Chapter 6

Further Outlook

As alluded before it is necessary to move towards a common localization standard. At the moment the emerging open systems technology spreads across all types of computer systems. In order to get this technology working across different national computer environments there should be a common standard for, e.g., character sets, locales and EDI (electronic data interchange). A development in this direction is the Unicode (ISO 10646) and the, not so common, TAD (Tron Application Databus character code set).

Other future developments, in Japan, are new input methods like pen input. In addition the development of internationalized software or computer environments will become more and more important. First I want to talk about the new character set proposals for world wide use.

6.1 Unicode & TAD

When the US computer hardware and software vendors discovered that about fifty percent of their revenue came from outside the USA they started thinking about a world wide character set standard. This is understandable because a world wide standard would satisfy all requirements of the different national language environments. In addition it would make the software development and localization for different languages much easier. One of the proposals is the ISO 10646 or, so called, Unicode ([18]).

The Unicode was developed mainly by US vendors. Nevertheless contributions came from all over the world. The vendors like Microsoft, SUN and Apple showed a great commitment to the development of this standard. The new versions of their operating systems will probably support this code. Unicode works with a BMP (Basic Multilingual Plane). Each plane has 256×256 character positions. There are two proposals for this character set. The UCS (Universal Coded character Set) 2-byte and the UCS 4-byte structure. The UCS-2 supports up to 65536 character positions and the UCS-4 supports a maximum of 2,147,483,647 character positions. Unicode provides the user with control characters, several different national characters, symbols, the, so called, unified CJK (Chinese Japanese Korean) ideographs and an area for private use.

The advantage of the Unicode is that it supports all alphabetical languages on the world and Kanji characters (see figure 6.1 on page 186). The, maybe, disadvantage is that the Chinese, Korean and Japanese Kanji character sets were moved together (unified). The ideas of collecting all Kanji characters from these languages and delete the characters which appear twice or more often sounds good, but you have to consider the different cultural backgrounds. A Kanji which has the meaning A in Japanese could have a totally different meaning in Chinese or Korean. A lot of Asian people are not happy with this approach.

Furthermore the UNIX Systems Laboratory sees, at the moment, no need to support this character standard. It is understandable because they spent a lot of time and money to develop the EUC and the supporting routines. Nevertheless it looks like that the Unicode is one of the best guesses of a new world wide character code set standard.

An other proposed character set is the TAD (Tron Application Databus character code set, see figure 6.2 on page 187). This character set is a proposal from the Tron laboratory of Ken Sakamura (Tokyo University). In this two byte matrix code set it will be possible to handle multiple languages simultaneously ([21]). The TAD character code set is a part of the Japanese Tron project which aims towards a total computer architecture. The Tron project will define, e.g., an international standard for formatting documents, data exchange, data-types and an international character code set (the TAD Multi-lingual character code set). It will support special control codes, different writing directions and a general framework for switching between this language specific environments. In addition it will provide a general framework for

character input.

At the moment it supports the English and the Japanese language. Unfortunately is Tron a development which takes mainly place in Japan. This causes that the Tron project is not well known outside Japans. In addition the most foreign vendors do not support any of the Tron features at the moment. Some foreign companies are observing the activity of the Tron research group, but at the moment the strongest support comes from Japanese companies.

6.2 Pen Computing

The difficulties of entering Kanji characters are only partly solved by the use of FEPs. In the office world a lot of the communication is still done by handwriting. The main problem is that it is not very comfortable to enter Kanji characters via a keyboard. The development in Japan goes towards better input methods.

One of the new input methods, with the most impact, is the pen input method. It would be like writing on paper and this would make it more comfortable for the Japanese to enter their Kanji characters. The Kanji characters are always drawn by a fixed stroke order. This makes easier for a pen input recognizing program to detect which character is meant. The development started in Japan about 1970 and today it looks like that pen input is the alternative input method of the future.

6.3 Asianization / Internationalization

In a fast moving and changing market like the computer business it will no longer be enough to be present in your home country. The internationalization or globalization of computer systems will open the world wide computer market.

Use this change or lose it. If you not start to work out a concept for the development of global software you will be soon hit by the high costs for adapting your software to different national language environments. In the meantime fifty percent of the revenues from US vendors come from abroad. This is a good sign that the new market for software is global.

6.3.1 Asianization

As you have seen in the market section the Japanese market is the biggest market in the Asia/Pacific basin. If you have successfully japanized a product you could use Japan as a base to conquer the rest of the Asian computer market.

Despite the fact that there are differences in the computer systems of Japan, Korea, China and Taiwan they have a lot of things in common. All of them must use a DBCS character set, because all of them rely on Kanji characters. There are different input methods, but in a system like the globalized UNIX versions, your system has not to take care of that. A tool like NLIO supports all of this different country specific specialities.

6.3.2 Internationalization / Globalization

If you work on software you should always consider to use the globalized approach. It does not take a long time to develop a product for the world market if you start using localization features from the beginning. If you later on want or have to localize / globalize your product it will be much more expensive and maybe impossible.

The one source concept, which is possible through the use of NLS and NLIO, makes it easy to do this from the beginning. It does not cost much but the reward and the money, which you will save in the future, is worth to do it.

In the PC/WKS market it is visible that both standard operating systems will support the development of internationalized software in the near future. UNIX is already prepared to support localization. The newer versions of MS-Windows show a strong move towards this direction.

I the near future software will be a global business and no longer restricted to one country or area. It is absolutely necessary to change the mind of software designers and programmers to the global approach for software development.

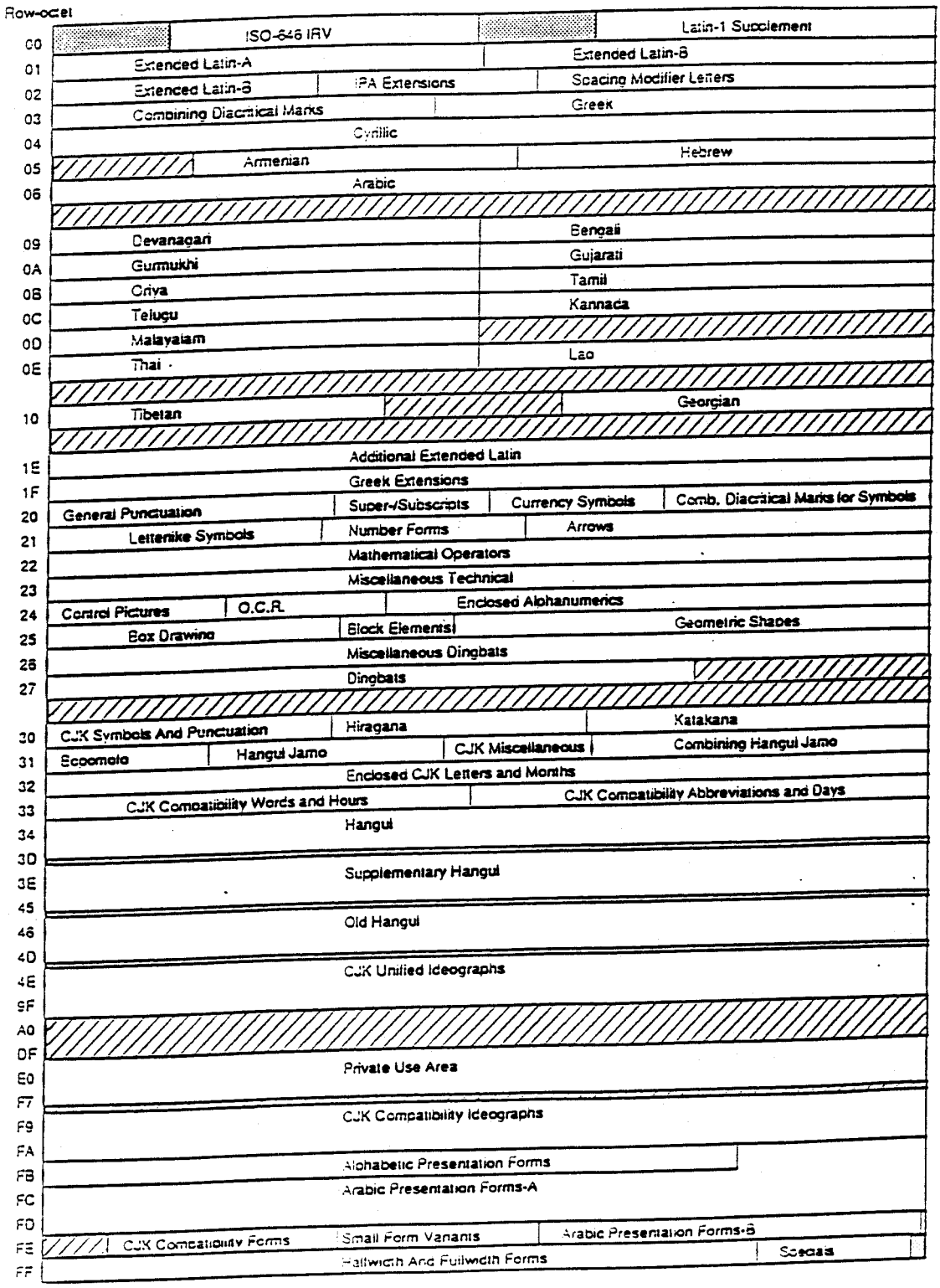
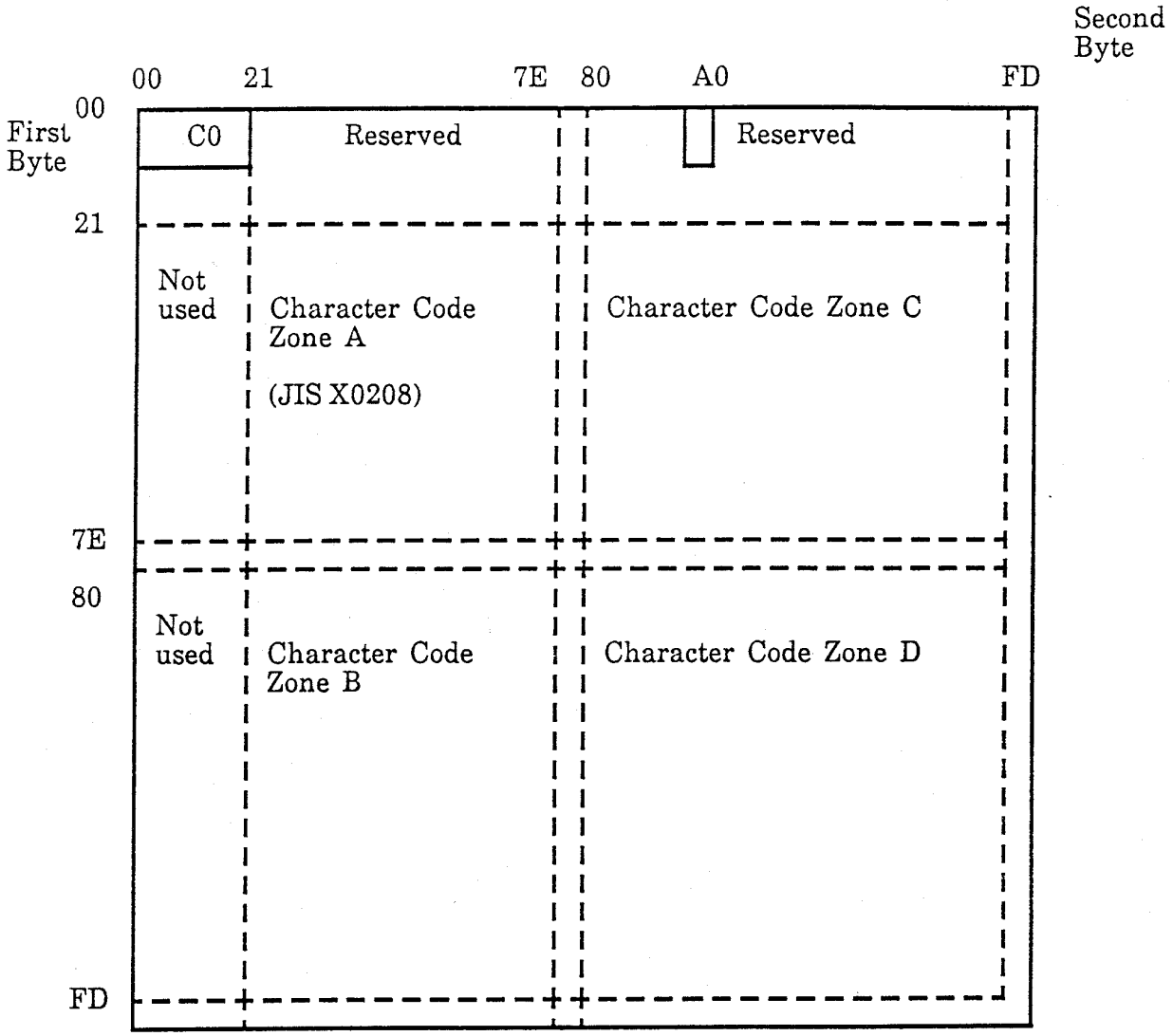


Figure 6.1: Unicode BMP



BTRON1 Japanese language Code Table

Figure 6.2: