocative efficiency* if "higher scale of output is produced at lower cost." A *technical efficiency* occurs

welfare rate. Focusing on this association (social preferences with efficiency of resource allocation), economics creates a consumers' welfare function as a value-free relationship. Instead of dealing with assumptions based on ethics, justice and political desirability, welfare economics[89] is most concerned with the measurement of efficiency or optimality of satisfaction of consumers' preferences.

As described by K.J. Arrow,[90] the social function is translated into a "constitution" or set of conditions to govern the welfare judgements. Arrow's idea was to transform individual desires into concrete social choices.[91] One procedure inferred from his theory was that no individual alone should be allowed to dictate the outcome. What Arrow's theorem in its entirety means is that no set of rules could possibly and consistently devise that judgement. His theory, nevertheless, proves, firstly, the inherent imperfection of any legal policy regulation on welfare grounds, and, secondly, that a sense of welfare which goes beyond the pure logic of economics does exist. There is, in other words, a strict and a broad sense of economics welfare. How much this broad sense lives up to the concept of welfare entailed in the intellectual property bargain is a matter for later consideration. For now, the study will concentrate on further explanation of the strict meaning of economics welfare. As such, welfare is a function of economics efficiency fully understood in the context of two other notions, market and competition. In order to clarify this point, a brief account on the whole competition context is necessary.

3.1.2 From perfect to imperfect competition

As a straightforward concept, the market describes a relationship between sellers (supply) and buyers (demand) subject to economic laws,[92] for instance, demand tends to increase as the price falls. Owing to individual preferences and income within a period of time, variations occur mostly because demand is a function of consumer income and price levels.[93] On the supply side, within a period of time and depending on the length and

---

when a firm using superior technical process compared to another produces the same level of output using less inputs. OECD, Glossary of Industrial Organisation Economics and Competition Law, p. 41

89 The expression is used to designate the study and evaluation of public policies designed to achieve maximization of human well-being.

90 Kenneth J. Arrow, General Equilibrium (Collected Papers), Blackwell, 1983. The author, an American Professor of Economics at Stanford University, was awarded the Nobel Prize in 1972.

91 K Arrow, ob cit, p. 222-225, heading "the theory of social choice".

92 Edwin Mansfield, *Microeconomics, Theory and Applications*, 6th ed, Norton, 1988, p. 20. In writing this expository section, I have much drawn from Mansfield's work, as well as from Roger D Blair & Lawrence W Kenny, *Microeconomics With Business Applications*, Wiley, 1987.

93 The measure of the sensitiveness of a product demand in a particular market is called in economics *price elasticity of demand*.

characteristics of that period, an increase in a price commodity is likely to work as an incentive for the producer to increase the quantity of commodity supplied. However, higher prices sooner or later act against the demand level. The picture may be changed when the introduction of new technology lowering the production cost, enables the producer to produce more cheaply and increases the quantity of supply. In addition to the interaction of demand and supply the price is an important element. Above an ideal price, part of the commodity supplied is unwanted (excess supply), and below that price demand tends to increase. At an ideal point, there is an *equilibrium* between quantity demanded and quantity supplied, and the price tends to be stable for a period of time. Although very formal, the notion of equilibrium in the model of perfect competition is important for a number of reasons, mainly for guiding a pricing policy (methods used by firms for determining their prices) which determines the behaviour of firms concerning the allocation of resources and shapes the competitive process, where the market is visualised as a relationship among rivals.

The conditions under which firms relate to each other define two broad classes of market: perfect competition on one hand and imperfect competition (monopoly, monopolistic competition, and oligopoly) on the other.

3.1.2.1 Perfect competition

Although appreciably sensitive to the welfare/efficiency claims lying behind the protection of intellectual property, the model of perfect competition is conceptually unfavourable to technical changes. In its static monotony, the model presents the following features: a large number of sellers have the same product; provided that the price is the same, purchasers do not care which seller they buys from, as both purchasers and sellers are so small in relation to the entire market, none of them acting alone are able to affect the product's price. The resources mobility is such that raw materials, for instance, cannot be monopolised; consequently, firms can enter and leave the market freely. There is a perfect share of knowledge regarding prices, technological data, and all the possible uses of the resources, so as consumers, firms and resource owners are able to take the best economic decision at an unfailing accuracy.

3.1.2.2 Pure monopoly, a contrasting approach

Opposite to perfect competition is the situation of pure monopoly, where "there must exist one, and only one, seller in a market."[94] The two states (perfect competition and pure monopoly) move from a point of a market

94 Mansfield, p. 280.

impersonally defined by a myriad of suppliers to the extreme of a market personality[95] based on a sole supplier. These theoretical models are so formal that one could hardly adopt one or other as a permanent policy. Nevertheless, monopolies occur for different reasons, some of them being that a single firm may:

- *control the entire supply of a basic input that is required to manufacture a given product;*
- *become a monopolist because the average cost of producing the product reaches a minimum at an output rate that is big enough to satisfy the entire market at a price that is profitable;*
- *acquire a monopoly over the production of a good by having patents on the product or on certain basic processes that are used in its production;*
- *become a monopolist because it is awarded a market franchise by a government agency. The firm is granted the exclusive privilege to produce a given good or service in a particular area.*[96]

Monopolies have the ability to change market conditions by affecting prices and output. Economists believe that under monopoly the use of resources tends to be less effective than under perfectly competitive industries. In the latter situation, output tends to be greater and prices lower than under monopoly.[97] One of the means through which monopoly may act is price discrimination, which economists regard as socially inefficient, but which is sometimes recommended.[98] Due to some type of indirect competition, however, monopolies rarely hold their position in the long run, giving room to intermediary market forms, such as *monopolistic competition* and *oligopoly*.

3.1.2.3 Monopolistic competition

Three conditions define monopolistic competition: the existence of a large number of firms, producing and selling similar products, and having the same level of demand and cost. For the sake of economic theory, firms producing similar products are arbitrarily grouped. Each firm has a degree of monopoly power over its own product, but not enough to enable the firm to

95 "The firm in a perfectly competitive market - says Mansfield - has so many rivals that competition becomes impersonal in the extreme; the firm under pure monopoly has no rivals at all." Ob. cit., p. 281.

96 Mansfield, ob. cit., p. 281, 282. The second situation (the competitive advantage of minimum cost production) above defines the so-called natural monopoly.

97 Mansfield, p. 297.

98 When a firm sells a commodity at more than one price, or sells similar products at prices in different rations to marginal costs, it is said that price discrimination occurs. Discrimination however is needed if without it the good can hardly be produced. See Mansfield, p. 301, 312.

threaten rivals. Each competitor's product is a little different from the others'. The variation is based on several elements, such as physical make-up and brand names, making the products or *dresses* very close substitutes.[99]

The model of monopolistic competition is supposed to operate under a degree of inefficiency, but close to perfect competition. From the above conditions one could infer that under monopolistic competition deterrence to entry is rather weak compared to an oligopoly industry.

3.1.2.4 Oligopoly

The main features of an oligopolist market are: a small number of firms (not necessarily large ones), great independence among them, and each firm's policy is likely to affect the other rival firms.[100] The oligopolistic firms tend to make entry difficult and pursue an economy of scale. Various theoretical models have been developed to explain the oligopolist behaviour. The duopolist equilibrium of output says that each firm tends to make profit-maximising choices on the assumption that the other competitor will not respond to change in output. The price rigidity theory assumes that a price cut by an oligopolist is likely to be followed by the others; conversely, competitors most unlikely change their prices to respond to an individually taken price increase.

Although these theoretical approaches do not take any form of collusion into consideration, oligopolist industries tend to come into collusive arrangements in order to increase profit, fight uncertainty, and make entries uneasy. Cartel arrangements designed to set price uniformization, distribution of sales, or to divide up a market, however, tend not to last for long because sooner or later firms are likely to cheat and breakdown the collusion.[101] This flows to the game theory which explains how decisions are made in the oligopoly environment where conflicts and co-operation take place. The competing game requires each player to set up its dominant strategy, and this sometimes includes cheating the other cartel members by cutting price, for instance.

Pricing policy under oligopoly is often guided by a dominant oligopolist who tends to determine the price of technology by negotiation[102] rather than on

---

99 Typical monopolistic competition include toothpaste, food, shoes, clothing, and furniture industries.

100 Some of the US oligopolies are IBM and Microsoft in the IT industry; GM, Ford, and Chrysler in the automobile industry; and GE and Westinghouse in the electrical equipment industry.

101 Based on this competitive behaviour, one may believe that the market itself is able to self regulate, thus making government intervention unnecessary. Historically, this conclusion has not been proven true.

102 See Yoo Soo Hong, UNCTAD/ITD/TEC/3, 12 Feb 1993, p. 35.
Price negotiation may take VER (voluntary export restraints) form, or bilateral agreement. In Europe, it is estimated that VERs cover 30 per cent of international trade in electronics. M.M. Kostechi, [1991] 14(4) World Competition 32. Warning about the debatable legality of such arrangements is found in

the basis of competitive market principle. The decision of the price leader affects the rest of oligopolist firms, and may work to bar entry. Barrier to entrants depends on the market size. Limit pricing may discourage newcomers to invest millions of dollars in order to establish and maintain, for instance, a sophisticated and modern foundry of integrated circuits. Entries, nevertheless, are not impossible in the long run.

3.1.2.5 Entries versus theory of contestable market

The theory of contestable market says that at a given time in a market there is a vulnerability to entry. Under the threat of newcomers, firms tend to behave as perfect competitors. They stop attempting to collude so as to prevent prices from rising, otherwise entry would be affordable.[103] The existing firms, however, may not be prepared to engage in a price-cutting policy which would work in the opposite way, i.e., would bar entry, but also would lead to a pricing war within the monopolistic industry, a dangerous and also unwanted outcome. All of these points lead to the assumption that in conditions of imperfect competition, firms are likely to behave in a way that affects price, output, and profits. When such a behaviour is coupled with the use of intellectual property right, the effects, although difficult to predict, will head to a loss of efficiency and welfare.[104]

3.1.3 Social welfare and efficiency

Turning now to its restrict concept, welfare is described as a measure of *consumer's supply*, i.e., "a net benefit received by the consumer."[105] Such benefit is translated into greater quantity of commodity the consumer is supplied with for the lowest price the producer can possibly charge, given certain conditions of supply and demand within a market and a period of time. While on the demand side the conditions are chiefly dictated by consumer preferences, the supply is considerably related, *inter alia*, to costs. These may include expenditure on R&D activities and royalties paid for intellectual property licensing. It is now understandable that increased consumer supply is a benefit arising from competition. Applied economics has developed fairly secure methods through which a learned technician is able to calculate the

---

"GATT Activities 1989", p. 18. See "Japan-Trade in Semi-Conductors" *in* GATT/BISD, 36th Supp. 1990, p. 116-163.

103 As Mansfield states, "if existing firms are charging a price in excess of marginal cost, it is profitable for an entrant to undercut the price of the existing firms." Ob. cit., p. 358. This describes the market contestability.

104 Mansfield, ob. cit., p. 359-362.

105 Mansfield, p. 100.

effects of a business practice on consumer supply. The theory, however, tells very little beyond the economic logic.

Theoretically, a static model of perfect competition assumes that the interrelated markets for all products are in a general equilibrium. At such a point, it is said that the firms apply the best combination of resources at the lowest cost, thus leading to maximisation of profit and utility. Efficiency, then, is synonymous with optimality. Optimal efficiency, in other words, is a concept which describes an optimal allocation of resources.[106] In order for that unreal world of general equilibrium to exist, it is assumed that consumers exercise different levels of preferences and consume, but the utilities flow from consumer groups to others without affecting the overall level of demand. This efficiency in exchange is a necessary condition for general equilibrium. Knowing all products with an unfailing accuracy, consumers are able to exercise a perfect substitution of products for others. This is another condition, i.e., efficiency in product substitution. A third condition for the general equilibrium relies on the efficiency in production, that is, the optimal allocation of resources remains unaffected, so the overall level of supply or production also remains unchanged. The whole picture gives a sense of optimal welfare distribution. Although too formal, the model provides for some practical lessons. One is that the concept of social welfare goes beyond the measurement of individual preferences.

In view of those three conditions of efficiency (efficiency in exchange, efficiency in product substitution, and efficiency in production) a situation called grand utility possibility frontier is created.[107] At this point it is said that some people have increased their utility to the maximum at the expense of the reduction of the utility of other people. The welfare frontier is an imaginary point representing the maximum well-being a person can enjoy "given the level of welfare enjoyed by the remaining members of the society."[108] Theoretically, it is not possible to establish the maximum point of the frontier, but outside of it no point is possibly attainable by society.

A situation of grand utility possibility frontier provides no more than a sense of optimal welfare distribution; it fails to establish a fair meaning of interpersonal satisfaction. A lesson can be drawn, nevertheless, which is that social welfare is desirable and is a function of consumer utility and resource

106 The concept was developed last century by Vilfrido Pareto (Pareto Optimal). Blair & Kenny, ob. cit. p. 457.

107 The economic analysis of the social welfare function is based on a model which takes into account a pair of goods and of consumers. Indifferent levels of the distribution of the commodities to both consumers are discussed. These levels represent a range of possibilities under optimal conditions of distribution of the total of the available quantities of the goods. See Mansfield, ob. cit., p. 474; Blair & Kenny, ob. cit., p. 465/66.

108 OECD, Glossary of Industrial Organisation Economics and Competition Law (1993).

allocation. An attempt to incorporate a sense of fairness into that functional relationship leads to the assumption that society as a whole is better off when a degree of utility is allocated from a consumer group to another. Scientific criteria, however, do not exist to guide a fair distribution of utility or income. This is arbitrarily developed either by a dictator or a parliament through a democratic process (majority rule). This suggests that a safeguarding policy affecting the exercise of the owner's right based on welfare and efficiency ends could hardly follow rigid criteria.

In so far as it is pursued, efficiency measure is supposed to conform to welfare, but other grounds are available to back the limitation of the exercise of intellectual property rights in the name of social welfare. The state knows to what extent a system should limit the use of intellectual property on basis of social welfare rather than efficiency, as much as it knows how heavily the middle class should pay taxes for the benefit of social welfare. One knows, therefore, that technological progress is desired to increment the level of community prosperity.

The maximisation of welfare and efficiency in a static sense cannot be fulfilled unless it is in conditions of a fixed level of technology. "That is, — Mansfield says — they show how inputs and commodities must be allocated if welfare is to be maximised, *given a fixed level of technology*. It is possible that an allocation of inputs and commodities that violates these conditions might lead to a higher level of consumer welfare than any allocation that meets these conditions, because it might result in a faster rate of technological change and productivity increase."[109] In this respect, it is suggested, "a perfectly competitive economy is likely to be inferior in a dynamic sense to an economy including many imperfectly competitive industries."[110] It follows that the introduction of new technologies is required to push forward the frontier of utility which in turn creates a paradox. The technical change, although desirable to the extent which it promises a new dimension of welfare and efficiency frontier, leads to an innovation process developed under conditions of imperfect competition where some degree of welfare/efficiency losses are greatly expected, if not unavoidable. It is on that paradoxical prospect of gains and losses in welfare and efficiency that both protection and limitation of intellectual property encounters the best justification from the economic rationale.

---

109 Mansfield, ob cit, p. 552.

110 Mansfield, idem. Economists do agree that competitive markets potentially favour greater social welfare, but monopoly is not necessarily bad. An effectively productive monopoly has the ability to operate technical change bringing a prospect of higher social welfare. In view of this, doubts exist whether perfect competition is desirable. Since such an unreâl model cannot be achieved in all markets, a sort of "workable competition" would be the target. But there is no consensual criteria to define this.

4. The nature of the innovation process and policies

4.1 Innovation under uncertainties

The uncertainties affecting the innovation process are not *per se* detrimental to the intellectual property bargain. The assumption made here is that in order to cope with uncertainties, innovative firms are likely to develop strategic behaviour,[111] the effects of which may ruin the intellectual property bargain.

Economists do not contend the uncertainty as an element of innovation activities.[112] Studies available on the matter are based on empirical analysis. Although these empirical and statistical studies are criticised for lack of completeness of information on which researchers elaborate, one survey published in Great Britain[113] is a very illustrative source from which the following assertions are briefly drawn.

The innovation process is described as being inherently surrounded by risks. Although low in 'adaptive' and 'imitative' types of project, the rate of uncertainty is reported to be considerably high. Three categories of uncertainties are identified; they relate to technical matters, market, and general business. These two latter categories are based on management of technology, involving a team of specialists with knowledge in interdisciplinary matters, including business affairs and potential demand forecast. The technical uncertainty "lies in the extent to which the innovation will satisfy a variety of technical criteria without increased cost of development, production or operation."[114] Uncertainty of this kind is normally associated with integration of R&D and manufacturing, product and interface standards, and product liability.

Uncertainty may be minimal, for instance, in "adapting electronic circuit designs to novel applications, but well within the boundaries of established technologies, or minor modifications of existing designs."[115] However, in general the scale of uncertainty is such that, it is argued, "most firms have a

---

111 In the course of Part One, it will be clear that the uncertainties of the innovation process themselves make for a strong point to claim protection of intellectual property. This is out of the question.

112 See F M Scherer, Innovation and Growth pp 94, 182, MIT, 1984.

113 Christopher Freeman, The Economics of Industrial Innovation, Pinter Publishers, reprinted in 1991. Freeman is a well known senior researcher of the University of Sussex, Science Policy Research Unity — SPRU. His book comments on a considerable number of surveys carried out in Europe and the United States.

114 Freeman, ob cit, p 149.

115 Freeman, ob cit, p. 151.

powerful incentive most of the time not to undertake the more radical type of product innovation."[116]

Due to these uncertainties, the bulk of investments concentrate on less risky projects accounting for minor improvements,[117] and profit-maximising firms are likely to develop strategic trade practices leading to block technical information. For instance, in a situation where the firm can make some profit by ultimately licensing to other firms the process or innovation, "there may be a deliberate preference for secrecy and not licensing."[118] Resorting to government assistance is another means of greatly reducing the technical and market uncertainties.[119] Yet, the remarkable, and to a certain extent debatable, finding is that high investments in radical long term innovation are likely to be confined to large firms enjoying oligopolistic competition.

4.2 Innovation under the context of imperfect competition

As aforementioned in section 2.3.1, perfect competition leads to maximisation of welfare and efficiency. The logical assumption to draw from the preceding discussion is that under imperfect competition a degree of loss in welfare and efficiency is expected, if not unavoidable. The evidence that the innovation process is carried out in a context of dynamic imperfect competition allows another assumption, that is, the more the innovation process is encouraged, the more incremental losses will be expected in a certain period of time. It thus follows that the innovation process entails a threat to the intellectual property bargain. The explanation made in the previous section has proved this remark to be true on a theoretical level. Large market share in itself, it may argued, does not upset the intellectual property bargain. Being large, however, means being able to capture economic resources and monopoly position thus holding the power to influence or manipulate market forces. Even being strategically advantageous or necessary to shield investments from the risks of innovation activities, such ability is *per se* a cause for concern about the achievement of the intellectual property bargain. The task now is to show the empirical evidence related to the environment of imperfect competition under which the innovation process is carried out.

---

116 Idem, p. 150.

117 Freeman, ob cit, p. 162.

118 Freeman, ob cit. p. 163.

119 On the government role in the innovation process see below in this Chapter, heading "The syndrome of the technical capability and policies".

4.2.1 The phenomenon of new competition

The case that innovation dynamics reflects an observable reality of, or leading to, an imperfect competition is historically supported by the termed "new competition" phenomenon.[120] In the middle of the last century, economic theorists were unfamiliar with the idea of industrial monopoly (large-firm competition). Economic studies were predominantly centred on the classical model of perfect competition opposed to monopoly. As that theoretical model did not expiain the behaviour of large-scale firms, economists of that period viewed the firms' "trustification" as an emerging reality which required a new economic theory of competition. Accounts on the nature of that phenomenon in America, focus on the growth in concentration and oligopolies of the late nineteenth century as a trend "associated in the contemporary mind with greater efficiency and lower prices," dominating large industrial sectors. This impressively challenged the economists' "perceptions of the nature of competition."[121] In fact the growing number of combinations, which appeared during that period as a result of the free competition, was later confirmed as part of a complex competitive reality which today's legal policies recognise and are designed not to condemn or revert but to control.

By establishing the first large-scale industry development of the late nineteenth and early twentieth centuries as marking the emergence of a "new competition", historians have contributed to the explanation that the phenomenon of intermediary imperfect competition is associated with the wave of the today's technical pace. This relationship between market structure and innovation, first suggested by Schumpeter followed by Galbraith,[122] illustrates that an imperfectly competitive economy will satisfy the conditions for a higher rate of technological change. There is, however, some controversy regarding the extent imperfect market is conductive to technological innovation. No analyst, however, has denied Schumpeter's proposition.

Holding a pessimistic view of the Schumpterian perspective, Scherer says that "rivalry normally accelerates the pace of technological research, development, and innovation, as long as the number of firms competing is not excessive." He, then, establishes his sense of balance by adding: "what is needed for a rapid rate of technological advance is the proper blend of competition and monopoly."[123] Concrete evidence is provided by Freeman, addressing the role of the firm's size's in the innovation process.[124]

---

120 See Mary S Morgan, Competing Notions of 'Competition' in Late Nineteenth-Century American Economics [1993] 25(4) History of Political Economy 563-604.

121 M Morgan, loc cit, p. 564, 565.

122 Joseph Schumpeter's work, "Capitalism, Socialism, and Democracy", was published in 1947 and is frequently cited by modern analysts. Similar strand was developed in 1952 by J. K. Galbraith in his work "American Capitalism".

Small firms established by inventor-entrepreneurs have made some good contributions "in the early days of the chemical industry, and the early days of the semiconductor and radio industries" and continue "to flourish in the minicomputer industry and in computer software."[125] The contribution, however, varies greatly from industry to industry and according to the level of innovation. Concerning the American semiconductor industry in particular, it has been pointed out that small firms have played exceptional role thanks to tactics of "technological entrepreneurs bringing with them ideas and half developed new products from a scientific environment in universities and government laboratories." However, when referring to "key innovations" large corporations continue to predominate.[126]

The contribution of small firms in types of innovations, such as "complex engineering products for which more than 10,000 components may be needed", including telephone exchanges and large computer systems, is beyond their resources.[127] In electronics, for instance, the "fairly significant contribution" British small firms have made is in printed circuit board for the electronics industry.[128] This consorts with the general assumption that in Europe, as in Japan, the innovation process has been greatly dominated by large corporations.[129]

Enjoying advantages such as more access to finance, ability to cope with government regulations, and specialist management expertise, large firms are more prepared to engage in long and costly R&D projects. This has been proven to be just as true in Europe as in the United States. Conclusive evidence from a study for the OECD shows "that the vast majority of small firms in OECD countries do not perform any organised research and development."[130] Similarly, a survey about R&D in America also suggested that "there is some tendency for R&D intensity to increase with size of firm with the largest size-groups."[131] The scale of research and development may suggest some relationship with patenting as a measure of scientific output. In this respect, information has not been found reliable, but has provided interesting findings.

123 F M Scherer, Innovation and Growth - Schumpeterian Perspectives, pp. 114, 127, The MIT Press, 1986.

124 C Freeman, The Economics of Industrial Innovation, 1991, chapter 6.

125 C Freeman, ob cit, p. 131. It is conventionally regarded as a small firm that with 200 or less employees.

126 C Freeman, ob cit, p. 138.

127 Idem, idem.

128 Idem, pp 141-143.

129 Idem, p. 138.

130 C Freeman, ob cit, p 132.

131 The survey conducted by Soete was published in 1979. See C Freeman, ob cit, p. 134.

4.2.2 Large firms' behaviour towards patenting

While "some firms attach great importance to patents and have large departments with a strong interest in patenting activity", others "either do not want to bother with patents or prefer to rely on secrecy", postponing filing patent applications.[132] Large firms, as a general assumption, are more strongly interested in patenting, confirming the historical view that patents represent a strategic tool in a large firm's hands. By 1945 in Britain, for instance, electrical engineering, chemical and pharmaceutical industries "accounted for 60 per cent of all patents."[133] The assumption of the large-firm propensity to patenting, however, is not plainly supported. Surveys carried out in the United States and Britain have suggested that propensity to patenting is higher among small firms. The conclusion is based on the fact that large firms depend on "patent sharing and know-how exchange arrangements" and small firms, in contrast to large ones, who "usually cannot afford not to patent and cannot afford to wait."[134] These studies just referred to did not take into account copyright and chip-designs, two very considerable forms of intellectual property protection in the field of information technology. The force of the surveys' outcome is thus very limited. They do not alter the monopolistic aspect of the intellectual property concentration as part of the nature of the innovation process. As a general rule, such a concentration is not only a reality at a firm level, but also observable at the spatial level of industrial structure.

4.3 Industrial structure and technical innovation

It has been assumed that barriers to entry is an element which works against efficiency and welfare. Due to the uneven nature of technical change, the innovation process has the effect of forming a structural barrier to entrants, thus, threatening the intellectual property bargain. The formation of this potential deterrence is now considered.

The analysis of the relationship between innovation and industrial structure has led theorists to compare the technical diffusion to a wave motion. Diffusion follows waves of development prospects determined by social and economic conditions, which vary from region to region. A consequential outcome is that technical changes are accelerated in selected industries or regions, and set back in those sectors and regions adversely affected by lack of adequate conditions.[135]

132 C Freeman, ob cit, p. 136.

133 Jonathan Liebenau, Patents and the chemical industry: tools of business strategy, in "The Challenge of New Technology, Innovation in British Business Since 1850", 135 at 136, edited by J Leibenau, Gower, 1988.

134 C Freeman, ob cit, p. 136.

A study of the industrial structure related to innovation in the United Kingdom has also confirmed the exacerbation of regional disparities associated with uneven technical diffusion.

It has been suggested that the unbalanced technical development is not simply a matter of the concentration of innovation activities. The reality is that technical revolutions induce instability because it is impossible for all regions to develop even rates of technical capability simultaneously. The fatality of capitalism is stated in these terms: "the constant drive to raise profits, the anarchy of the market and the inability to plan production in consonance with the market all lead to uneven development between individual firms."

Resulting from an accumulation of conditions such as a skillful work force and competitive muscles, the disequilibrium is a determining factor in the nature of the innovation activity as a process of gains and losses. Some enterprises of different regions lose out at the expense of others in the same product market.

The stigma of the imbalance of industrial structure in the OECD areas has also been discussed. Showing his concern in this respect, a representative of Japan stated:

If technological innovation were to take place uniformly in every field, there would be no problem. However, advanced technology innovation is bound to centre on selected industries; there will inevitably be a lack of equilibrium in the development of industries due to the time lag caused in the process of the spread of technological innovation from one industry to another. The present situation is causing a domestic and international disequilibrium in structure between the field which remains in the dark and the field which is in the limelight and where technological innovations are rapidly taking place and towards which capital and human resources gravitate.!138

Two contributing factors to that imbalance and particularly associated with information technology are speed of technical change and economy of scope. For instance, in the case of personal computers with potential applications to industrial use, the lapse of time for upgrading performance has become shorter than a twelve-month period. Furthermore, describing the

---

135 Carlota Perez, "Microelectronics, Long Waves and World Structural Change: New Perspectives for Developing Countries" [1985] 13(3) World Development 441.

136 Ash Amin & John Goddard, "Technological Change, Industrial Restructuring and Regional Development", p. 3, Allen & Unwin, 1986.

137 Amin & Goddard, ob cit, p. 2, 10. In order to tackle the problem, state intervention is contemplated on the assumption that "what is happening due to the operation of market forces in the growth areas can be reproduced through public interventions in the crisis regions." The authors, nevertheless, do not take it for granted.

138 The statement has been made at the OECD forum by Mr G. Takanashi when he was Chairman of the Fair Trade Commission -Japan "Competition Policy and Technological Innovation", p. 23.

technical speed in the computer business, an IBM representative testified in these terms: "the art is growing and changing with blinding speed that if the automobile industry had progressed on the same curve as computer in the fifteen years, we would now have been able to buy for twenty dollars a self-steering car that would attain speeds up to four hundred miles per hour and be able to drive the length of California on one gallon of gasoline."[139] As to the economy of scope, the impact[140] is on production management, requiring a ready response.

Economy of scale (single production line of uniform products) is based on cost-efficient large-scale investment in production facilities, mass production and mass sales of standard or homogeneous products. Yet, today development of microprocessors has made possible *production management of different products on a single production line* a possibility.[141] The management of this economy of scope includes:

- *collection of information about consumer demands at point of sales (POS);*
- *analysis of the customer data by POS computing system; and*
- *data communication from the distribution system to manufactures.*[142]

The features of the economy of scope is that it allows prompt identification of diversified demands, accommodation of consumer needs through manufacture of different (related) products, an increasing variety of business opportunities, and full operation of the small and medium-sized firms capabilities. This dynamic environment illustrates a performance only attainable by selected technologically equipped industrial segments. In order to tackle distortions of this kind, the limitation to intellectual property seems to be a valid assistantial policy, and in this respect the role of the state has been rather noticeable.

4.4 The syndrome of the technical capability and policies

The theoretical and economic background has been developed to support the existence of welfare and efficiency claims framing the underlying intellectual property bargain. As much as this bargain is associated with technical change relying by definition on an unstable economic structure, the

139 Testimony of Ralph Gommery, CONTU Report, p. 35.

140 The rapid development in the fields of hardware and software has been identified as giving rise to problems of compatibility or interoperability of equipments. See Karl H. Pilny, *Legal Aspects of Interfaces and Reverse Engineering - Protection in Germany, the United States and Japan, [1992] 23(2) IIC 196.*

141 The text follows the explanation stated by G. Taghanashi, OECD Report W.00050/D.390, 2473, p. 22.

142 Idem.

welfare and efficiency ends anticipated by the protection of intellectual property are kept under impairing conditions. While the economic rationale makes a case for limiting the use of intellectual property, it is now argued that the limitation depends on the assistantial role of the state acting along side the private enterprises. Control on IPR use is not, or should not be, an isolated policy. The point here is this: if technical capability and change are part of a country's policies and law, to not ensure proper protection of intellectual property does not seem logical. The social objectives of protection however are hardly achieved, unless IPR use is controlled. Neither does it seem logical to not have a policy to safeguard those objectives. Concerning incentive to innovation, protection and safeguarding policies together are only meaningful within a complex arrangement where state and societal acting forces work together. The observations of the way the incentive for new technologies are organised support and justify these remarks, as well as the degree of state intervention. The question now is how much do state assistantial policies matter.

They matter where they rectify the defects of the market economy and complement it, so as to respond to the Nation's will to catch up with, or to maintain leadership in technology by supporting innovation strategic activity; to ensure that firms act, and society's resources are free from undue restraints; to preserve or promote social welfare by making the improvement of living standards possible.

4.4.1 Justification of state intervention

Although plainly justified in the light of modern liberalism,[143] government intervention has always been a very controversial matter due to the distortive effects it may have.[144] The influence of the increasing role of the government, nevertheless, in the creation and diffusion of new technologies is very strong.[145] In exercising influence, as a consumer, regulator or underwriter, the state acts either in partnership with the industry, or by leading actions to create conditions for industrial development and competitiveness.[146] The ways in which state support is organised vary from country to country and depend on historical and contextual reasons.

---

143 R. Eccleshall, "Liberalism", pp. 37-78, in *Political Ideologies - An introduction*, 1984.

144 For some accounts see comments by Yoo Soo Hong in the "Report of Ad Hoc Expert Group on Technology Policies in Open Developing Country Economies", p. 33-37, UNCTAD/ITD/TEC/3, 12/FEB/1993.

145 See abstracts of significant articles appeared in the period of 1972-1991 in [1993] 22 Research Policy 101.

146 For details about the US policies for the incentive of new technologies, see "*1990 DOC Technical Survey*" [footnote 3] and John Street, Politics and Technology, Macmillan, 1992.

4.4.2 Four examples of state intervening partnership

In Japan, for instance, the coalition of state and business contrasts with the partnership of business and labour in Germany.[147] This relationship between government, business and labour is qualified by both, the catalyst function of the state and societal commitment. In some cases the state assumes a dominant position, such as in France, distinguishing from the business-dominant system of the United States.[148] In any case the state acting alone, i.e., without private alliance, would hardly conform to the liberal ideal.

A commitment and a choice rest in the core of that alliance. Both state and citizens are aware of the technological dilemma, that is to miss the technological race seems to jeopardise the welfare of the Nation. Conversely, the risks of sharing the race are several, at least in short term. Some welfare and efficiency losses may occur due to disruptions in market structure, and a number of jobs may be put at stake because of the displacement caused by automation. The consent to technology, if it occurs, invariably leads to a syndromic technical capability characterised by the particular attitude of the country as a whole to catch up with, or keep the leadership in technology and competitiveness. Such an attitude is reflected, for instance, in government policies and law. Three examples illustrate the point.

In passing the American Technology Pre-eminence Act of 1991, designed to speed technical development and maintain economic competitiveness, the US Senate stated that the decline in both technological leadership and market share of the US industries could not be allowed to continue in prejudice of the "Nation's standard of living."[149] The desire for technical leadership was also expressed in the High-Performance Computing Act of 1991,[150] thus shaping the American technology policy. One consequence of this policy[151] affecting the use of intellectual property is that the title to any intellectual property arising from joint R&D programme supported by the government shall vest in, and cannot be transferred except to, a company incorporated in the United States. The legislation also outlines a range of administrative measures and Government-funded programmes which are part of a complex framework.

---

147 Jeffrey Hart, "The Effects of State-Societal Arrangements on International Competitiveness: Steel, Motor Vehicles and Semiconductors in the United States, Japan and Western Europe", [1992] (22) British Journal of Political Science 255-300.

148 Jeffrey Hart, loc cit.

149 P.L. 102-245, H.R. 1989, Senate Report No. 102-157. The ATPA 1991 traces the national needs in technology, sets out programmes, allocates funding, organizes the technology administration, and refers to other four Acts which form the legal framework of the US technology policy.

150 Public Law No. 102-194, S. 272.

151 Senate Report (Commerce, Science, and Transportation Committee) No. 102-157, p. 17, Sept. 24, 1991.

of this policy[152] affecting the use of intellectual property is that the title to any intellectual property arising from joint R&D programme supported by the government shall vest in, and cannot be transferred except to, a company incorporated in the United States. The legislation also outlines a range of administrative measures and Government-funded programmes which are part of a complex framework.

The best example to illustrate the syndrome of technical leadership in Japan is in the integrated circuit business, where the Japanese industry is regarded as a strong rival to the United States'. As soon as very-large scale integrated circuits (VLSI) appeared in the 1970s, — J. Hart comments —.

it became policy of both the major Japanese firms and the Japanese government to beat the Americans in process technology so as not to be dealt out of the competition in VLSI products. The government committed itself to this enterprise not just because it was concerned about semiconductors, but also because it believed that overtaking the United States in semiconductors was the key to improving Japanese competitiveness in all major downstream industries such as consumer electronics, computers and telecommunications equipment. Thus, in the transition from LSI to VLSI in semiconductors, the connection between state-societal arrangements and technological innovation was extremely clear.[153]

The strong desire for a rapid economic growth has not only been a Japanese post-war commitment set up by the government and businessmen, but also a "central political goal to which all other Japanese policies have been subordinated,[154]" including the intellectual property policy which has become a weapon Japan's developmental system.

4.4.3 An account of the Brazilian ground rule

In Brazil, the Nation's will for catching up with technology is part of the constitutional framework,[155] creating a state duty and preordaining objectives and means. While the responsibility for developing and commercially applying new technologies lies within the private sector, the Constitution charges the State with the duty to promote and foster scientific development, research and

152 Senate Report (Commerce, Science, and Transportation Committee) No. 102-157, p. 17, Sept. 24, 1991.

153 Jeffrey Hart, "The Effects of State-Societal Arrangements on International Competitiveness: Steel, Motor Vehicles and Semiconductors in the United States, Japan and Western Europe" [1993] (22) B.J. Pol. S., at 281.

154 Michael Borrus, *Macroeconomic Perspectives on the Use of Intellectual Property Rights in Japan's Economic Performance*, in "Intellectual Property Rights in Science, Technology, and Economic Performance", p. 261 at 264, edited by F Rushing and C Brown, 1990.

155 1988 Constitution, Articles 218 and 219.

Despite the non self-executing character of the constitutional provisions,[155] the clear fundamental purpose has been to create the foundations of a state-societal covenant shared by the State and the society at large, expressing a strong desire to catch up with technology and to develop a technological market. The institutional agencies, nevertheless, have so far failed to respond to these economic and social ambitions effectively.

Brazilian technology policy has always been implemented under a carrot basis, i.e., on the basis of financial incentive and market restrictions which, in the recent past, put the country under severe foreign pressure. Set up under the self-reliance assumption, market restrictions were much criticised.[156] Due to the dynamic nature of high-technology, IT policy has fallen far behind the technical pace. There is no precise reason for this failure. Strong suggestions however refer to either lack of confidence or will of foreign firms to transfer advanced technology, or the State's subsidising policy being carried out on a carrot basis rather than on a carrot-and-stick basis under which some penalty would be imposed upon national firms for not pursuing technical capability. Moreover, the targets set up in the informatics programme (National Plan of Informatics and Automation — PLANIN), *per se* too ambitious, have never been met.[157]

Apart from some optimistic views,[158] a general feeling of doubt among the business sector has always existed as to the efficacy of the IT policy, coupled with the belief that to promote high-technology foreign co-operation is indispensable. A major inconsistency or weakness in the overall policy, therefore, has been in not enacting proper protection for intellectual property.

It is part of the constitutional covenant that Brazilian and foreigners, as provided in law, shall be ensured temporary protection of rights on works and industrial inventions (including any intellectual property with industrial applications). As to the latter, protection aims at both social interest and technological and economic development of the country.[159] The rationale underlying such a clause is that protection of intellectual property is mandatory to the degree it works for the welfare of the Nation. In this regard, the lack of a steady intellectual property policy has made the country ill-equipped to

155 M. G. Gonçalves Ferreira Filho, "Fundamental Aspects of the 1988 Constitution", pp. 11-25, *in* A Panorama of Brazilian Law, Dolinger & Resenn (ed), North-South Center/EEL, 1992.

156 Gallangher, The United States-Brazilian Informatics Dispute, [1989] 23(3) The International Lawyer 505; Ellene Felder & Andrew Hurrell, "The US-Brazilian Informatics Dispute", FPI/School of Advanced International Studies, 1988.

157 Law 8,244, 16 October 1991.

158 For some accounts, see Hubert Schmitz & José Cassiolato (eds), "High-Tech for Industrial Development — Lessons from the Brazilian experience in electronics and automation", Routledge, 1992.

159 1988 Constitution, Article 5 (XXVII and XXIX)

comply with the fundamental agreement and the Nation's syndromic pursuit of technological autonomy.

4.4.4 Guidance to the state catalyst function

The challenge in limiting the exercise of intellectual property rights rests on the country's ability to combine policies, i.e., to safeguard the social objectives behind the protection of intellectual property as a component tunefully integrated with the overall policies put forward as a means of fostering the development of new technologies and technical change. In connection with this, two principles guide the function of the catalyst state: surveillance as a means to improve, in terms of welfare and efficiency, the state assistantial machinery; and planning as a process to justify legal measures or reliefs which especially affect the use of intellectual property.

Surveillance is necessary to the extent that it makes state action effective by capturing the best opportunity to act, while the process of justification makes the implementation of the legal policy, in a particular situation, reasonably acceptable to the parties concerned. The force of these principles relies on the need to identify circumstances where welfare-improving interventions are likely to be feasible in practice. In dealing with this, in Part Two the study identifies the intellectual property policy as a complementary instrument to explore concrete possibilities of improving welfare and efficiency.

Especially in the information technology sector, the policy is translated into legal measures made available to facilitate new entries, to increase the bargaining power between rivals, to discourage abusive behaviour, to encourage regional development, and to foster high-performance computing for the improvement of state services such as education, public transport, national health, and basic and applied research. In this respect, a balancing protection of chip designs should be included as a state commitment.