Part I

Patent Systems and Patent Law

Chapter 2

Introduction

The first part gives a short overview over general aspects of the patent system. Especially because the patent system as such is not the primal focus of the book, it is important to understand how the different parts interact in the whole patent system. Only then, results derived from theoretical models that are introduced in the second and the third part can be meaningfully translated into realistic policy recommendations.

Chapter 3 briefly relates the history of the patent system from two different perspectives. First, a general overview over the evolution of the patent system is given. This not only shows us which events shaped the system. It reveals also its strengths and weaknesses, and it reminds us why certain provisions have been introduced. This knowledge is especially useful when evaluating the relative performance of the patent system. This chapter also pays special attention to international patent treaties which form the background for the third part of this book. Currently, three treaties exist which are open for every country to join. Each of them have different aims.

Chapter 4 introduces the two most influential schools of thought that have frequently been used to justify the existence of patent systems as well as their design. Economist are trained to adhere to a utilitarian point of view where welfare is the universally accepted standard according to which all changes in the whole economic system are measured. Utilitarianism has exerted less influence in other areas of social sciences where non–utilitarian schools of thought are more influential. Introducing the natural rights theory as a justification for patent systems in addition to the utilitarian patent theories helps to understand a lawyer's perspective towards intellectual property rights.

Chapter 5 aims at familiarizing economists with the main legal issues of patent systems. This is accomplished by tracing an invention through the whole system, i.e. from filing the application to a judge's ruling in infringement or validity cases.

Subsequently, attention is drawn to patent law and its economic effects, the principal subject of the book. The main provisions governing various aspects of a patent are treated as legal policy instruments, i.e. the main channels through with the quality of inventions eligible for patents as well as the strength of the conferred rights can be adjusted. Finally, it is demonstrated how these legal policy instruments modelled in economics.

Chapter 3

A short history of the patent system

Our modern patent systems are the result of a long evolution. Here, we recount the history of patents very briefly mainly to demonstrate two important points. Firstly, the patent system has more or less continuously been changing over the centuries. Each of the more markedly adjustments reflect responses to advances in known technology fields, the creation of new technology fields or broader developments in society. Patent systems have mastered each of these challenges. They certainly are not working efficiently, but judged from a purely historical point of view, they perform reasonably well.

Secondly, prizes and awards have been used together with privileges to exclusively trade in certain goods and services or at special conditions. Privileges can be seen as the predecessors of patents as a reward system. The patent system prevailed whereas prizes and awards are rarely used today. The reason for the unequal development of the different policy instruments can be seen in their ability to cope with imperfect information and uncertainty. Clearly, both hypotheses are interrelated and are taken up in the next chapter.

The second part of this section is dedicated to the three most important international treaties which every country may join.[1] For each of them, the circumstances are described which lead to the negotiation table. Subsequently, the main achievements and most salient features of the treaties are explained. This second part of the chapter is important for the third part of this book that models international aspects of patent policy.

3.1 A general overview

The idea of patents is quite old. Early traces reach as far back as to the fourteenth century, and the patent system seems still to be far from optimal as the numerous demands for a patent reform in various countries prove.[2] It is far beyond the scope of this work to attempt to give a comprehensive overview over the history of patents. Instead, only certain aspects are presented that set the background against which the models in the subsequent chapters are developed.

The section starts with the early history of patents and mostly describes the developments in England. Subsequently, the history of the patent system in the

United States (US) is introduced because the remainder of this work mainly refers to the US system. Finally, the most important international treaties are discussed.

3.1.1 The first patents

Hippodamus of Miletus (498 BC – 408 BC) apparently proposed a reward system to foster innovations, including legal innovations. Aristotle (384 BC – 322 BC) criticized the idea in his *Politics* (Merges and Duffy 2002: 1). Despite the fact that the concept of a reward system was obviously known in Ancient Greece, the first patents can only be traced to the Renaissance. Given the eighteen hundred years that elapsed until the first patents were granted, it is natural to ask why a patent system did not develop earlier in the history of Western civilization.

The nature of a patent

The answer lies in the very nature of a patent. Firstly, a patent is a document that bestows upon the owner certain rights which guarantee above–normal profits.[3] Secondly, the patent system and therefore also patents itself are part of the institutional framework of our societies. As such, it shares the properties and characteristics of all institutions. In particular, an institution never adapts frictionlessly to changes in society. Instead, a certain degree of tension between the altered environment and the old institutions has to be build up before the latter finally adjusts. Due to the inherent inertia, the reasons for the appearance of the patent system must lie in certain developments in the Middle Ages that eventually turned a well adjusted institution into an obstacle: the guilds.

3.1.2 The guilds in the Middle Ages

In the Middle Ages, guilds became very strong in Europe. Guilds were an important institution and they performed a number of tasks. They settled disputes among members and often maintained a minimum quality standard. Guilds also protected the common trade interests of their members. Their regulations were quite rigid: setting up one's own business in a particular trade was only possible for members.[4] In addition, an invention could be used by any guild member without compensating the inventor (Bernhardt and Kraßer 1986: 41).

This custom highlights the medieval concept of a class society. Persons are first and foremost members of their class. Their rights and obligations were defined by the position of their class in social hierarchy. Guilds were an integral part of the medieval social system and its rules and regulations balanced rights and duties for each member.

At the dawn of the Renaissance, the rigid class structures loosened. More and more, persons perceived themselves as individuals and not as members of a particular class. Contrary to the times when guilds were first formed, the advantage for the owner of an invention seemed no longer to be the reasonable price for the protection guilds offered. This shift in perspective was also driven by advances in the legal system. Thus, at the end of the Middle Ages, guilds became an obstacle to innovation.

The antecedent of a patent system

The privilege system was at least originally introduced to counterbalance the increased power of the guilds that now stifled competition and entrepreneurship. Privileges conferred the right to practise the trade for which it was granted by the King or the Sovereign. In England, the first *letter of patent* was signed in 1311 (Bainbridge 2002: 311) (see also Table 3.1).

According to today's standards, only a minority of privileges were issued for true inventions. Often, they were granted for introducing a trade or technology that was well–known and practised abroad into the country (Bernhardt and Kraßer 1986: 41). In contrast to later years, privileges were not only bestowed on citizens, but also on foreigners provided they settled in the country that issued the privilege. Although a person could *apply* for a privilege, a formalized set of rules according to which they were granted did not exist, and a person had no enforceable right to it (Bernhardt and Kraßer 1986: 42).

The Venetian Senate's Act

The *Venetian Senate's 1474 Act* is the first recorded attempt to introduce an administered patent system. All essential features of a modern patent system were present: patent requirements were specified, a definition of the rights that were conferred was given, patent prosecution and patent enforcement as well as remedies to be applied in case of a violation of rights were put forth (Merges and Duffy 2002: 4).

It is subject to debate whether the Venetian Senate's Act is a historical curiosity without any practical importance even at the time or whether it influenced neighbour states and trade partners. Bernhardt and Kraßer (1986: 45) argue that the well–established privilege system remained the principal instrument to protect inventions and new trades in Venice. In contrast, Penrose (1951: 2) points out that nearly a hundred of these *special* privileges have been granted between 1475 and 1500. Merges and Duffy (2002: 4) even suggests that the idea of a systematic legal protection for inventions was spread by Italian craftsmen who travelled across Europe and were finally brought to England.

Developments in England

In the first part of the reign of Queen Elizabeth I, letters of patent were often used as a means to attract foreign artisans and, thus, to import technological knowledge (Merges and Duffy 2002: 5). By the end of her reign, the privilege system had been reduced to a full–fledged monopoly system. Although never strong, the link between a privilege and some aspect of usefulness or novelty of the subject matter was abandoned altogether. Instead, privileges were bestowed on courtiers or other persons that were in the sovereign's favour.[5] Partly, *letters of patent* were granted in order to raise revenues for the Crown that was perpetually short of money (Bainbridge 2002: 312).

These developments became detrimental for commerce which eventually led to the *1623 Statute of Monopolies* in England. This Statute explicitly outlawed all monopolies except the ones granted for true inventions. The Statute of Monopolies recognized a utility requirement, but a written description was not necessary to obtain a *letter of patent*. However, by the early eighteenth century,

the written description requirement was an acknowledge standard (Bainbridge 2002: 312), yet the system remained registry–based.

With the beginning of the Industrial Revolution, the notion as to the nature of an inventor's contribution to society that justified monopoly rights changed. Formerly, it was the new manufacture or new art itself, and it was at least in the beginning common practise that the inventor had to demonstrate the operability of his invention. Now, the focus shifted towards new and useful information so that the written description requirement became even more important and was more rigorously applied (Merges and Duffy 2002: 6).

3.1.3 The development in the United States

The British colonies in North America

England brought the concept of a patent system to its American territories.[6] The power to shape laws lay with the colonies that were formed. Consequently, the patent system was not unified across the colonies and conflicts regarding ownership and priority inevitably arose. There were several instances in which a patent was issued for the same invention to different persons in different states. Naturally, the patent owner in one state would sue the patent owner residing in another state when the latter tried to sell his products in the former one's state.

The 1790 Patent Statute

By the time the British colonies declared their independence, the problems of an ununified patent system had become obvious. It is, therefore, not surprising that patent law became federal law in the newly born American nation. According to Article I, Section 8 of the Constitution, the Congress is given the right and the obligation

> To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries.
>
> (Merges and Duffy 2002: 7)

The first patent statute was introduced in 1790, and the first patent was issued in the same year.

Patent rights were granted for new and useful inventions. The novelty requirement stipulated that the device or process in question must not have been used or known in the US at the time of registration (cf *infra*). Similar to the Venetian Senate's Act, a formal patent examination did not take place; inventors registered their inventions at the patent office. Whether or not the invention did fulfil a patent requirements at the time of registration was investigated by courts only if someone challenged a patent.

The nineteenth century

In the nineteenth century, two major changes were introduced. The registry–based system was exchanged for an examination system and the non–obviousness criterion was developed.

At the first glance, it is not obvious that starting to determine whether an invention satisfies the patent requirement before granting patent rights was the optimal response to the increase in patent registrations that was driving the Industrial Revolution. Indeed, receiving a patent in a registry–based system requires minimal resources. Patent rights can be enjoyed almost immediately even with a vastly increased number of patent registrations. However, expensive and prolonged lawsuits over infringement cases and allegedly invalid patents increased immensely. When registering, a patentee could never be certain that the patent would be upheld in court when challenged. Worse, the patent could not be contested for years, thereby prolonging the uncertainty. In a registry–based system, a marked raise in registrations will increase uncertainty, make patents in general less valuable and, therefore, undermine the effectiveness of the whole patent system. Thus, introducing a formal patent examination system in 1936 created more certainty for the rightful patentees and reduced the workload of the courts.

The increasing number of patent applications revealed another shortcoming of the early modern patent system. Inventions must be new and useful, but novelty in itself does not measure the degree of technological progress the invention does achieve. As such, the novelty requirement is insufficient to ensure that, amongst the vast number of applications, only *large* inventions were granted patents. Thus, the non–obviousness standard (cf *infra*) was developed requiring inventions to be non–trivial. From the mid–nineteenth century on, the non–obviousness standard was used by courts, but it was only codified in the 1952 revision.

Early twentieth century

In the early twentieth century, the mode of doing research in our modern form was born.[7] Increasingly, an invention was no longer the result of one individual's efforts. Instead, it was the output of a research department established by large firms to protect and secure the market dominance they had previously achieved.[8] They started to build in some cases considerable patent portfolios.

In the 1920s and 1930s, the public reacted to this ongoing development with mistrust and an anti–trust movement formed. It was felt that large companies built those patent portfolios to stifle competition. Patents were no longer seen as society's just reward for generating useful inventions and knowledge.

After World War II

After World War II, the 1952 Patent Act reinforced most of the fundamental principles laid out in the 1790 Patent Statute and its 1836 Revision. However, the anti–patent sentiments that originated in the 1920s and 1930s continued during the 1960s and 1970s.[9] They resulted in a low probability with which patents were upheld in court.

Before the *Court of Appeals for the Federal Circuit* (CAFC) was founded in 1982, patent cases were heard before one of the district courts. Their rulings differed widely, and often parties tried to bring a claim before a district court that had a history of ruling in the parties interest in similar cases.[10] At the same time, it was feared that the US industry was losing competitiveness against Japanese firms. Thus, partly to overcome diverging rulings of the district courts,

but mainly to strengthen patents and the patent system as whole, the CAFC was established. Right from the beginning, the CAFC was more inclined to affirm the validity of a patent (Merges et al. 2003: 111).

The 1990s

In 1995, the US government signed the treaty on *Trade Related Aspects of Intellectual Property* (TRIPs Agreement) and implemented it immediately. As a consequence, the US patent law had to be amended in some respects to bring it into accordance with the TRIPs regulations (Merges et al. 2003: 296). In particular, the patent term is now 20 years from the application date as opposed to 17 years from the issuing date. Patent rights were extended to include the exclusive rights of selling and importing goods associated with a product or process patent. In addition, US patent applications those subject matter is also covered by foreign applications are now published.

3.2 International patent treaties

3.2.1 The Paris Convention

The consequences of different patent systems

In the second half of the nineteenth century, patent laws had been implemented in many countries.[11] They differed widely in almost all respects, and some provisions had a particularly strong impact on international relations as the following two examples show.

Most countries required the patentee to *work* the patent within certain time, i.e. usually within less than a year. If the subject matter was a process, it must be used after this time. If the patent was granted for a product, production must have started in the issuing country within this time span. In case the patent owner did not comply with this provision, authorities had the power to revoke the patents. Thus, it was regularly impossible to retain patent protection *and* import products covered by a patent into the issuing country.

By the mid–nineteenth century, novelty had become a standard patent requirement. Consequently, an invention was only eligible for patent protection if it has previously not been known or used (cf *infra*). Novelty was destroyed by a number of circumstances. Among other instances, an invention was regarded to be not new if it had been described in a foreign patent application or had previously been presented as e.g. a prototype at an exhibition.

The inception

Before this background, it is hardly surprising that the International Exhibition in Vienna in 1873 was the occasion that brought about the first international patent treaty. Austria was one of those countries that regarded the novelty requirement to be forfeited if prototypes were shown at an exhibition. The American delegation was aware of the fact that all applications made after the Vienna Exhibition for a subject matter used for a prototype would be declined.

Consequently, they pressed Austria to adapt the patent law accordingly. Although Austria did not bow to the pressure of the US, they passed a special law expressly protecting the exhibited inventions.

The negotiation process

In fact, it took several conferences and the establishment of a permanent committee to reach an agreement. The first conference was held in Vienna directly after the International Exhibition. Although the unofficial meeting took place when anti–patent sentiments were still strong in Europe, the resolutions of the conference were decidedly pro–patent. In addition, a committee that organized the next unofficial conference was formed.

In 1878, the second unofficial conference took place in Paris. The ambitious goal was to install an International Patent Union which would replace national patent laws by a uniform international patent for all members. However, it became soon clear that opinions on patent policy differed to such a degree that the high aims had to be exchanged for a feasible approach of harmonizing certain aspects of patent laws.

Partly because the conference was unofficial and partly because consent could still not be reached on a number of issues, a permanent committee was formed. It took on the arduous task of drafting the legal text and to organize the first official meeting on which the draft was to be discussed. This conference took place in Paris in 1880. Here, the draft was officially adopted, yet some alterations were necessary. The Paris Convention was finally approved and signed in 1883.

The achievements

Among the most important achievements of the Paris Convention are the *national treatment* of foreigners, priority and the independence of patents.

According to the provision of *National Treatment*, foreign and domestic inventors have to be treated equally. The importance of this provision lies in the fact that it did not rest on reciprocity: any country that offered patent protection had to confer exclusive rights on e.g. Dutch inventors. Yet, their own citizens could not obtain Dutch patents since the Netherlands had abandoned their patent system at that time. Hence, the participants agreed to a provision which meant that outflows of profits due to patent protection for foreigners did not meet inflows of profits from that particular country. This was certainly a tremendous step at the time.

The priority rule provides for a certain period, usually twelve months from the filing date of first patent application, in which novelty of the invention cannot be forfeited by actions of third parties. In particular, this means that a patent application in one country cannot obstruct the novelty for an application in another country because they are treated as if they were all filed at the same, prior date. In addition, every other person is barred from applying for the same invention in a country within this specified period of time.

The provision on the independence of patents means that invalidation or revocation of a patent in one country does not entail the same effect for patents of the same invention in other countries.

Administration

As to 2009, the Paris Convention has 173 members. The most recent member is Thailand which joined in August 2008. The Paris Convention has been revised six times. Currently, the Stockholm revision of 1967 in the amended form of 1979 is in force. The Convention is administered by the World Intellectual Property Organization (WIPO) in Geneva.

3.2.2 The TRIPs Agreement

The problem

By the early 1980s, firms from industrialized countries had build up a sizable stock of foreign direct investments.[12] In this, pharmaceutical firms in general and the Pfizer Inc. in particular were not different from others. Chemical and pharmaceutical firms, however, faced a particular problem. Even among members of the Paris Convention or the Patent Cooperation Treaty (cf *infra*), patent laws differed widely. In some developing countries, therapeutic substances itself were not patentable. Patents were only granted for the technological processes that produced these substances. Therefore, e.g. Indian firms were able to find other methods of producing these substances and so to offer generics at lower prices in the domestic market legally. Moreover, Indian pharmaceutical firms were free to export those generics to other countries that did not consider active ingredients to be patentable subject matters.

The WIPO proved to be an unfortunate venue to push for stronger intellectual property rights (IPRs). Developing countries had joined forces to make their interests as chief importers of technologies heard at the WIPO. In addition, even if future negotiations would bring about stronger intellectual property rights to the existing treaties governed by the WIPO one problem would remain: none of those treaties had an effective enforcement mechanism like the one the GATT offered.

Putting an idea on an agenda

Allegedly, Pfizer Inc.'s CEO, Edmund Pratt, first came up with the idea to push the notion of intellectual property rights as being trade–related once again.[13] To achieve their ends, Pfizer's executives needed support within the business community. They used their networks to spread the idea at business forums and (non–governmental) trade associations. In addition, they lobbied the Advisory Committee on Trade Negotiations and thereby introduced the idea into the political sphere.

Even with the support of the business community and US–American politicians, making intellectual property rights part of the GATT required that they were put on the agenda of the next round of trade negotiations. The Ministerial Conference of Contracting Parties of the GATT sets the agenda, and the next meeting was scheduled for September 1986 in Punta del Este (Uruguay). Certainly, the US exerted a considerable influence, but the support of the European Community and Japan had to be ensured to have a realistic chance of linking intellectual property rights to trade aspects. So, Pfizer's and IBM's representatives created the Intellectual Property Committee (IPC) that consisted of 13 major US corporations in March 1986.[14] More than half of the members were

highly interested in removing the patent bar to chemical and pharmaceutical *substances* or new plants in developing countries. The protection of the production processes were regarded to be insufficient since inventing around patented processes was possible so that the therapeutic substances could legally be produce and, to a certain extend, even exported. In June and August, members of the IPC travelled to Europe and Japan to induce their competitors and business partners to make a case for the IPR–trade link with their governments.

Intellectual property rights were put on the Agenda of the Uruguay Round of the GATT negotiations and the TRIPs Agreement became part of the Final Act (Annex 1C) that was signed in April 1994 in Marrakesh.

The significance of the TRIPs

The TRIPs Agreement constitutes a minimum standard for the protection of intellectual property rights that every WTO member has to comply with.[15] Hence, it not only applies for patents, but also for copyrights, trademarks, geographical indications, industrial designs and trade secrets. Virtually all member countries had to adapt their national laws to some extend.

Keeping the history of the TRIPs Agreement in mind, the following provisions bear witness to the efficiency of the IPC's lobbying efforts: According to Article 27,

> [...] patents shall be available for any inventions, whether products or processes, in *all fields of technology*, provided they are new, involve an inventive step and are capable of industrial application. [...] patents shall be available and patent rights enjoyable without discrimination as to the place of invention, the field of technology and *whether products are imported or locally produced*.[16]

As the wording of Article 27 makes clear, an exclusion of certain subject matters from patentability is, with two exceptions, not longer permissible.[17] Countries, as e.g. India that did not protect therapeutical substances, must now grant patents for pharmaceutical products given the application does satisfy the other requirements (cf *infra*).[18] In fact, patents must be available for all (existing and future) fields of technology so that the patent system is at least theoretically highly adaptable.

In addition, patent protection must equally apply whether products based on an invention are imported or locally produced.[19] This effectively declares the provision according to which a patent must be worked in the issuing country unlawful.

A true novelty constitutes Article 28 paragraph 1 (b):

> A patent shall confer on its owner the following exclusive rights:
>
> [...] where the subject matter of a patent is a process, to prevent third parties not having the owner's consent from the act of using the process, and from the acts of: using, offering for sale, selling, or importing for these purposes at least the product obtained directly by that process.
>
> (WTO, 1994)

As a consequence, new products for which only one production process is yet known are now protected twofold. This is especially important for the pharmaceutical and the chemical industry: even if the patent on the substance itself (product patent) is later revoked for some reason it is still protected by the process patent as long as no other way exists to produce it.

Some scholars have attempted to evaluate the effects of the TRIPs Agreement. McCalman (2005) estimates e.g. the net cross–country profit flows induced by the Agreement for a selected number of countries in the short and long run. He finds, that although developing countries may gain in the long run, industrialized countries are certainly benefiting most from the TRIPs Agreement.[20]

3.2.3 The Patent Cooperation Treaty

Nature and aim

While both, the Paris Convention and the TRIPs Agreement are concerned with harmonizing patent law and other laws protecting intellectual property, the Patent Cooperation Treaty (PCT) has different aims. According to Article 1 of the PCT, a Union is established 'for cooperation in the filing, searching, and examination of applications for the protection of inventions' (WIPO, 1972: 11). The PCT provides a simplified application procedure for applicants who wish to receive a national and/or at least one international patent for the same invention.

The starting point

After World War II during which patentable research activities were halted, national and international patent applications increased again. Especially for those countries that heavily exported technologies, as e.g. the US, filing patent applications for the same invention in several countries was an arduous task, not only because of divergent material requirements. Every application in a foreign country had to be filed in the respective national language. If e.g. an US inventor were to file an application for the same invention in the US, Germany, Russia and Japan he would need three translations of the description of the subject matter apart from altering the respective applications to meet local patent requirements.

In addition, many countries required foreign applicants to be represented by a local patent attorney. From an applicant's point of view, multiple international applications for the same subject matter were burdensome and extremely costly. On the other hand, the patent offices had to perform the necessary searches for prior art and examinations for applications that had already been examined by other patent offices. From a patent office's point of view, the usually huge back–log could be cut back at least a little if they were allowed to rely on the search results of other patent offices.

Towards the PCT

It is therefore not surprising that the US initiated a dialogue to simplify international filing and examination procedures. A formal proposal was presented to the Bureaux Internationaux Reunis pour la Protection de la Propiete Intellectuelle (BIRPI) in 1966.[21] The BIRPI commissioned a study to identify and

assess possibilities to reduce duplicate filings and examination procedures. This was the basis for the first preliminary draft in 1967. During the following years, a number of consultations, at which not only state representatives, but also intergovernmental organizations and non–governmental organizations (NGOs) participated, took place. Expert opinions were heard and the first draft was amended several times. The final draft was signed at the Washington Diplomatic Conference in 1970. Formally, the PCT was enforced in January 1978 and originally included eighteen members.[22]

The principles of a PCT application

An international PCT application consists of an international and a national phase (WIPO, 1972: 746). Inventors of a signatory state usually file the PCT application at the local patent office. However, they may do so at any of the member states' patent offices. With the application, the international phase is entered. Although applicants file only one application, they apply for a patent in each of the designated countries. Thus, several national procedures are opened upon entering the international phase. If the application satisfies the formal requirements, all national procedures are halted for at least 20 months during which an authorized patent office undertakes the international search for relevant prior art and delivers the international search report to the applicant and to the patent office that received the application. The international application and the international search constitute the first stage of the international phase and are mandatory.

The second stage of the international phase is optional and consists of an international preliminary examination. Applicants may demand a preliminary examination after they have received the international search report. Upon entering the second stage of the international phase, national application procedures are suspended for at least another five months.

Either after the first or the second stage, the application enters into the national phase. Only then, the applicant is required to provide translations in case national patent offices ask for them. The international search report and, if applicable, the international preliminary examination report are sent to all the involved national offices.

Administration

As to 2009, the PCT numbers 141 contracting states; Peru and Chile joined in 2009. The PCT was amended three times, the last being in force since 2002. This treaty is also administered by the WIPO in Geneva.

3.A Important development steps of patents

In Table 3.1, important development steps of patent systems, historical events as well as certain inventions are given. The bold dates in the first column refer to developments of patent systems whereas the other dates stand for the year in which the patents for the inventions in the last column have been issued.

Table 3.1: History of the patent system, important inventions and historical events

Table chapter3-t-1

Notes

1 One of the most important international patent treaty is not described in this chapter — the European Patent Convention. This treaty is a regional one as its membership was and is restricted to European countries. Other such regional agreements exist.

2 See e.g. Merrill et al. (2004) or Jaffe and Lerner (2004) for the US.

3 Formerly, they consisted of practising a trade, buying or selling goods at a pre–specified prices. Today, patents are more complex structures; they confer on the owner the exclusive right '[...] to prevent third parties not having the owner's consent from the acts of: making, using, offering for sale, selling, or importing for these purposes' (TRIPs 1994: Article 28) the protected products or processes.

4 At least at the beginning, these regulations served as a minimum quality standard. In addition, guild–internal (dispute) settlement processes existed. Therefore, guilds and their regulations provided at least partly informal rules that are now part of modern legal systems. As such, they were necessary prerequisites for any modern form of business.

5 Privilege systems were common all over Europe which also experienced its perversion into a system to reward courtiers.

6 This section is mainly based on Merges and Duffy (2002: 7). See also Khan and Sokoloff (2001).

7 See e.g. MacGarvie and Furman (2005) for further information on the birth of industrial research laboratories.

8 See e.g. MacGarvie and Furman (2005) for an excellent description of factors that lead to formation of the private research laboratories.

9 The resentments against patents of the *big business* never quite stopped as the work of Reik (1946) and Vaughan (1948) who advocate the (limited) use of compulsory licensing shows.

10 See e.g. Hunt (1999b: 20), Lerner (2003: 2) or Menell and Scotchmer (2005: 40) for empirical evidence.

11 This section is based on Penrose (1951).

12 Most of the section bases on Drahos (2003). For detailed information see the references therein. For an alternative view, see e.g. Yu (2003).

13 According to Dutfield (2002), the Levi Strauss Corporation tried unsuccessfully to include intellectual property rights into the Tokyo Round of the GATT (1973–1979). Whereas Pfizer Inc.'s attempts were focused on patents, Levi Strauss' interests lay with trademarks.

14 The members were: Bristol–Myers (pharma), DuPont (chemicals and health care), FMC Corporation (chemicals), General Electric, General Motors, Hewlett–Packard, IBM, Johnson & Johnson (pharma), Merck (pharma), Monsato (GM), Pfizer (pharma), Rockwell International and Warner Communications.

15 Among others, Fink and Primo Braga (2005) investigated in the effects of stronger international patent rights empirically.

16 WTO, 1994, emphasis added by the author.

17 The permissible field restrictions are for: (1) diagnostic, therapeutic and surgical methods and (b) plants and animals other than microorganisms (paragraph 3). Field restrictions define certain technology fields for which patents are granted. Naturally occurring plants, minerals and substances, natural laws and phenomenon as well as mathematical algorithms are still not patentable, but those exceptions are not considered to be field restrictions. They constitute a bar across all technology fields. Paragraph 2 of Article 27 provides a *safety–hole* in that inventions may also be excluded from patentability when the public order or human, animal or plant life are threatened.

18 The US brought dispute settlement cases against Pakistan (1996), India (1996) and Argentina (1999, 2000) that dealt with the protection of pharmaceutical and agricultural and chemical produts (see Dispute Settlement Cases at the WIPO webpage).

See also Abbott (2005) on the Doha Declaration and similar declarations. They try to mitigate problems that may arise in developing countries that would be unable to pay for patented medicines.

19 As explained above, many countries revoked a patent if the owner did not work the patent in the issuing country after a certain period had elapsed. Only with the TRIPs Agreement this practise was abandoned in member countries of the WTO. In 2000, the US brought a dispute settlement case against Brazil because Brazil required the working of patents in the country.

20 Deng (2006) comes to a similar conclusion. Chadha (2009) studies the connection between the introduction of the TRIPs Agreement and the patenting activity in the Indian pharmaceutical industry.

21 The BIRPI was the international organization that administered the Paris Convention at this time and may be regarded as the predecessor of the WIPO.

22 Article 63 of the PCT provided that the Treaty enters into force three months after eight States, four of which must be large ones, ratified the PCT. This explains why it took eight years to establish the PCT. By the end of 2006, 133 countries have signed the PCT.

Chapter 4

Foundations of patent systems

The last chapter has revealed that the objectives as well as the instruments of patent systems changed over time. The *letter of patents* was introduced as a mercantilistic instrument, whereas today patents are frequently seen as an instrument to overcome market failure. Early patent systems required a demonstration that the invention indeed worked. Nowadays, a proof that the invention works is not necessary. Instead, a detailed description of the invention has become a patent requirement. To the most important adjustments in recent history of patents undoubtedly belong the TRIPs Agreement, as well as the possibility to obtain patents in new technology fields such as software, business methods, genetically modified plants and animals as well as gene sequences in a number of countries. Undoubtedly, new developments in science, technology and society will continue to challenge the patent systems and force the latter to adopt. Indeed, a number of scholars point out the urgency for patent reforms.[1] Yet, how are shortcomings and inefficiencies identified, and how are reasonable alterations to the system selected?

To evaluate the performance of a patent system, the objectives that are to be achieved have to be clearly defined. Various (philosophical) schools and a number of positive and normative theories have attributed different purposes to the patent system. Here, two schools, the natural rights theory and the utilitarian theory, are introduced. Since modern patent systems seem to draw more heavily on the latter, two patent theories belonging to the utilitarian school are described in greater detail.

4.1 Natural rights theory

4.1.1 John Locke and the natural rights theory

John Locke is considered to be the most prominent representative of the natural rights theory.[2] In the second of the *Two Treatises of Government* he develops his theory of political power, i.e. the beginning of societies, their aims, organization, justification.

The starting point of the analysis is the state of nature given by God. For

Locke, this state is not merely an abstract idea. It neither refers to an era before societies formed. According to Locke's libertarian view, every persons are born in this state and the laws of nature are their birthright. The laws of nature include the rights (1) to preserve their life, liberty and estates and (2) to prevent others from violating these rights and to punish others for a violation (Locke 1824: 179, § 87).

In the state of nature, every person has

> [...] a property in his own person: nobody has any right to but himself. The labour of his body, and the work of his hands we may say, are properly his. Whatsoever then he removes out of the state that nature that provided and left it in, he hath mixed his labour with, and joined to it something that is his own, and thereby makes it his property.
>
> (Locke 1824: 145, § 27)

Thus, property is an immediate consequence of the law of nature. Therefore, natural rights not only justify the existence of property, but also define the rights that should be attached to it: no one has the right to take another person's property (Locke 1824: 213, § 138). Locke's view on property is absolute. Hence, it is fair to assume that he meant the owner not only to have the exclusive right to possess the property, but also to enjoy the services of it and to transfer the property to others — the modern definition of property rights.[3] Natural rights and property are firmly interlinked and are central to Locke's political theory.

The next step is to explain, why societies formed. Although persons have absolute freedom in the state of nature, they must also be prepared to defend their freedom frequently. Locke argues that societies are created with the main purpose to preserve the members' natural rights in general and their property in particular. To this end, society must be endowed with certain rights and powers: legislative and executive powers. However, both are limited to the extend that

> [...] the law of nature stands as an eternal rule to all men, legislators as well as others.
>
> (Locke 1824: 210, § 135)

4.1.2 Intellectual property and natural rights

The distinction between tangible and intellectual property is meaningless as far as natural rights theory is considered. As long as labour is mixed with unowned resources, property is created and must be protected.[4] Thus, a patent system protecting intellectual property is not only justified according to natural rights theory. It also lays down the principles for its provisions.

For a natural rights theorist, intellectual property should have the same level of protection as tangible property has. This obviously includes the exclusive rights of possessing it, enjoying its services by e.g. licensing it, and transferring it to others. Indeed, these are acknowledged standards for all types of property today. There are also other normative aspects of the equal treatment or all types of property that are unfamiliar since they have not been implemented. In accordance with tangible property, intellectual property rights should not be limited in duration. Most importantly, patents would not be protected against

independent discovery (cf Oddi 1996), since this would interfere with other persons property.

Unquestionably, the natural rights theory did play a role in the history of modern patent systems. In modern patent law, however, no traces are left.

4.2 Utilitarian theory

4.2.1 Jeremy Bentham and Utilitarianism

According to Eatwell et al. (1998: 770), *utilitarianism* has three features:

> (1) individual well–being ought to be the end of moral actions; (2) each individual is to *count for one and no more than one*; (3) the object of social action should be to maximize general utility.

Jeremy Bentham laid the foundations to utilitarianism as we know it today. Although utilitarianism and natural rights theory share some ideas, as e.g. that all human beings are equal and that they pursue good and avoid evil, Bentham rigorously rejected natural rights theory. For him, rights that predate or precede positive laws do not exist — *every* right is created by law and, thus, by a government. Perhaps even more importantly, natural rights theory lacks any guidance as to which measure from the vast number of possible actions a government should take in order to increase social welfare. Bentham's concept provides both, the ends and the means. According to the utilitarian theory, the ultimate object is to maximize general utility, and general utility also helps to decide which actions to choose.

4.2.2 Utilitarianism and patent systems

According to Bentham's utilitarian theory, policymakers should maximize general welfare. This principle also establishes a yardstick that aids in answering the following questions: (a) should the government intervene to strengthen intellectual property, (b) which set of rules should be installed, and (c) how should it be designed?

Too many or too few inventions?

At least in principle, the first question has a simple answer. A policymaker should impose a set of rules if the actual and the welfare–maximizing level of inventions do not coincide. Although it is impossible to calculate the optimal level of inventions, we can identify incentives that cause over- or under–investment into R&D.

First of all, most inventions are regarded to be socially valuable since they satisfy a human desire more effectively (product innovations) or perform a certain task more efficiently (process innovation). In addition, high R&D and innovation rates are necessary to reach or sustain appropriate economic growth paths according to one branch of modern growth theory.[5] Therefore, we might conclude that governments want to encourage R&D.

Indeed, early utilitarians highlighted either the information good characteristic (Jeremy Bentham and John Stuart Mill) or the public good characteristic

(Arthur Cecil Pigou) of inventions. The former refers to the notion that imitation costs are always lower than development costs. Consequently, imitators would successfully drive inventors out of the market if inventions were not protected in some form. Rational inventors anticipating the inventors actions may decide to divert their resources away from inventions and thereby reduce society's welfare.

The public good argument rests on the fact that inventions essentially consist of information which is non–rival and non–excludable without intellectual property rights. Since inventors are unable to appropriate the full social benefit, they will provide an inefficient level of inventions and innovations. In contrast to most other public goods, inventions affect social welfare in two different ways. Firstly, inventions increase social welfare directly by providing new products and processes. The second channel emphasizes knowledge as an important input factor for further research which itself leads to new products and processes. Thus, welfare is increased indirectly. In the latter function, knowledge is socially desirable in itself and point to a particular problem arising from the cumulative nature of inventions (cf *infra*).

Both, the information good and the public good argument, arrive at the conclusion that there will be too few innovations as compared to the social optimum unless the government intervenes.

By the time the mechanisms of imperfect markets had been understood better, possible negative consequences of the patent system were pointed out. Plant (1953) e.g. claims that all inventions are spontaneous. Thus, they are made with or without intellectual property rights. Other protective mechanisms are available and have been used for a long time: e.g. trade secrecy, first–mover advantage or lead time. Granting exclusive rights, so Plant, would lead to over–investment, especially when the research environment is competitive. Although Plant's argument is supported by empirical evidence, a patent system may still be justified since trade secrecy prevents inventions from performing their second material function: disclosing the embodied knowledge so that subsequent inventions can be made.

From utilitarian point of view, government intervention is justified whenever the level of inventions is sub–optimal. Unfortunately, there are incentives for both over- as well as for under–provision of inventions, and they typically are present at the same time. However, it seems undebated that without a regulatory system more inventions would be protected by trade secrecy which deprives society from valuable knowledge.

Patents, prizes or awards

Even if market failure can be attested, a patent system might not be superior to other alternatives. Thus, the second question is concerned with the relative performance of patents, prizes, awards and procurements.[6] Although the patent system has frequently been criticized, only few scholars tried to compare the different systems.[7]

Suzanne Scotchmer shows that prizes provide the same incentive to innovate as patents, but at lower costs under certain circumstances.[8] Prizes and awards do not entail deadweight loss which inevitably arises in a patent system. Given that prizes and awards are even older than patents and have been used to reward inventors to the present day (Scotchmer 2004: 41–46), why does the

patent system dominate other alternatives? In order to provide the socially efficient incentives to innovate using prizes or awards, the value of the desired invention must be known. It may be more easily established in some cases. For the vast majority of cases however the value of an invention remains highly uncertain even after the patent has been issued. This is particularly true in the pharmaceutical industry. Typically, several years elapse until a patented therapeutical substance has been developed into a medicine suitable for clinical tests.[9] These tests are necessary to gain market approval, and the value of the active ingredient is considerably higher if market approval is granted than if it is denied.

The advantage of patents over prizes and awards consists in a patent system's ability to perform reasonably well with comparably little information.[10] This makes the patent system more flexible than either of the alternatives.

The optimal design

The last question is concerned with the optimal design of a patent system. In pursuing this task, one cannot exclusively focus on either social benefits or social costs of inventions since policy instruments increasing the former may also increase the latter. Thus, the question of how to achieve the ultimate object can be re–addressed as follows (see Besen and Raskind 1991):

1. Do individuals and enterprises have sufficient incentives to invent?
2. Is the balance between the incentives to invent and the dissemination of knowledge optimal?
3. Are inventions generated at minimum *social* costs?

Each of the numerous (normative) patent theories have to answer these questions. However, only two are introduced here: the incentive theory and the disclosure theory.[11] They have been chosen because they each corresponds to one of the functions of a patent system: to facilitate technical progress and to disseminate knowledge.

4.2.3 Two patent theories

All utilitarian patent theories have to demonstrate market failure in order to justify the intervention into the market mechanism. Identifying the reasons for market failure is also the starting point for laying down the guiding principles of an appropriate patent system. By nature, they will be incomplete because each theory emphasizes a single factor. In reality, however, all factors of market failure are present at the same time.

Incentive Theory

This patent theory focuses on the firms' incentives to innovate. For firms, to engage in R&D is just another, highly uncertain investment decision. As every other investment decision, the aim is to keep at least pace with competitors and at best to out–perform the rivals in future. Incentive theory emphasizes that creating a piece of information or knowledge is more expensive than to copy

it. Consequently, the profitability of research projects crucially depends on the firms' ability to protect their inventions and know–how.

Using the standard neoclassical framework, it can be shown that the incentives to innovate are optimal only if firms are able to appropriate the entire social benefits from the invention (Scotchmer 1991). In particular, it is assumed that there is one firm that has the potential to innovate in a perfectly competitive environment where imitation is costless. Neither legal protection nor alternative protective mechanisms exist. In a frictionless world, firms are unable to recoup their research expenditures. Rational actors will anticipate this situation and refrain from R&D. However, if the firm were to receive the entire social benefit, e.g. due to an effective and efficient patent system, it would exert the optimal effort into the R&D project.

Alternative protective mechanisms such as trade secrecy and lead time have been used for a long time to protect inventions.[12] Thus, some research projects will be profitable even without legal protection. Yet, since expected payoffs are lower for both alternative mechanisms, the incentives to invent will not be socially efficient. Strong, legal patent protection seems therefore justified. If there was only one invention in a particular line of technology, an infinite patent term would induce sufficient incentives. However, there is usually a sequence of inventions in the same technology line where later inventions perfect previous ones or find new applications. It is therefore impossible to give every inventor the full social benefit of is invention (Scotchmer 1996). Thus, with cumulative inventions in mind, patent law has reduce the incentives for every inventor to achieve the maximal possible incentives for all inventors in the long run.

Incentive theory has always been criticized for both theoretical and empirical reasons. One may doubt whether the existence of patent protection does increase the number of inventions and, therefore, welfare. According to (Cohen et al. 2000), firms usually rely on a mixture of legal and alternative protective measures. Thus, the same inventions may have been created without patent protection and therefore at lower social costs. Empirical evidence as well as surveys yield mixed results (cf e.g. Penrose 1951; Kortum and Lerner 1998).

Even if patents foster industrial progress as intended, they may also create social costs by creating too much incentive or the wrong kind of incentives. If the market structure is oligopolistic and several firms try to find a solution for a known problem, research costs are duplicated, but the social benefit of the invention remains the same. The incentives created by the patent system may generate competition in research that eventually dissipates the entire social benefit of an invention (cf e.g. Grady and Alexander 1992; Oddi 1996). In addition, firms use their patents not solely for the purpose of recouping their up–front investments, but also to pursue (aggressive or defensive) strategies to weaken their competitors (cf Shapiro 2001). In particular, an attractive patent portfolio may become a necessary prerequisite to negotiate license agreements in different technology fields. Although this strategy may actually reduce development and production costs for incumbents when existing patent portfolios lead to favourable cross–licensing agreements, it might constitute a prohibitive entry barrier for other firms. The importance of strategic patenting greatly differs across industries. However, it is well documented for the subgroups of consumer electronics industry and the software industry (Hall and Ziedonis 2001; Brodley 1990).

Whether or not the patent system does create *the right amount* of incentives

to generate inventions largely remains an empirical question. Presently, the evidence is at best mixed.[13]

Disclosure Theory

The second patent theory introduced here looks upon patents as a contract between two parties, society and the inventor.[14] Each party has its rights and obligations. The *traded object* is the knowledge embodied in the invention. The *price* society pays for it is granting the exclusive patent rights. Different form the incentive theory that centres on the welfare effects of new products or processes, the disclosure theory focuses on making information publicly available.

Information has a very peculiar production function.[15] It almost entirely consists of fixed costs. The variable costs for reproducing the same information are negligible. Under these circumstances, the efficient allocation of information is to provide it free of charge to everyone who is interested.[16] However, the creators of information have to cover the fixed costs so that they will usually charge a positive price. An inefficient allocation of information and, therefore, market failure will arise. In case the creators anticipate that potential consumers are not willing to pay for the information, they will not produce the information in the first place unless their private utility for the information exceeds the fixed costs.

Inventions are basically information that are extremely expensive to produce partly because the research process is highly uncertain. Different from creators of many other kinds of information, inventors may sill recoup their R&D costs if they can protect their inventions appropriately, i.e. usually by trade secrecy. Thus, as along as a patent system succeeds in revealing the information that would otherwise have been kept secret it reduces the level of market failure Beckerman-Rodau (2002). Since patents are published 18 months after the issuance, patented inventions are invariably revealed.[17]

Publishing inventions may increase welfare further. Especially when R&D processes are envisaged, knowledge is undoubtedly an important input factor. Once a patent has been granted, other researchers may build upon the knowledge and generate new inventions. This makes it possible that e.g. applications for a basis invention can be discovered and brought to the consumer sooner as compared to a situation where the invention was kept secret.[18] Modern growth theory is build upon this concept and links information to economic growth and welfare.[19]

Admittedly, a patent system would not be the first choice of correcting market failure under these circumstances. Prizes that reward inventors and put the invention into a public domain would fare much better as deadweight loss is avoided. As discussed above, however, a prize system is difficult to establish in an uncertain environment, i.e. when the value of the information cannot be assessed. Again, a patent system may perform reasonably well with inventions.

Disclosure theory is rather flexible as far as specific patent regulations are concerned. Different from natural rights theory which would imply an infinite patent term when applied strictly, disclosure theory would specify the lowest possible patent life that induces inventors to use patents rather than trade secrecy to protect their intellectual property. This also describes the general implications of this patent theory for patent systems: to use the necessary and least invasive policy instruments to achieve the disclosure of information embodied

in inventions.

Notes

1 See e.g. Merrill et al. (2004), Jaffe and Lerner (2004) or Bessen and Meurer (2008) for suggestions for the US.

2 See also Bix (2000) for a modern interpretation.

3 See e.g. Demsetz (1967) for an economic view on property rights.

4 See also Claeys (2006).

5 For an overview on modern growth theory see e.g. Romer (1990), Grossman and Helpman (1991) Barro and Sala-I-Martin (1995) or Aghion and Howitt (1998, 1992). For an empirical study see e.g. Falvey et al. (2004).

6 See e.g. Scotchmer (2004) and Gallini and Scotchmer (2002) among others. Wright (1983) explicitly compares the performance of patent, prizes and research contracts in an environment with and without information asymmetry. Scotchmer (1996) reveals that patent protection is not necessary to encourage second–generation products under certain circumstances.

7 A more fundamental critique is put forth by Crampes and Langinier (2005) who show that the pure existence of an intellectual property rights system may be detrimental to the flow of useful innovations.

8 See Scotchmer (2004) and the references therein.

9 See Abrantes–Metz et al. (2006) for a study on the development phases and the subsequent success of pharmaceutical products.

10 In a recent paper, Acemoglu et al. (2008) show that subsidies which are in some sense similar to prizes cannot achieve the social efficient outcome. Well designed patent systems, however, can.

11 See e.g. Oddi (1996) and the references therein for other utilitarian theories.

12 See e.g. Cohen et al. (2000); Cohen (2004), Bulut and Moschini (2006) or Hussinger (2004). Atallah (2004) models the firms' decision of using either patents or trade secrecy as a protective mechanism.

13 See e.g. Graham and Mowery (2001), Lerner (2002) or Hall (2004).

14 See Denicolò and Franzoni (2006), who studies the differences between both patent theories.

15 For an introduction to information economics see e.g. Hirshleifer and Riley (1992).

16 Creane (2006) argues that dissemination of knowledge can be excessive and thus decrease welfare.

17 The US form an exception. Here, only inventions which are (to be the) subject of foreign patent applications are published.

18 See e.g. Bhattacharya and Guriev (2004, 2006).

19 See e.g. Grossman and Helpman (1991) or Romer (1990) on modern growth theory. See e.g. Horii and Tatsuro (2005), Futagami and Iwaisako (2007), Gould and Gruben (1996) or Grupp and Stadler (2000) on works that directly connect intellectual property to economic growth.

Chapter 5

An introduction to patent law and policy instruments

All components of a patent system potentially affect firms' decisions on research projects, the question which instruments are to be used to protect the intellectual property as well as the market structure itself. It is a quite complex system so that it seems expedient to describe it in greater detail to understand its workings and mechanisms. Although the TRIPs Agreement (WTO, 1994) has harmonized patent law and therefore also the patent system to a certain degree, there remain considerable differences in patent systems. These distinctions are inconsequential for the analysis in the following chapters so that we decided to refer to only one system, the US patent system.[1] The first section traces an invention's way through the patent system, i.e. from filing the application to the court rulings concerning remedies for the infringement of a patent.

Some aspects of the patent system, as e.g. the way in which damages are determined or the judges' tendency to uphold challenged patents, can have a huge impact on firms' inclination to use patents. However, the book focuses on certain provisions of patent law that determine the strength of the patent rights conferred. Consequently, the second section gives a short introduction to patent law. It covers familiar patent requirements that inventions must be new and useful as well as more unfamiliar ones of disclosure and enablement. For each of the included provisions, the principles are outlined according to which patent examiners or courts determine whether the requirements are met in a particular case. Subsequently, potential economic effects are sketched.

Patent law has to concern itself with numerous details to ensure that all cases are treated equally, and it must remain flexible enough to accommodate future developments without stifling the patent system. In addition, patent law has to take care that present inventors are not rewarded at the expense of future ones. Economic models on the other hand cannot and should not deal with all the legal details. They generalize effects and demonstrate the how single provisions work. Unsurprisingly, the details of patent law do not easily translate into economic models. Therefore, the third section shows which legal provisions are to be analyzed in the remainder of the book, and how they can be treated in a way suitable for economic modelling.

5.1 The patent system

The patent system comprises all steps and provisions that govern patents from filing an application to the issuance of the patent as well as all conflicts over patent rights. Frequently, the patent system is divided into prosecution, i.e. the whole application process, and enforcement. The former includes all administrative steps leading to the issuance or a rejection of a patent. The latter negotiates conflicting interests between a patentee and other parties which possibly arise from the administrative act of granting patent rights.

5.1.1 Patent prosecution

Patent prosecution starts with the application. Except for applications under the Patent Cooperation Treaty (PCT) or the European Patent Convention (EPC) they are filed at the domestic patent office. Applications under the PCT or the EPC can be filed at any patent office of a contracting state.

The application describes the invention and contains the principal information on which examiners base their decisions. As the patent itself, the application is a highly standardized document. In fact, an application and the corresponding patent may differ in the description and the claims of the invention but not in the structure. Therefore, it seems appropriate to discuss the parts of a patent in some detail.

The parts of a patent

A patent has three main parts: the bibliographic data section, the body and the claims.

The bibliographic data

The bibliographic data section (cf Figure 5.1) provides general and technology–related information.[2] In the headline, the issuing country and the first inventor are on the left–hand side. On the right–hand side of the headline, the patent number and the issuing date of the patent are given. The general information in

Figure chapter5-f-1

Figure 5.1: Bibliographic data of a patent

the main part consist of a short descriptive title, the inventors and the assignee as well as the application number and the filing date.

Since the full description of the invention is given in the body of the patent, the technology–related information in the bibliographic data section is very brief. It reveals the technology fields to which the invention belongs by listing the appropriate numbers of the domestic and international classification code. Further, the technological fields that the examiners searched for relevant prior art are detailed.[3] Inventors are required to specify all relevant sources that were used in making the invention in the application. These references are part of the bibliographic section and frequently consist of other patents and published scientific papers. Finally, a concise abstract that must convey the essence of the invention as well as its use is given.

The body

The body contains the main information concerning the invention. Usually, drawings of the main parts of the invention or flow charts in case of software–related inventions are presented first. It follows the specification of the invention (cf Figure 5.2) which gives exhaustive information on three points: written description, enablement and the best mode (cf *infra*). Together, the elements

Figure chapter5-f-2

Figure 5.2: The body of a patent

constitute the disclosure and enablement requirement which must be met at the filing date of the patent and cannot be amended to include new subject matter during the examination process.

Because all inventions draw upon previously created knowledge, inventors have to describe precisely how their invention is related to the prior art in the written description section. They also have also to make clear what sets their invention apart from the prior art and the cited references, i.e. they have to demonstrate an inventive step.

For a patent to be granted, the invention must be useful which also means that a specific use for the invention has to be given in the application. Both, the enablement and the best mode requirement are concerned with making and using the invention. Specifically, inventions must be outlined in such a way that a person skilled in the technology line is able to make and use the inventions without undue experimentation.

Naturally, all inventors want their patents to be as broad as possible. One way of achieving it is by listing e.g. a range of materials that can be used instead of naming a single one. The best mode requires inventors to disclose the configuration that has proven to be most successful in the research process.

The claims

The last part of a patent consists of one or more claims (cf Figure 5.3, emphasis added). They state in a formalized manner what the patentee claims to have

Figure chapter5-f-3

Figure 5.3: The claims of a patent

invented. The claims must not contain any information that is not described in the body of the patent. Frequently, they are considered to be the most important part of the patent since they define the scope of the patent rights that are conferred. Naturally, they are rewritten several times in negotiations between the examiner and the inventor or the inventor's attorney during the examination process.

Other steps of the prosecution process

After the application has been filed, examiners have to assess whether the application satisfies all patent requirements (cf *infra*). Therefore, they have to

search for prior art above and beyond what the inventors states in the application. Frequently, internal and international patent databases as well as scientific literature are consulted. On the basis of the search result, the examiners determine whether the invention constitutes sufficient technological progress to justify a patent.

In case the examiner rejects the patent application, the claims can be rewritten two times. Usually, this stage is entered when claims were initially poorly defined and too broad. Negotiations between examiner and inventors help to define the proper patent scope or to clarify that the invention does not merit a patent. At the end of this process, examiners either accept the application in which case the patent is granted or they reject it. In the latter case, the applicants may abandon their efforts to receive a patent or file a continuation patent.

The form of continuation patents is unique to US patent law. Here, the application re–enters the examination process and can be changed or split into several patents to meet the patent requirements. From the patentee's point of view, this procedure has two advantages: the continuation patent retains the 'parent' patent's priority date and continuation patents are usually not published.[4]

Those applications for which patent protection is not sought abroad need not to be published in the US. All other patent applications are to be published 18 months after the filing date.

5.1.2 Patent enforcement

Patent enforcement is concerned with matters after a patent has been issued. It settles conflicting interests between patentees and (potential) infringers. In general, two situations may arise. (1) A patentee puts forth an infringement action against a party which, so the patent owner believes, violates his patent rights. (2) A party, as e.g. a potential infringer or another patent owner, brings a declaratory judgement action against the patentee because the former believes that some or all claims are invalid.

In both cases, the validity of the patent can and usually is challenged. However, the (potential) infringer has always the burden of proof. Since a patent has undergone a thorough administrative examination, it is presumed that patents are valid until there is sufficient evidence for doubt.

If the infringement action is successful, the accused party is held to be liable for infringement. Three kinds of remedies may be imposed (cf *infra*): (1) an injunction, where the infringer has to give up producing the product or using the infringing technology, (2) a damage has to be paid, and (3) the legal expenses of the winning party have to be paid. Often, a combination of those measures are imposed.

Since the (potential) infringer usually defends himself by claiming that patent is invalid in the first place, part of the claims can be declared invalid even if the patentee wins his case.

5.2 Legal policy instruments and their economic effects

The main purpose of this section is to introduce the *legal patent instruments* and to describe their economic effects.[5] Only some of them, i.e. patent eligibility and patent requirements, form the principal subject of the following chapters. Therefore, they are only briefly discussed in this section. The legal patent instruments that are not studied in the following chapters are examined in greater detail here.

Before introducing the main components of patent law, a description of the patent rights are given. According to §154 of USC 35 a patent constitutes

> a grant to the patentee, his heirs or assigns, of the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States, and, if the invention is a process, of the right to exclude others from using, offering for sale or selling throughout the United States, or importing into the United States, products made by that process, referring to the specification for the particulars thereof.

5.2.1 Patent prosecution

Patent fees

Application fees

When inventors file a patent application they have to pay application fees. Throughout the whole examination process other fees have to be paid which depend on certain characteristics. Typical fees are e.g. search fees, examination fees and fees payable if the patent application has more than a certain number of claims (cf Table 5.2 in the Appendix).[6]

From an economic perspective, a rational inventor will only apply for a patent if the expected value exceeds the expected cost of a patent. Principal components of the expected patent costs are the total application costs as well as litigation costs that may arise in future over validity or infringement issues. To which extent litigation costs will become relevant depends largely on the characteristics of the invention and the patentee (cf e.g. Harhoff and Reitzig 2004). Application costs comprise application fees as well as the costs for a patent attorney and are sunk.

Application fees are not primarily imposed to increase the patent office's budget (Jaffe and Lerner 2004: chapter 5), but to stop inventions that do not contribute to the progress of the technology field from entering the patent system. Trivial inventions would not satisfy the patent requirements and would therefore not be granted a patent. Nevertheless, they would bind resources. Although application fees are by far the smallest part of the total expected patent costs, they have the effect of an entry barrier. This effect is amplified by the fact that as soon a patent application is pursued attorney costs will also accrue.

The presumption that only trivial inventions are kept out of the patent system is subject to dispute. On the one hand, it has been argued that application

fees are still too high for small or young enterprises as well as individual inventors although they are eligible for reduced fees in the US (see Table 5.2 in the Appendix). On the other hand, concerns about the raising number of poor quality patents (Merges 1999) are voiced.

Maintenance fees

Patent fees do not only apply during the examination procedure. Although the patent term is 20 years from the filing date, maintenance fees have be paid to keep the patent in force (Table 5.2). They are progressive, and only about one of three patents patents cover the entire patent term.

Maintenance fees also work as a sorting mechanism protecting valuable inventions longer than less valuable ones. Rather than imposing a large fee when the patent is issued, smaller instalments are well adapted to the fact that information affecting the value of the patent is gradually revealed over time. Thus, e.g. a pharmaceutical firm may keep the patent in force while clinical tests for a new medicine are carried on. After a rejection of market approval, however, the firm may want to abandon the patent.

Patentable subject matter

After the application reaches the examiners, they have to determine whether the invention is eligible for patent protection, i.e. whether the discovery belongs to a patentable subject matter. Patentable subject matters are broad technology classes for which patents can in principle be obtained. §101 of USC 35 rules that:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefore, subject to the conditions and requirements of this title.

Different from European patent law, US law does not explicitly exclude any fields or classes of technology.[7] However, the wording of the provision leaves no doubt that patent were only intended for technical inventions in a broad sense, i.e. not only for technological processes and machinery, but also materials and biological and chemical substances. This excludes not only artistic expressions that are protected by other forms of intellectual property rights. It has also been the standard of patent law from its beginnings that laws of nature, physical phenomena, abstract ideas, and naturally occurring plants and minerals are generally not patentable. These exceptions are known as the *natural law doctrine*.[8]

History has shown that the §101 USC 35 was initially narrowly construed so that e.g. computer software, business methods and living organisms were initially unpatentable. Over time, however, each of these subject matters became patentable in the US.[9]

Since the expansion of patentable subject matters in an international framework is studied in Chapter 12, general considerations are here limited to the national perspective, i.e. under autarky. For illustrative sake, take the example of biotechnology and assume that the patent system indeed spurs innovations. Making biotechnology a patentable subject matter has two effects. Firstly, it is

important to note that the absence of patent protection does usually not imply that there are no inventions in that field. University research is largely state–funded so that basic research is conducted in every technology class. In addition, private firms that made a discovery in a new technology field may simply resort to alternative protective mechanisms such as e.g. trade secrecy or lead time.[10] In a situation where technological knowledge is growing fast and trade secrecy is used, crucial information remains scattered since trade secrecy prevents its dissemination (cf *supra*).[11] As a consequence, improvements on e.g. a method to alter a particular gene sequence appear on average later, and consumers have to wait longer for practical applications such as new drugs. Biotechnological progress can be expected to increase if the government decides to grant patents in this field. Indeed, Merges (1995: 107) understands intellectual property rights as device for industrial policy that is capable of subsidizing creative firms in this particular field without burdening the budget.

Secondly, private enterprises are less likely to direct their research efforts to new technology fields for which they cannot obtain patents. Again, a higher technological progress can be expected when a subject matter becomes patentable, not because of more readily available knowledge, but because research expenditures and investments increase in the particular technological class. Regibeau and Rockett (2005) argue that granting patents on *pure gene sequences* would redirect research toward basic research and away from the search of useful applications.

Utility requirement

The utility requirement is derived from §101 USC 35 (cf *supra*), where it is explicitly stated that inventions must be useful to be patentable. In this provision, the utilitarian roots of the patent system become obvious: only inventions that increase social welfare merit the intellectual property rights conferred by patents (cf *supra*).

Formerly, this patent requirement has been used to exclude immoral and dangerous inventions from patentability. Today, the USPTO largely focuses on technical utility and leaves other government agencies to deal with moral issues and the protection of consumer interests.

The modern view on the issue is that an invention has to pass three tests in order to comply with the utility requirement: (1) practical utility, (2) beneficial utility, (3) operability.

The *Operability* test examines whether the invention is capable of accomplishing the utility claimed in the application. Since it is not longer necessary to demonstrate the operability of the invention, the test mainly serves to exclude fantastic claims, as e.g. to have invented the perpetuum mobile, or claims that base on accidental mistakes made by the inventor.

The *beneficial utility test* assesses whether the invention is in general desirable for society. By this standard, discoveries directed at immoral or socially harmful purposes are unpatentable. Indeed, patent applications on gambling devices have been rejected on moral grounds prior to 1977. However, the USPTO has become more lenient as far as immorality is concerned.

Finally, the invention must have a practical purpose, i.e. *practical utility*. The standard has never been very high. Nevertheless, it frequently becomes relevant for pharmaceutical, chemical or biotechnological inventions.[12] To obtain a

patent, it is not sufficient to change e.g. substances or materials that have proved to be socially useful and hope that the new ones will finally turn out to have some purpose — the substance in itself is not useful for society. The inventor must credibly demonstrate how it can be used to the general advantage.[13]

The utility requirement is seldom the stumbling block for an applicant. Economically, the requirement guarantees a minimum level of benefit society receives in exchange for the exclusive rights it confers to the inventor.

It should be pointed out that the USPTO does not consider social usefulness and commercial value to be synonymous during the examination phase. It recognizes that a new substance e.g. to treat malaria is socially desirable even if it does not survive the clinical test phase and, thus, does not have an immediate commercial value.[14] This substance might serve as a starting point for the development of new substances that eventually get market approval and successfully cures malaria. Hence, the original substance serves a socially useful purpose although it has no commercial success and, thus, deserves patent protection.

In addition, the utility requirement discourages researchers from filing patents too early and encourages them to find new products and processes *as well as* keeping a practical application in mind.

Novelty and statutory bars

Novelty

The novelty requirement is not only derived from §101 USC 35 which states that inventions must be new. §102 of USC 35 stipulates that

> A person shall be entitled to a patent unless—
>
> (a) the invention was know or used by others in this country, or patented or described in a printed publication in this or a foreign country before the invention thereof by the applicant for patent [...]

The short passage points at two connected problems: (1) how to measure novelty, and (2) to which point in time the measure should be applied.

With respect to the second question, two concepts are used. Either the invention has to satisfy the novelty requirement at the *invention date* or at the *filing date* of the application. Except for the US, all countries implemented the second concept. Clearly, both concepts differ as a patent application might not be filed immediately for various reasons. Inventors may want to perfect their invention or broaden the potential patent scope by establishing new uses for the invention. Alternatively, the invention might be kept secret for strategic reasons.

A closely related issue is that of who shall be considered to be the inventor of a subject matter. The two above mentioned concepts are used which are also known as the first–to–invent and the first–to–file method in this context. Again, the person who invented the subject matter and the one who applied for the patent first need not to be identical. In holding on to the first–to–invent method, the US aim at rewarding the true inventor. Disputes on who did invent a certain subject matter do occasionally arise, and court trials over and again show that the first–to–invent method makes measuring novelty much more complex.

Yet how is novelty actually measured? §102(a) provides that an invention is regarded to be new unless it has been known or used or described in a printed publication before the invention date (cf *supra*). Ultimately, all inventions draw upon knowledge created by others so that parts of them are well–known and used. Then, the question is really about how much of an invention must be described in a reference of the prior art to destroy novelty. The US apply the concept of anticipation. The principle is simple and straightforward. For an invention to be anticipated by prior art, a single published reference must describe each and every element of the claimed invention. In this case, the invention fails the novelty test and is unpatentable. If some elements of the invention are described in one reference and the rest in another publication, the application for the subject matter passes the novelty test. By combining two or more references, the inventor has created something new. Whether such combinations constitute a sufficient technological advance to merit a patent then becomes a matter of obviousness.

From an economic point of view, it is important that patents are only issued for inventions that generate social benefit. Disseminating technical knowledge is one way of increasing welfare. Although not necessarily rewarding the true inventor, the first–to–file method does encourage inventors to patent early. With the first–to–invent method, additional provisions, i.e. the statutory bars (cf *infra*), have to be put into force to achieve a similar and frequently not identical effect (Scotchmer and Green 1990).

The novelty requirement as such also ensures that society only grants exclusive rights for inventions that benefit society. By conferring patent rights for an invention that already exists does not create an extra benefit. To the contrary, it would result in social costs. In case the subject matter had been patented before, monopoly rights have to be born longer than necessary. Consumer prices for associated products are higher and the deadweight loss will be greater when a patent is granted for an invention that does not pass the novelty requirement and was not patented before.

Statutory bars

Statutory bars to patents are laid down in §102(b–d) USC 35. They are not primarily concerned with any technological considerations but with external events that have taken place before the filing of the patent application and destroy novelty. In particular, such events are that subject matter has already been patented (abroad), it has been described or publicly used at least a year before the filing for a US patent or it has been abandoned.

At a first glance, the anticipation standard which determines novelty seems too narrow to guarantee a sufficiently high social benefit associated with the protected invention. Before the 1952 US Patent Act, the novelty requirement alone was responsible for preventing trivial innovations from being rewarded patents. Although the doctrine of *inventiveness* had been used for a long time, only the 1952 amendment of the Patent Act introduced a sharp distinction between novelty and non–obviousness.

Non–obviousness

Most commentators consider the non–obviousness requirement to be the heart of modern patent systems. Novelty and utility, although important, are easily met by most inventions. To satisfy the non–obviousness requirement is much harder.

The non–obviousness requirement aims at measuring an invention's contribution to the technological progress in the field and is ruled by §103 USC 35:

> A patent may not be obtained [...] if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains.

The provision states clearly how a sufficient inventive step is measured: as the difference between the invented subject matter and the prior art. Thus, an invention is unpatentable if a person skilled in the art would find the solution to the underlying problem obvious. Trivial technical changes as well as obvious combinations of prior art that would satisfy the novelty requirement are barred from patent protection by this provision. In practise, the non–obviousness requirement is not easy to apply since the *person skilled in the art* is a hypothetical construct.

Although courts have increasingly ruled in patent validity cases that an invention must have been non–obvious because it had commercial success, the non–obviousness criterion is a measure of technological progress in a specific field.[15] The latter may raise social welfare in several ways. Firstly, in case a sophisticated invention is successful in the market, consumers enjoy new or improved products. Alternatively, they may be able to buy known products at lower prices because production is possible at lower costs. In both cases, the consumers' surplus is higher even though exclusive rights inevitably create a deadweight loss. Secondly, if the invention is not a marketable product, new technological know–how is created so that follow–on research becomes possible sooner. These efforts may eventually lead to marketable products or processes that have commercial success.[16]

Since there is evidence that technical quality and social welfare are positively correlated (Reitzig 2005), should the non–obviousness standard be as high as possible to maximize social welfare? Weak non–obviousness requirements might have severe *side–effects* as the following example shows. A firm holds process patents for a therapeutical substance and has commercial success, and a second firm improves the production process slightly.[17] If the non–obviousness standard is weak, even trivial improvements receive a patent. As a consequence, the second firm is able to realize profits that do not correspond to the social value it added. However, the value of the patent and the profits of the first firm decline. Under those circumstances, rationally behaving firms may pursue suboptimally small, trivial inventions. Clark (1927) points out that firms may find it profitable to wait for a competitor to make an invention in case intellectual property is insufficiently protected. Weak non–obviousness standards may lead to this situation.

On the other hand, setting the non–obviousness standard too high might also yield suboptimal results. Assume that the second firm did achieve a major

improvement on the production process of the substance so that the drug has fewer side–effects. Denying a patent to such improvements will certainly result in lower welfare. A more detailed analysis of the *optimal non–obviousness standard* is dedicated to Chapters 10 and 13.

Disclosure and enablement

The practise of diligently describing an invention arose during the Industrial Revolution when the perception of what society gains in exchange for the exclusive rights changed (cf *supra*). Previously, it was the product or the trade that was introduced and a demonstration of operability or a prototype was frequently required. With the Industrial Revolution, society became more interested in knowledge so that it became necessary that inventions are described in such a way that they can be reproduced. In detail, §112 USC 35 states that

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same, and shall set forth the best mode contemplated by the inventor of carrying out his invention.

The main parts of the disclosure and enablement requirement are (1) the definiteness of claims, (2) the written description, (3) enablement, and (4) the best mode. With the exception of the definiteness of claims, all other requirements apply to the specification in the body of the patent application (cf *supra*).

What exactly inventors claim to have invented can be found in the specification section and the claims section of the patent. The claims are the essence of a patent and constitute the legal definition of the patent, i.e. the patent scope. They are first consulted when conflicts, as e.g. infringement or a validity challenge, arise. Therefore, the *claim definiteness* requires that claims are written in such a way that the boundaries of the legal rights that are conferred by the patent can easily be understood by someone skilled in the art.

In the specification section, the invention is more extensively described. To satisfy the *written description* requirement inventors have to make clear what sets the invention apart from existing work. At the same time, there must not be a disparity of what the inventor claims in the specification and claims section of the patent.

According to the *Enablement* requirement, applicants have not only to describe the invention itself. They have also to give extensive details on how to make the invention and how to use it. It must be written in such a way that someone skilled in the technology field can reproduce and use the invention without undue experimentation. Inventors are thus not allowed to withhold important pieces of information that enables them alone to work the invention after the patent lapses or expires.

Finally, according to the *best mode* requirement, the inventor must disclose the best mode to make and use the invention when there are certain choices, as e.g. to the pressure to be applied, materials to be used or the way certain parts are joined.

The requirements of disclosure and enablement regulate how much technological know–how is revealed to the public.[18] If the standards are low, inventors

will keep part of the information secret and thereby enjoy the best of both protective mechanisms.[19] As a consequence, researchers that want to improve the invention or develop complementary devices have to spend (socially wasteful) resources in order to find out how the original invention works in the first place. Hence, lax disclosure and enablement standards inevitably increase the social costs of the invention (see e.g. Trajtenberg 1990). Followers have not only to spend resources to rediscover the withhold information, but it takes longer until the improvements or complementary devices become available to consumers and the public at large.[20] Therefore, technological progress and economic growth can be expected to be slower when disclosure and enablement are not applied strictly.

At least in principle, high standards will also minimize disputes over patent rights.[21] The written description and the definiteness of claims requirement ensure that (1) the legal rights cover no more than what the applicant invented and (2) that the legal rights are clearly drawn. Consequently, high standards of disclosure and enablement lead to a higher certainty and encourage inventors to pursue fruitful research *and* to disclose the knowledge by seeking patent protection. At the same time, at least part of costly litigation procedures can be avoided. However, this also depends on a number of other factors.

5.2.2 Patent enforcement

After a patent has been issued, two possibilities arise for legal actions: (1) A potential infringer may seek to invalidate the patent, or (2) the patentee may sue an infringer. Usually, infringers defend themselves by claiming that the patent is invalid in the first place. Since potential infringers have the burden of proof, they have to collect evidence to cast sufficient doubt as to the validity of the patent. Hence, part of the judges' task will be to examine whether the patent indeed meets the patent requirements discussed above.

During an infringement action, courts have to establish whether the infringer's product or process does indeed violate the patentee's rights. Upon a positive finding, they have also to decide on appropriate remedies.

Infringement

§271 USC 35 specifies as to what infringement constitutes:[22]

> [...] whoever without authority makes, uses, offers to sell, or sells any patented invention, within the United States, or imports into the United States any patented invention during the term of the patent therefore, infringes the patent.

It is worth mentioning that it is irrelevant whether the accused infringers knew that the patent exists and whether they intended to violate the patentee's rights.[23] Infringement is a matter of fact. If intention is considered at all, they affect the remedies chosen by the court.

Patent claims define the legal boundaries of a patent so that they have also to be consulted when determining whether or not an accused infringer's product or process violates the patentee's rights. Claims are always literally construed, where all the relevant patent documents (description, drawings, and the entire history of the patent) as well as additional materials such as e.g. dictionaries

or expert testimonies may be used in guiding the court. However, literal infringement is rare. If infringers did know the patent they would surely change details to increase their chances in court. On the other hand, the probability that independent development leads to the same invention in each and every element is small. Yet, the patent scope as derived from the literal interpretation of the claims can be broadened or narrowed by the *Doctrine of Equivalents* and the *Reverse Doctrine of Equivalents* respectively.

The Doctrine of Equivalents rests on the presumption that if the same *functions* are used in the same *way* to obtain the same *result* the process or product must be the same. This is known as the triple identity test. To declare a patent to be infringed on the basis of the Doctrine of Equivalents, two further prerequisites have to be met. Firstly, equivalence must be established for all elements of the patent and not just some of them. Secondly, a person skilled in the art must have known that the elements in question can be used interchangeably when infringement began. Thus, the Doctrine of Equivalents protects the patent owner against minor, insubstantial variations.

The Reverse Doctrine of Equivalents achieves the opposite effect: it protects the accused infringer.[24] The triple identity test can also be used as a guideline here. Different from the Doctrine of Equivalents, however, the Reverse Doctrine of Equivalents is invoked if the same function is employed in a *substantially different way* to obtain the same result. Under these circumstances, a patent may be found not to be infringed even though the accused device falls into the literal scope of the claims.

Accused infringers usually defend themselves by challenging the patent's validity. Hence, courts ultimately decide whether or not an invention merits a patent grant. This two–step system (the administrative and the judicial part) to ensure that the invention indeed satisfies the patent requirements has advantages and disadvantages. It might actually keep overall costs low. Despite the patent offices examination, there will always be some patents those subject matter does not meet the patent requirements and have therefore been wrongly issued. However, most of them are not (commercially) valuable so that they are not disputed even if infringed. Yet, social costs in form of higher prices or increased deadweight loss are negligible because the patent is a commercial failure. Only a minority of cases where inventions are valuable are decided in courts. Therefore, the two–step procedure helps to keep the overall costs of the patent system low.

Yet, the system also generates a certain amount of uncertainty since a patent holders must face a certain probability that their patents are declared invalid by a court. Accordingly, the most sophisticated patent statute may not accomplish to foster innovations when the courts do not uphold the standards prescribed by law. In case patents are frequently revoked, inventors may turn to trade secrecy to protect their creative work. On a macro–economic level, research expenditures may be directed to other fields of technology or remain on a suboptimally low level across all fields of technology.

As far as infringement is concerned, it serves a as fine–tuning mechanism for the non–obviousness standard, apart from protecting patent holders' rights. The principle tools employed are the Doctrine of Equivalents and the Reverse Doctrine of Equivalents.

If the Doctrine of Equivalents is applied liberally, most of the patent cases will be found to be infringed. Since the doctrine effectively broadens the scope of a patent, follow–on research may be halted as even larger improvements on an

existing subject matter will probably violate the original patent. On the other hand, if the Reverse Doctrine of Equivalents is widely deployed, most of the patent cases will be found non–infringing. As a result, the expected value of the patented invention will be lower thereby generating only insufficient incentives to search for an invention in the first place or to resort to trade secrecy if possible.

Since both, the Doctrine of Equivalents and its counterpart adjust the non–obviousness standard, the considerations in Chapter 10 and 13 also apply.

Remedies

The question of remedies becomes only relevant if a patent has been found valid *and* infringed. According to the Supreme Court's opinion, the patentee should be fully compensated for the suffered injury due to the infringement. Courts have several possibilities to accomplish this, but they have a considerable discretion on what they consider appropriate and how the compensation is to be determined. The principle tools are: (1) injunction relief, (2) reasonable royalty damage, (3) lost profits damage, and (4) wilful infringement.

After a patent has been found valid and infringed, society which, after all, grants the exclusive rights in exchange for the associated technological know–how, has an obligation to ensure that no further violation of the patentee's rights continues, i.e. the patent owner has a right to an *injunction relief*. To grant a permanent injunction relief is the usual procedure. However, there are some cases in which courts denied it on reasons of public interest. Since the patentee is still entitled to compensation, the patentee and the infringer have to agree on a license. If both parties cannot reach agreement, the court will determine a suitable royalty.[25]

Generally, patent law aims at stopping any wrongdoing as soon as possible. Accordingly, the patentee may seek a preliminary injunction relief that stays in force until the trial ends.[26] However, courts are very cautious in granting a preliminary injunction relief since this could cause irrevocable damage to the accused infringer's business. After all, the findings of the court may eventually establish that the patent is either invalid or not infringed. Although the defendant has a right to be compensated for the interruption of his business, the particular line might be destroyed. The patentee has therefore to convincingly show that his patent is likely to be found valid and infringed by the court in order to get a preliminary injunction.

Reasonable royalty and lost profit damages are measures to compensate the patent owner for the injury he suffered on account of the infringement.[27] Naturally, patentees seek full compensation for any lost profits. The patentee has the burden of proof and the standards are rather high to prevent the infringer of being unduly punished. In case the court finds the patent owners evidence on the lost profits unconvincing, it calculates the damage on basis of a hypothetical license contract both parties would have struck at the time when infringement began.[28]

Finally, if the court's findings demonstrate that the infringer was aware of the patent, *wilful infringement* is presumed. In this case, the infringer is liable to treble damage.

Remedies are imposed to ensure that infringers do not profit from his violation of patentees' rights. It also encourages license agreements. However,

Table 5.1: Legal and economic patent policy instruments
Table chapter4-t-1

apart from the patentee's willingness to license, this requires that the infringer is aware of the patent before he decides to infringe. The mere awareness of the patent is sufficient to establish wilful infringement in which case the infringer is liable to treble damage. Hence, wilful infringement in its present form leads to a general disincentive to search for patents at any stage of research and development. Not only does it stand in the way of license agreements, but it is also at odds with a patent system that is otherwise designed to disseminate technological knowledge.

5.3 From patent law to economic modelling

Legal patent policy instruments do not easily translate into economic policy measures.[29]. The problem arises because the disciplines serve totally different purposes. Patent law has to define rules and norms which ensure equal treatment of all patent applications and a certain degree of consistency without losing its flexibility. The law must be such that each and every patent case can justly be dealt with. Moreover, each and every case is considered on an individual basis. Economics, on the other hand, has to abstract from particular cases to find general effects and mechanisms. Thus, patent law provides necessarily more details than economic models can and should consider.

The present work focuses only on certain legal policy instruments. They are the ones most directly related to patents as opposed to the entire patent system, i.e. patentable subject requirements, patent requirements, patent term and patent scope. In Table 5.1, they are grouped according to whether the primary test is undertaken by the patent office, i.e. before the patent is issued, or by the courts, i.e. after the patent has been issued.

The question of whether patentable subject matter should be extended is straightforward to model. More interesting insights can be gained in an international framework though. Thus, this question is relegated to Part III.

Utility, novelty, and non–obviousness are different patent requirements that ensure that a patent is only granted for inventions that make a sufficient contribution to the technological progress of a particular field. In economics, they can hardly be distinguished. The only requirement that can be operationalized is the non–obviousness standard which measures some sort of quality improvement in a technology field. It is studied in the national as well as international framework, where the results of the former also serve as a reference point for the latter.

The seemingly straightforward question of the patent life is a more complex problem in economics. We distinguish between the statutory patent life and the effective patent term. This differentiation becomes important if one considers that only around one third of all patents are maintained for the entire patent life (cf *infra*). Although a number of factors influence the effective patent life, the maintenance fee schedule is the most direct one.

As far as patent scope and infringement is concerned, it is only evaluated if disputes arise and, therefore, after a patent has been issued. In economics, a

distinction between patent scope and patent breadth is reasonable. Patent scope measures how different an alternative production process or a product must be to not infringe a certain patent. It protects against competing, *contemporary* technologies and products and, thus, against imitation.

In contrast, patent breadth is assumed to protect against *succeeding inventions* and therefore against trivial quality improvements in the same technology field. The problems of hold–up or blocking patents are closely related but not considered here (see e.g. Ellingsen and Johannesson 2004).

5.A A comparison of patent fees

Table 5.2 presents the minimum fees an inventor would have to pay to receive and maintain a patent in the US and the EU. In the US, small entities (e.g. individuals) get a 50 percent reduction on all fees. Depending on how many dependent and independent claims a patent contains, additional fees would become relevant. Here, it is assumed that the hypothetical patent has one independent claim and at least one dependent claim, but no more than ten dependent ones.

In Europe, all fees except for the designation fee are paid only once, independent of the number of EU member countries for which the patent is applied for. The maximum designation fee is 560 €, i.e. for seven countries.

Table 5.2: Patent fees in the US and the EU
Table chapter5-t-2

Source: http://www.uspto.gov/web/offices/ac/qs/fee2006may15.htm for the USPTO, http://www.european-patent-office.org/epo/fees1.htm for EPO, and Annex to §2(1) PatKostG (Law on the Fees of the German Patent and Trademark Office) for Germany.

Notes

1 Guellec and van Pottelsberghe (2007) give an in–depth inside into the European patent system.

2 See e.g. the webpage of the European Patent Office (http://www.epo.org) or of the US Patent and Trademark Office (http://www.uspto.gov) for published patent applications as well as patent documents.

3 The term *prior art* is used for the entire published body of knowledge connected to the invention.

4 See Bessen and Meurer (2008) for a critique of continuation patents and how they undermine the patent system as a whole. See also Graham (2002), Graham and Mowery (2005) or Hegde et al. (2007) for the use of continuation patents.

5 This section largely draws from Merges and Duffy (2002) for the explanations on the provisions. See also Besen and Raskind (1991).

6 For details on the patent fees see http://www.uspto.gov/web/offices/ac/qs/fee2006may15.htm.

7 See Condon and Sinha (2005) and Beckerman-Rodau (2002) for problems that arise for poorest countries when medicines belong to patentable subject matters.

8 Burk and Lemley (2003: 126) argue that the natural law doctrine is rather a restriction to permissible scope which affect all industries than a field restriction.

9 See e.g. Hahn (2005) or Lemley (2005b) who discusses intellectual property rights in new technology fields such as software, biotechnology or nanotechnology. Plotkin (2004), Campbell-Kelly and Valduriez (2005) or Burk and Lemley (2005) focus specifically on software patents.

10 See Merges (1995: 107) or Cohen et al. (2000) who both argue that only few industries rely on patents alone to capture the returns for an invention.

11 Note that trade secrecy does not protect against independent discovery. Hence, patent protection is especially attractive for products or processes that can be re–engineered easily (cf *infra*).

12 See also Plotkin (2004) on software patents and practical utility.

13 According to the US Patent Statute, new uses for known substances are protected by the patent on the known substance. Nevertheless, the new use can be protected as a process patent, i.e. the process of using the old substance for the new purpose is patented. However, the owner of the process patent must obtain a license from the holder of the product patent to work his invention (cf *infra* and Chapter 9).

14 Usually, applications on new medical substances are filed long before clinical tests are carried out. However, in principle, the utility requirement could be construed so that appropriate clinical tests become a prerequisite for a patent. The fact that the courts and Congress never required such tests seems to be a conclusive proof that they always distinguished social utility from commercial value.

15 Hunt (2004) e.g. sharply criticizes this tendency since it might effectively lower the non–obviousness standard.

16 A similar development as described by Hall and Ziedonis (2001) for computer chips is observed in the pharmaceutical and biotechnological industry: small (biotechnology) firms specialize on finding new therapeutical substances which are no marketable products yet. Subsequently, they try to find a (pharmaceutical) cooperation partner that is large and experienced enough to develop a new drug, undertake the necessary (clinical) test and produce the drugs in case a market approval is obtained.

17 A similar example is described in Merges and Duffy (2002: 995).

18 See e.g. Jansen (2005)on the economic effects of disclosure regulations, and Ghose (2006) or Gick (2004) on the effects on firms' incentives .

19 See e.g. Anton and Yao (2004) for a formal model describing this phenomenon.

20 See e.g. Bhattacharya and Guriev (2004, 2006) who model this problem.

21 Bessen and Meurer (2005) find that patent litigation is increasing tremendously. In relatively new patentable subject matters this might be expected since reliable databases for prior art are not existing yet and patent examiners in this field are still inexperienced. In older technology fields, such developments may indeed point to the use of patents for strategic reasons. See e.g. Lanjouw and Schankerman (1991). On broader issues see e.g. Lerner (2008) or Ménière and Parlane (2008).

22 The *experimental use doctrine* forms an exception. Everybody is allowed to use patented intellectual property for private experimentation, e.g. to test, perfect or improve the invention. See e.g. Hagelin (2005), Rowe (2005) or Mueller (2005).

23 See e.g. Lemley (2005a) on controversial aspects concerning the infringement. Margolis (2006) on possible benefits from infringements.

24 Some authors do not distinguish between the Doctrine of Equivalents and the Reverse Doctrine of Equivalents, but include both functions under the former expression.

25 See e.g. Lemley and Shapiro (2006) on the effects of a thread of an (preliminary) injunction relief on the outcome of license negotiations.

26 On the effects of preliminary injunction relief see e.g. Lanjouw and Lerner (2001).

27 See Choi (2006) for a discussion on how to determine reasonable royalties.

28 Schankerman and Scotchmer (2001) show e.g. that lost profit damage or reasonable royalty damage may be superior to other forms of damage payments and, in fact, benefit the patent holder.

29 See e.g. (Kahn 1940; Lemley 2004) or Burk and Lemley (2003)