Chapter 1

Introduction and Overview

1.1 Introduction

The topic of the book is the art or science of **Automatic Translation**, or **Machine Translation** (MT) as it is generally known — the attempt to automate all, or part of the process of translating from one human language to another. The aim of the book is to introduce this topic to the general reader — anyone interested in human language, translation, or computers. The idea is to give the reader a clear basic understanding of the state of the art, both in terms of *what* is currently possible, and *how* it is achieved, and of what developments are on the horizon. This should be especially interesting to anyone who is associated with what are sometimes called "the language industries"; particularly translators, those training to be translators, and those who commission or use translations extensively. But the topics the book deals with are of general and lasting interest, as we hope the book will demonstrate, and no specialist knowledge is presupposed — no background in Computer Science, Artificial Intelligence (AI), Linguistics, or Translation Studies.

Though the purpose of this book is introductory, it is not *just* introductory. For one thing, we will, in Chapter 10, bring the reader up to date with the most recent developments. For another, as well as giving an accurate picture of the state of the art, both practically and theoretically, we have taken a position on some of what seem to us to be the key issues in MT today — the fact is that we have some axes to grind.

From the earliest days, MT has been bedevilled by grandiose claims and exaggerated expectations. MT researchers and developers should stop over-selling. The general public should stop over-expecting. One of the main aims of this book is that the reader comes to appreciate where we are today in terms of actual achievement, reasonable expectation, and unreasonable hype. This is not the kind of thing that one can sum up in a catchy headline ("No Prospect for MT" or "MT Removes the Language Barrier"), but it is something one can absorb, and which one can thereafter use to distill the essence of truth that will lie behind reports of products and research.

With all this in mind, we begin (after some introductory remarks in this chapter) with a description of what it might be like to work with a hypothetical state of the art MT system. This should allow the reader to get an overall picture of what is involved, and a realistic notion of what is actually possible. The context we have chosen for this description is that of a large organization where relatively sophisticated tools are used in the preparation of documents, and where translation is integrated into document preparation. This is partly because we think this context shows MT at its most useful. In any case, the reader unfamiliar with this situation should have no trouble understanding what is involved.

The aim of the following chapters is to 'lift the lid' on the core component of an MT system to give an idea of what goes on inside — or rather, since there are several different basic designs for MT system — to give an idea of what the main approaches are, and to point out their strengths and weaknesses.

Unfortunately, even a basic understanding of what goes on inside an MT system requires a grasp of some relatively simple ideas and terminology, mainly from Linguistics and Computational Linguistics, and this has to be given 'up front'. This is the purpose of Chapter 3. In this chapter, we describe some fundamental ideas about how the most basic sort of knowledge that is required for translation can be represented in, and used by, a computer.

In Chapter 4 we look at how the main kinds of MT system actually translate, by describing the operation of the 'Translation Engine'. We begin by describing the simplest design, which we call the **transformer** architecture. Though now somewhat old hat as regards the research community, this is still the design used in most commercial MT systems. In the second part of the chapter, we describe approaches which involve more extensive and sophisticated kinds of linguistic knowledge. We call these **Linguistic Knowledge** (LK) systems. They include the two approaches that have dominated MT research over most of the past twenty years. The first is the so-called **interlingual** approach, where translation proceeds in two stages, by analyzing input sentences into some abstract and ideally language independent meaning representation, from which translations in several different languages can potentially be produced. The second is the so-called **transfer** approach, where translation proceeds in three stages, analyzing input sentences into a representation which still retains characteristics of the original, source language text. This is then input to a special component (called a transfer component) which produces a representation which has characteristics of the target (output) language, and from which a target sentence can be produced.

The still somewhat schematic picture that this provides will be amplified in the two following chapters. In Chapter 5, we focus on what is probably the single most important component in an MT system, the dictionary, and describe the sorts of issue that arise in designing, constructing, or modifying the sort of dictionary one is likely to find in an MT system.

Chapter 6 will go into more detail about some of the problems that arise in designing and building MT systems, and, where possible, describe how they are, or could be solved. This

chapter will give an idea of why MT is 'hard', of the limitations of current technology. It also begins to introduce some of the open questions for MT research that are the topic of the final chapter.

Such questions are also introduced in Chapter 7. Here we return to questions of representation and processing, which we began to look at in Chapter 3, but whereas we focused previously on morphological, syntactic, and relatively superficial semantic issues, in this chapter we turn to more abstract, 'deeper' representations — representations of various kinds of representation of meaning.

One of the features of the scenario we imagine in Chapter 2 is that texts are mainly created, stored, and manipulated electronically (for example, by word processors). In Chapter 8 we look in more detail at what this involves (or ideally would involve), and how it can be exploited to yield further benefits from MT. In particular, we will describe how standardization of electronic document formats and the general notion of standardized **markup** (which separates the content of a document from details of its realization, so that a writer, for example, specifies that a word is to be emphasised, but need not specify which typeface must be used for this) can be exploited when one is dealing with documents and their translations. This will go beyond what some readers will immediately need to know. However, we consider its inclusion important since the integration of MT into the document processing environment is an important step towards the successful use of MT. In this chapter we will also look at the benefits and practicalities of using **controlled languages** — specially simplified versions of, for example, English, and **sublanguages** — specialized languages of sub-domains. Although these notions are not central to a proper understanding of the principles of MT, they are widely thought to be critical for the successful application of MT in practice.

Continuing the orientation towards matters of more practical than theoretical importance, Chapter 9 addresses the issue of the **evaluation** of MT systems — of how to tell if an MT system is 'good'. We will go into some detail about this, partly because it is such an obvious and important question to ask, and partly because there is no other accessible discussion of the standard methods for evaluating MT systems that an interested reader can refer to.

By this time, the reader should have a reasonably good idea of what the 'state of the art' of MT is. The aim of the final chapter (Chapter 10) is to try to give the reader an idea of what the future holds by describing where MT research is going and what are currently thought to be the most promising lines of research.

Throughout the book, the reader may encounter terms and concepts with which she is unfamiliar. If necessary the reader can refer to the Glossary at the back of the book, where such terms are defined.

1.2 Why MT Matters

The topic of MT is one that we have found sufficiently interesting to spend most of our professional lives investigating, and we hope the reader will come to share, or at least understand, this interest. But whatever one may think about its intrinsic interest, it is undoubtedly an important topic — socially, politically, commercially, scientifically, and intellectually or philosophically — and one whose importance is likely to increase as the 20th Century ends, and the 21st begins.

The *social* or *political* importance of MT arises from the socio-political importance of translation in communities where more than one language is generally spoken. Here the only viable alternative to rather widespread use of translation is the adoption of a single common 'lingua franca', which (despite what one might first think) is not a particularly attractive alternative, because it involves the dominance of the chosen language, to the disadvantage of speakers of the other languages, and raises the prospect of the other languages becoming second-class, and ultimately disappearing. Since the loss of a language often involves the disappearance of a distinctive culture, and a way of thinking, this is a loss that should matter to everyone. So translation is necessary for communication — for ordinary human interaction, and for gathering the information one needs to play a full part in society. Being allowed to express yourself in your own language, and to receive information that directly affects you in the same medium, seems to be an important, if often violated, right. And it is one that depends on the availability of translation. The problem is that the demand for translation in the modern world far outstrips any possible supply. Part of the problem is that there are too few human translators, and that there is a limit on how far their productivity can be increased without automation. In short, it seems as though automation of translation is a social and political necessity for modern societies which do not wish to impose a common language on their members.

This is a point that is often missed by people who live in communities where one language is dominant, and who speak the dominant language. Speakers of English in places like Britain, and the Northern USA are examples. However, even they rapidly come to appreciate it when they visit an area where English is not dominant (for example, Welsh speaking areas of Britain, parts of the USA where the majority language is Spanish, not to mention most other countries in the world). For countries like Canada and Switzerland, and organizations like the European Community and the UN, for whom multilingualism is both a basic principle and a fact of every day life, the point is obvious.

The *commercial* importance of MT is a result of related factors. First, translation itself is commercially important: faced with a choice between a product with an instruction manual in English, and one whose manual is written in Japanese, most English speakers will buy the former — and in the case of a repair manual for a piece of manufacturing machinery or the manual for a safety critical system, this is not just a matter of taste. Secondly, translation is expensive. Translation is a highly skilled job, requiring much more than mere knowledge of a number of languages, and in some countries at least, translators' salaries are comparable to other highly trained professionals. Moreover, delays in translation are costly. Estimates vary, but producing high quality translations of difficult material,

a professional translator may average no more than about 4-6 pages of translation (perhaps 2000 words) per day, and it is quite easy for delays in translating product documentation to erode the market lead time of a new product. It has been estimated that some 40-45% of the running costs of European Community institutions are 'language costs', of which translation and interpreting are the main element. This would give a cost of something like £ 300 million per annum. This figure relates to translations actually done, and is a tiny fraction of the cost that would be involved in doing all the translations that could, or should be done.[1]

Scientifically, MT is interesting, because it is an obvious application and testing ground for many ideas in Computer Science, Artificial Intelligence, and Linguistics, and some of the most important developments in these fields have begun in MT. To illustrate this: the origins of Prolog, the first widely available logic programming language, which formed a key part of the Japanese 'Fifth Generation' programme of research in the late 1980s, can be found in the 'Q-Systems' language, originally developed for MT.

Philosophically, MT is interesting, because it represents an attempt to automate an activity that can require the full range of human knowledge — that is, for any piece of human knowledge, it is possible to think of a context where the knowledge is required. For example, getting the correct translation of *negatively charged electrons and protons* into French depends on knowing that protons are positively charged, so the interpretation cannot be something like "negatively charged electrons and negatively charged protons". In this sense, the extent to which one can automate translation is an indication of the extent to which one can automate 'thinking'.

Despite this, very few people, even those who are involved in producing or commissioning translations, have much idea of what is involved in MT today, either at the practical level of what it means to have and use an MT system, or at the level of what is technically feasible, and what is science fiction. In the whole of the UK there are perhaps five companies who use MT for making commercial translations on a day-to-day basis. In continental Europe, where the need for commercial translation is for historical reasons greater, the number is larger, but it still represents an extremely small proportion of the overall translation effort that is actually undertaken. In Japan, where there is an enormous need for translation of Japanese into English, MT is just beginning to become established on a commercial scale, and some familiarity with MT is becoming a standard part of the training of a professional translator.

Of course, theorists, developers, and sellers of MT systems must be mainly responsible for this level of ignorance and lack of uptake, and we hope this book will help here — one motivation for writing this book was our belief that an understanding of MT is an essential part of the equipment of a professional translator, and the knowledge that no other book provided this in accessible form.

We are reminded of this scale of ignorance every time we admit to working in the field of MT. After initial explanations of what MT is, the typical reaction is one of two contra-

[1] These estimates of CEC translation costs are from Patterson (1982).

dictory responses (sometimes one gets both together). One is "But that's impossible — no machine could ever translate Shakespeare." The other is "Yes, I saw one of those in the Duty Free Shop when I went on holiday last summer." These reactions are based on a number of misconceptions that are worth exposing. We will look at these, as well as some correct conceptions, in the next section.

1.3 Popular Conceptions and Misconceptions

Some popular misconceptions about MT are listed on page 7. We will discuss them in turn.

• "MT is a waste of time because you will never make a machine that can translate Shakespeare".

The criticism that MT systems cannot, and will never, produce translations of great literature of any great merit is probably correct, but quite beside the point. It certainly does not show that MT is impossible. First, translating literature requires special literary skill — it is not the kind of thing that the average professional translator normally attempts. So accepting the criticism does not show that automatic translation of non-literary texts is impossible. Second, literary translation is a small proportion of the translation that has to be done, so accepting the criticism does not mean that MT is useless. Finally, one may wonder who would ever *want* to translate Shakespeare by machine — it is a job that human translators find challenging and rewarding, and it is not a job that MT systems have been designed for. The criticism that MT systems cannot translate Shakespeare is a bit like criticism of industrial robots for not being able to dance Swan Lake.

• "There was/is an MT system which translated *The spirit is willing, but the flesh is weak* into the Russian equivalent of *The vodka is good, but the steak is lousy*, and *hydraulic ram* into the French equivalent of *water goat*. MT is useless."

The 'spirit is willing' story is amusing, and it really is a pity that it is not true. However, like most MT 'howlers' it is a fabrication. In fact, for the most part, they were in circulation long before any MT system could have produced them (variants of the 'spirit is willing' example can be found in the American press as early as 1956, but sadly, there does not seem to have been an MT system in America which could translate from English into Russian until much more recently — for sound strategic reasons, work in the USA had concentrated on the translation of Russian into English, not the other way round). Of course, there are real MT howlers. Two of the nicest are the translation of French *avocat* ('advocate', 'lawyer' or 'barrister') as *avocado*, and the translation of *Les soldats sont dans le café* as *The soldiers are in the coffee*. However, they are not as easy to find as the reader might think, and they certainly do not show that MT is useless.

• "Generally, the quality of translation you can get from an MT system is very low. This makes them useless in practice."

Some Popular Misconceptions about MT

- **False:** MT is a waste of time because you will never make a machine that can translate Shakespeare.
- **False:** There was/is an MT system which translated *The spirit is willing, but the flesh is weak* into the Russian equivalent of *The vodka is good, but the steak is lousy*, and *hydraulic ram* into the French equivalent of *water goat*. MT is useless.
- **False:** Generally, the quality of translation you can get from an MT system is very low. This makes them useless in practice.
- **False:** MT threatens the jobs of translators.
- **False:** The Japanese have developed a system that you can talk to on the phone. It translates what you say into Japanese, and translates the other speaker's replies into English.
- **False:** There is an amazing South American Indian language with a structure of such logical perfection that it solves the problem of designing MT systems.
- **False:** MT systems are machines, and buying an MT system should be very much like buying a car.

Far from being useless, there are several MT systems in day-to-day use around the world. Examples include METEO (in daily since 1977 use at the Canadian Meteorological Center in Dorval, Montreal), SYSTRAN (in use at the CEC, and elsewhere), LOGOS, ALPS, ENGSPAN (and SPANAM), METAL, GLOBALINK. It is true that the number of organizations that use MT on a daily basis is relatively small, but those that do use it benefit considerably. For example, as of 1990, METEO was regularly translating around 45 000 words of weather bulletins every day, from English into French for transmission to press, radio, and television. In the 1980s, the diesel engine manufacturers Perkins Engines was saving around £ 4 000 on each diesel engine manual translated (using a PC version of WEIDNER system). Moreover, overall translation time per manual was more than halved from around 26 weeks to 9-12 weeks — this time saving can be very significant commercially, because a product like an engine cannot easily be marketed without user manuals.

Of course, it is true that the quality of many MT systems is low, and probably no existing system can produce really perfect translations.[2] However, this does not make MT useless.

[2]In fact, one can get perfect translations from one kind of system, but at the cost of radically restricting what an author can say, so one should perhaps think of such systems as (multilingual) text creation aids, rather than MT systems. The basic idea is similar to that of a phrase book, which provides the user with a collection of 'canned' phrases to use. This is fine, provided the canned text contains what the user wants to

First, not every translation has to be perfect. Imagine you have in front of you a Chinese newspaper which you suspect may contain some information of crucial importance to you or your company. Even a very rough translation would help you. Apart from anything else, you would be able to work out which, if any, parts of the paper would be worth getting translated properly. Second, a human translator normally does not immediately produce a perfect translation. It is normal to divide the job of translating a document into two stages. The first stage is to produce a draft translation, i.e. a piece of running text in the target language, which has the most obvious translation problems solved (e.g. choice of terminology, etc.), but which is not necessarily perfect. This is then revised — either by the same translator, or in some large organizations by another translator — with a view to producing something that is up to standard for the job in hand. This might involve no more than checking, or it might involve quite radical revision aimed at producing something that reads as though written originally in the target language. For the most part, the aim of MT is only to automate the first, draft translation process.[3]

• "MT threatens the jobs of translators."

The quality of translation that is currently possible with MT is one reason why it is wrong to think of MT systems as dehumanizing monsters which will eliminate human translators, or enslave them. It will not eliminate them, simply because the volume of translation to be performed is so huge, and constantly growing, and because of the limitations of current and forseeable MT systems. While not an immediate prospect, it could, of course, turn out that MT enslaves human translators, by controlling the translation process, and forcing them to work on the problems it throws up, at its speed. There are no doubt examples of this happening to other professions. However, there are not many such examples, and it is not likely to happen with MT. What is more likely is that the process of producing draft translations, along with the often tedious business of looking up unknown words in dictionaries, and ensuring terminological consistency, will become automated, leaving human translators free to spend time on increasing clarity and improving style, and to translate more important and interesting documents — editorials rather than weather reports, for example. This idea borne out in practice: the job satisfaction of the human translators in the Canadian Meteorological Centerimproved when METEO was installed, and their job became one of checking and trying to find ways to improve the system output, rather than translating the weather bulletins by hand (the concrete effect of this was a greatly reduced turnover in translation staff at the Center).

• "The Japanese have developed a system that you can talk to on the phone. It translates what you say into Japanese, and translates the other speaker's replies into English."

The claim that the Japanese have a speech to speech translation system, of the kind described above, is pure science fiction. It is true that speech-to-speech translation is a topic of current research, and there are laboratory prototypes that can deal with a very restricted range of questions. But this research is mainly aimed at investigating how the

say. Fortunately, there are some situations where this is the case.

[3] Of course, the sorts of errors one finds in draft translations produced by a human translator will be rather different from those that one finds in translations produced by machine.

various technologies involved in speech and language processing can be integrated, and is limited to very restricted domains (hotel bookings, for example), and messages (offering little more than a phrase book in these domains). It will be several years before even this sort of system will be in any sort of real use. This is partly because of the limitations of speech systems, which are currently fine for recognizing isolated words, uttered by a single speaker, for which the system has been specially trained, in quiet conditions, but which do not go far beyond this. However, it is also because of the limitations of the MT system (see later chapters).

• "There is an amazing South American Indian language with a structure of such logical perfection that it solves the problem of designing MT systems."

The South American Indian language story is among the most irritating for MT researchers. First, the point about having a 'perfectly logical structure' is almost certainly completely false. Such perfection is mainly in the eye of the beholder — Diderot was convinced that the word order of French exactly reflected the order of thought, a suggestion that non-French speakers do not find very convincing. What people generally mean by this is that a language is very simple to describe. Now, as far as anyone can tell all human languages are pretty much as complicated as each other. It's hard to be definite, since the idea of simplicity is difficult to pin down, but the general impression is that if a language has a very simple syntax, for example, it will compensate by having a more complicated morphology (word structure), or phonology (sound structure).[4] However, even if one had a very neat logical language, it is hard to see that this would solve the MT problem, since one would still have to perform automatic translation into, and out of, this language.

• "MT systems are machines, and buying an MT system should be very much like buying a car."

There are really two parts to this misconception. The first relates to the sense in which MT systems are machines. They are, of course, but only in the sense that modern word processors are machines. It is more accurate to think of MT systems as *programs* that run on computers (which really are machines). Thus, when one talks about buying, modifying, or repairing an MT system, one is talking about buying, modifying or repairing a piece of *software*. It was not always so — the earliest MT systems were dedicated machines, and even very recently, there were some MT vendors who tried to sell their systems with specific hardware, but this is becoming a thing of the past. Recent systems can be installed on different types of computers. The second part of the misconception is the idea that one would take an MT system and 'drive it away', as one would a car. In fact, this is unlikely to be possible, and a better analogy is with buying a house — what one buys may be immediately habitable, but there is a considerable amount of work involved in adapting it to one's own special needs. In the case of a house this might involve changes to the decor and plumbing. In the case of an MT system this will involve additions to

[4] Of course, some languages have larger vocabularies than others, but this is mainly a matter of how many things the language is used to talk about (not surprisingly, the vocabulary which Shakespeare's contemporaries had for discussing high-energy physics was rather impoverished), but all languages have ways of forming new words, and this has nothing to do with logical perfection.

the dictionaries to deal with the vocabulary of the subject area and possibly the type of text to be translated. There will also be some work involved in integrating the system into the rest of one's document processing environment. More of this in Chapters 2 and 8. The importance of customization, and the fact that changes to the dictionary form a major part of the process is one reason why we have given a whole chapter to discussion of the dictionary (Chapter 5).

Against these misconceptions, we should place the genuine facts about MT. These are listed on page 11.

The correct conclusion is that MT, although imperfect, is not only a possibility, but an actuality. But it is important to see the product in a proper perspective, to be aware of its strong points and shortcomings.

Machine Translation started out with the hope and expectation that most of the work of translation could be handled by a system which contained all the information we find in a standard paper bilingual dictionary. Source language words would be replaced with their target language translational equivalents, as determined by the built-in dictionary, and where necessary the order of the words in the input sentences would be rearranged by special rules into something more characteristic of the target language. In effect, correct translations suitable for immediate use would be manufactured in two simple steps. This corresponds to the view that translation is nothing more than word substitution (determined by the dictionary) and reordering (determined by reordering rules).

Reason and experience show that 'good' MT cannot be produced by such delightfully simple means. As all translators know, word for word translation doesn't produce a satisfying target language text, not even when some local reordering rules (e.g. for the position of the adjective with regard to the noun which it modifies) have been included in the system. Translating a text requires not only a good knowledge of the vocabulary of both source and target language, but also of their grammar — the system of rules which specifies which sentences are well-formed in a particular language and which are not. Additionally it requires some element of **real world knowledge** — knowledge of the nature of things out in the world and how they work together — and technical knowledge of the text's subject area. Researchers certainly believe that much can be done to satisfy these requirements, but producing systems which actually do so is far from easy. Most effort in the past 10 years or so has gone into increasing the subtlety, breadth and depth of the linguistic or grammatical knowledge available to systems. We shall take a more detailed look at these developments in due course.

In growing into some sort of maturity, the MT world has also come to realize that the 'text in $\rightarrow$ translation out' assumption — the assumption that MT is solely a matter of switching on the machine and watching a faultless translation come flying out — was rather too naive. A translation process starts with providing the MT system with *usable* input. It is quite common that texts which are submitted for translation need to be adapted (for example, typographically, or in terms of format) before the system can deal with them. And when a text can actually be submitted to an MT system, and the system produces a

Some Facts about MT

- **True:** MT is useful. The METEO system has been in daily use since 1977. As of 1990, it was regularly translating around 45 000 words daily. In the 1980s, The diesel engine manufacturers Perkins Engines was saving around £ 4 000 and up to 15 weeks on each manual translated.

- **True:** While MT systems sometimes produce howlers, there are many situations where the ability of MT systems to produce reliable, if less than perfect, translations at high speed is valuable.

- **True:** In some circumstances, MT systems can produce good quality output: less than 4% of METEO output requires any correction by human translators at all (and most of these are due to transmission errors in the original texts). Even where the quality is lower, it is often easier and cheaper to revise 'draft quality' MT output than to translate entirely by hand.

- **True:** MT does not threaten translators' jobs. The need for translation is vast and unlikely to diminish, and the limitations of current MT systems are too great. However, MT systems can take over some of the boring, repetitive translation jobs and allow human translation to concentrate on more interesting tasks, where their specialist skills are really needed.

- **True:** Speech-to-Speech MT is still a research topic. In general, there are many open research problems to be solved before MT systems will be come close to the abilities of human translators.

- **True:** Not only are there are many open research problems in MT, but building an MT system is an arduous and time consuming job, involving the construction of grammars and very large monolingual and bilingual dictionaries. There is no 'magic solution' to this.

- **True:** In practice, before an MT system becomes really useful, a user will typically have to invest a considerable amount of effort in customizing it.

translation, the output is almost invariably deemed to be grammatically and translationally imperfect. Despite the increased complexity of MT systems they will never — within the forseeable future — be able to handle all types of text reliably and accurately. This normally means that the translation will have to be corrected (post-edited) and usually the person best equipped to do this is a translator.

This means that MT will only be profitable in environments that can exploit the strong points to the full. As a consequence, we see that the main impact of MT in the immediate future will be in large corporate environments where substantial amounts of translation are performed. The implication of this is that MT is not (yet) for the individual self-employed translator working from home, or the untrained lay-person who has the occasional letter to write in French. This is not a matter of cost: MT systems sell at anywhere between a few hundred pounds and over £ 100 000. It is a matter of effective use. The aim of MT is to achieve faster, and thus cheaper, translation. The lay-person or self-employed translator would probably have to spend so much time on dictionary updating and/or post-editing that MT would not be worthwhile. There is also the problem of getting input texts in machine readable form, otherwise the effort of typing will outweigh any gains of automation. The real gains come from integrating the MT system into the whole document processing environment (see Chapter 2), and they are greatest when several users can share, for example, the effort of updating dictionaries, efficiencies of avoiding unnecessary retranslation, and the benefits of terminological consistency.

Most of this book is about MT today, and to some extent tomorrow. But MT is a subject with an interesting and dramatic past, and it is well worth a brief description.

1.4 A Bit of History

There is some dispute about who first had the idea of translating automatically between human languages, but the actual development of MT can be traced to conversations and correspondence between Andrew D. Booth, a British crystallographer, and Warren Weaver of the Rockefeller Foundation in 1947, and more specifically to a memorandum written by Weaver in 1949 to the Rockerfeller Foundation which included the following two sentences.

> "I have a text in front of me which is written in Russian but I am going to pretend that it is really written in English and that it has been coded in some strange symbols. All I need to do is strip off the code in order to retrieve the information contained in the text."

The analogy of translation and decoding may strike the sophisticated reader as simplistic (however complicated coding gets it is still basically a one-for-one substitution process where there is only one right answer — translation is a far more complex and subtle business), and later in the memorandum Weaver proposed some other more sophisticated

views,[5] but it had the virtue of turning an apparently difficult task into one that could be approached with the emergent computer technology (there had been considerable success in using computers in cryptography during the Second World War). This memorandum sparked a significant amount of interest and research, and by the early 1950s there was a large number of research groups working in Europe and the USA, representing a significant financial investment (equivalent to around £,20 000 000). But, despite some success, and the fact that many research questions were raised that remain important to this day, there was widespread disappointment on the part of funding authorities at the return on investment that this represented, and doubts about the possibility of automating translation in general, or at least in the current state of knowledge.

The theoretical doubts were voiced most clearly by the philosopher Bar-Hillel in a 1959 report, where he argued that fully automatic, high quality, MT (FAHQMT) was impossible, not just at present, but in *principle*. The problem he raised was that of finding the right translation for *pen* in a context like the following:

(1) Little John was looking for his toy box. Finally he found it. The box was in the pen. John was very happy.

The argument was that (i) here *pen* could only have the interpretation *play-pen*, not the alternative *writing instrument* interpretation, (ii) this could be critical in deciding the correct translation for *pen*, (iii) discovering this depends on general knowledge about the world, and (iv) there could be no way of building such knowledge into a computer. Some of these points are well taken. Perhaps FAHQMT is impossible. But this does not mean that any form of MT is impossible or useless, and in Chapter 7 we will look at some of the ways one might go about solving this problem. Nevertheless, historically, this was important in suggesting that research should focus on more fundamental issues in the processing and understanding of human languages.

The doubts of funding authorities were voiced in the report which the US National Academy of Sciences commissioned in 1964 when it set up the Automatic Language Processing Advisory Committee (ALPAC) to report on the state of play with respect to MT as regards quality, cost, and prospects, as against the existing cost of, and need for translation. Its report, the so-called *ALPAC Report*, was damning, concluding that there was no shortage of human translators, and that there was no immediate prospect of MT producing useful translation of general scientific texts. This report led to the virtual end of Government funding in the USA. Worse, it led to a general loss of morale in the field, as early hopes were perceived to be groundless.

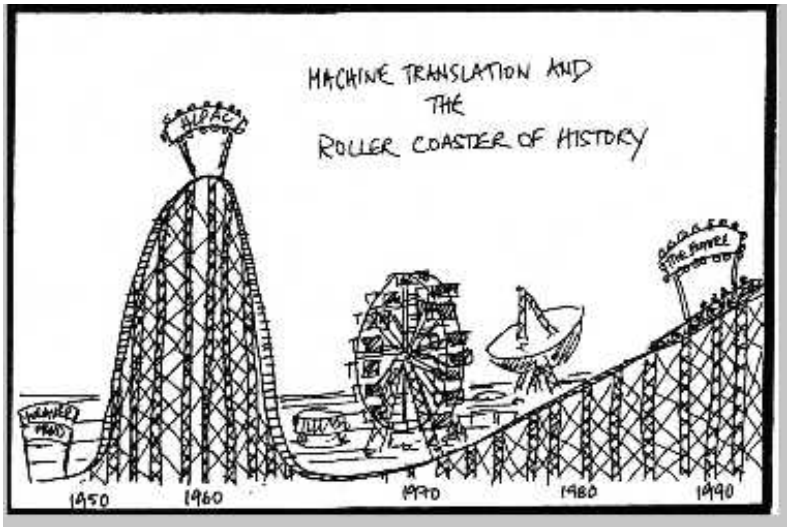
The spectre of the ALPAC report, with its threats of near complete withdrawal of funding, and demoralization, still haunts workers in MT. Probably it should not, because the achievements of MT are real, even if they fall short of the idea of FAHQMT all the time

[5] Weaver described an analogy of individuals in tall closed towers who communicate (badly) by shouting to each other. However, the towers have a common foundation and basement. Here communication is easy: "Thus it may be true that the way to translate ... is not to attempt the direct route, shouting from tower to tower. Perhaps the way is to descend, from each language, down to the common base of human communication — the real but as yet undiscovered universal language."

— useful MT is neither science fiction, nor merely a topic for scientific speculation. It is a daily reality in some places, and for some purposes. However, the fear is understandable, because the conclusion of the report was almost entirely mistaken. First, the idea that there was no need for machine translation is one that should strike the reader as absurd, given what we said earlier. One can only understand it in the anglo-centric context of cold-war America, where the main reason to translate was to gain intelligence about Soviet activity. Similarly, the suggestion that there was no prospect of successful MT seems to have been based on a narrow view of FAHQMT — in particular, on the idea that MT which required revision was not 'real' MT. But, keeping in mind the considerable time gain that can be achieved by automating the draft translation stage of the process, this view is naive. Moreover, there were, even at the time the report was published, three systems in regular, if not extensive, use (one at the Wright Patterson USAF base, one at the Oak Ridge Laboratory of the US Atomic Energy Commission, and one the EURATOM Centre at Ispra in Italy).

Nevertheless, the central conclusion that MT did not represent a useful goal for research or development work had taken hold, and the number of groups and individuals involved in MT research shrank dramatically. For the next ten years, MT research became the preserve of groups funded by the Mormon Church, who had an interest in bible translation (the work that was done at Brigham Young University in Provo, Utah ultimately led to the WEIDNER and ALPS systems, two notable early commercial systems), and a handful of groups in Canada (notably the TAUM group in Montreal, who developed the METEO system mentioned earlier), the USSR (notably the groups led by Mel'čuk, and Apresian), and Europe (notably the GETA group in Grenoble, probably the single most influential group of this period, and the SUSY group in Saarbrücken). A small fraction of the funding and effort that had been devoted to MT was put into more fundamental research on Computational Linguistics, and Artificial Intelligence, and some of this work took MT as a long term objective, even in the USA (Wilks' work on AI is notable in this respect). It was not until the late 1970s that MT research underwent something of a renaissance.

There were several signs of this renaissance. The Commission of the European Communities (CEC) purchased the English-French version of the SYSTRAN system, a greatly improved descendent of the earliest systems developed at Georgetown University (in Washington, DC), a Russian-English system whose development had continued throughout the lean years after ALPAC, and which had been used by both the USAF and NASA. The CEC also commissioned the development of a French-English version, and Italian-English version. At about the same time, there was a rapid expansion of MT activity in Japan, and the CEC also began to set up what was to become the EUROTRA project, building on the work of the GETA and SUSY groups. This was perhaps the largest, and certainly among the most ambitious research and development projects in Natural Language Processing. The aim was to produce a 'pre-industrial' MT system of advanced design (what we call a Linguistic Knowledge system) for the EC languages. Also in the late 1970s the Pan American Health Organization (PAHO) began development of a Spanish-English MT system (SPANAM), the United States Air Force funded work on the METAL system at the Linguistics Research Center, at the University of Texas in Austin, and the results of work at the TAUM group led to the installation of the METEO system. For the most part, the history of the 1980s in MT is the history of these initiatives, and the exploitation of results



Machine Translation and the Roller Coaster of History

in neighbouring disciplines.

As one moves nearer to the present, views of history are less clear and more subjective. Chapter 10 will describe what we think are the most interesting and important technical innovations. As regards the practical and commercial application of MT systems. The systems that were on the market in the late 1970s have had their ups and downs, but for commercial and marketing reasons, rather than scientific or technical reasons, and a number of the research projects which were started in the 1970s and 1980s have led to working, commercially available systems. This should mean that MT is firmly established, both as an area of legitimate research, and a useful application of technology. But researching and developing MT systems is a difficult task both technically, and in terms of management, organization and infrastructure, and it is an expensive task, in terms of time, personnel, and money. From a technical point of view, there are still fundamental problems to address. However, all of this is the topic of the remainder of this book.

1.5 Summary

This chapter has given an outline of the rest of the book, and given a potted history of MT. It has also tried to lay a few ghosts, in the form of misconceptions which haunt the enterprise. Above all we hope to convince the reader that MT is possible and potentially useful, despite current limitations.

1.6 Further Reading

A broad, practically oriented view of the field of current MT by a variety of authors can be found in Newton (1992a). Generally speaking, the best source of material that takes an MT user's viewpoint is the series of books titled *Translating and the Computer*, with various editors and publishers, including Lawson (1982a), Snell (1979), Snell (1982), Lawson (1982b), Picken (1985), Picken (1986), Picken (1987), Picken (1988), Mayorcas (1990), Picken (1990), and Mayorcas (Forthcoming). These are the published proceedings of the annual Conference on Translating and the Computer, sponsored by Aslib (The Association for Information Management), and the Institute for Translation and Interpreting.

By far the best technical introduction to MT is Hutchins and Somers (1992). This would be appropriate for readers who want to know more technical and scientific details about MT, and we will often refer to it in later chapters. This book contains useful discussions of some of the main MT systems, but for descriptions of these systems by their actual designers the reader should look at Slocum (1988), and King (1987). Slocum's introduction to the former, Slocum (1986), is particularly recommended as an overview of the key issues in MT. These books all contain detailed descriptions of the research of the TAUM group which developed the METEO system referred to in section 1.3. The METEO system is discussed further in Chapter 8.

A short assessment of the current state of MT in terms of availability and use of systems in Europe, North America, and Japan and East Asia can be found in Pugh (1992). An up-to-date picture of the state of MT as regards both commercial and scientific points of view is provided every two years by the *Machine Translation Summits*. A report of one of these can be found in Nagao (1989). There is a description of the successful use of MT in a corporate setting in Newton (1992b).

On the history of MT (which we have outlined here, but which will not be discussed again), the most comprehensive discussion can be found in Hutchins (1986), though there are also useful discussions in Warwick (1987), and Buchmann (1987). Nagao (1986) also provides a useful insight into the history of MT, together with a general introduction to MT. The ALPAC report is Pierce and Carroll (1966). The work of Wilks' that is referred to in section 1.4 is Wilks (1973).

For general descriptions and discussion of the activity of translation (both human and machine) Picken (1989) is a useful and up-to-date source. This contains references to (for example) works on translation theory, and gives a great deal of practical information of value to translators (such as lists national translators' and interpreters' organizations, and bibliographies of translations).

For up-to-date information about the state of MT, there is the newsletter of the International Association for Machine Translation *MT News International*. See the list of addresses on page 207.