Technical Difficulties in dealing with Asian Languages

Presented by: Don Shin

1-Stop Translation

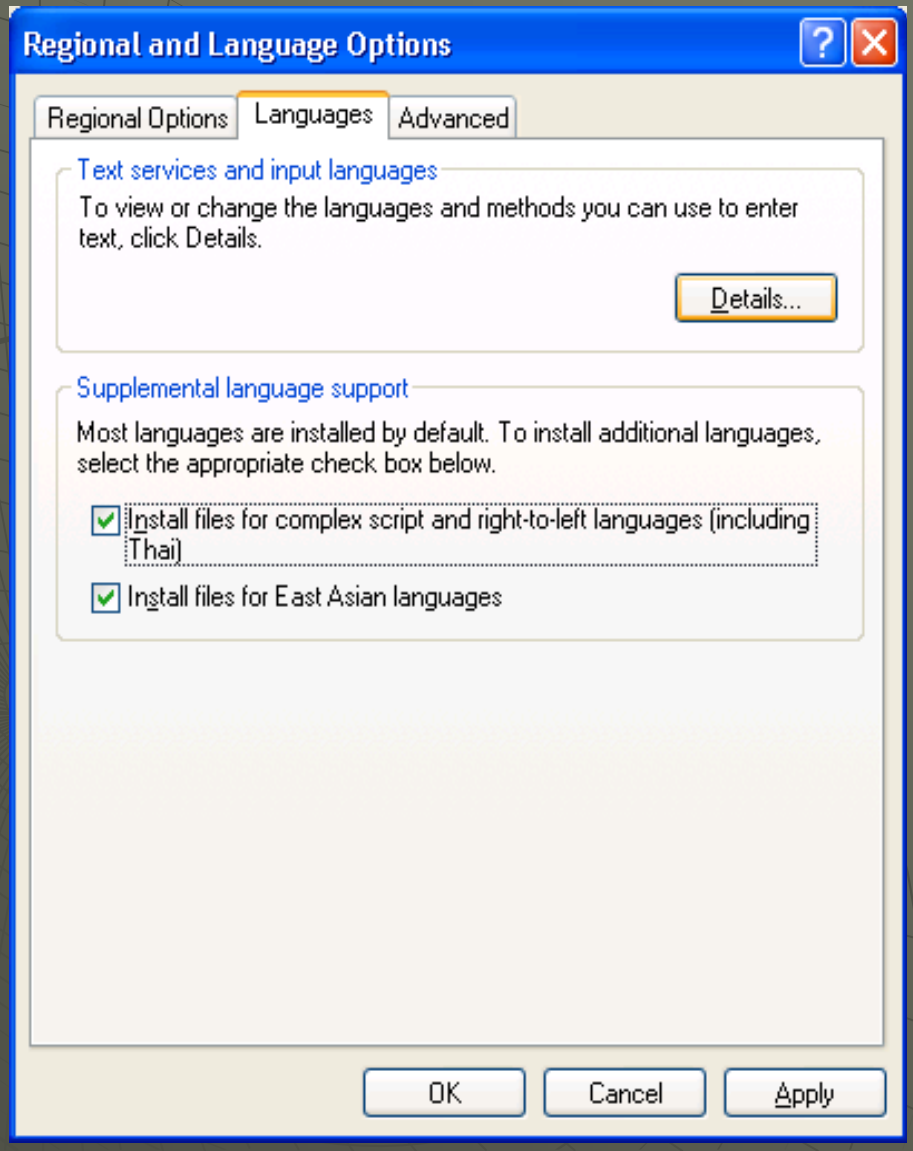
Using CJK in English System

1. Windows 2000

Double-click My Computer, Control Panel and Regional Settings. Check Chinese, Japanese and/or Korean under Language Settings in the General tab and click OK.

2. Windows XP

XP: start-control panel-Date, Time, Language, and Regional Options-languages-details



You will find the following TTC and TTF files in your WINDOWS\FONTS directory:

Traditional Chinese	***mingliu.ttc***	MingLiU & PMingLiU
Simplified Chinese	***simsun.ttc***	SimSun & NSimSun
	simhei.ttf	SimHei
Japanese	***msmincho.ttc***	MS Mincho & MS PMincho
	msgothic.ttc	MS Gothic, MS PGothic & MS UI Gothic
Korean	***batang.ttc***	Batang, BatangChe; Gungsuh, GungsuhChe
	Gulim.ttc	Gulim, GulimChe; Dotum, DotumChe

3. Microsoft Office

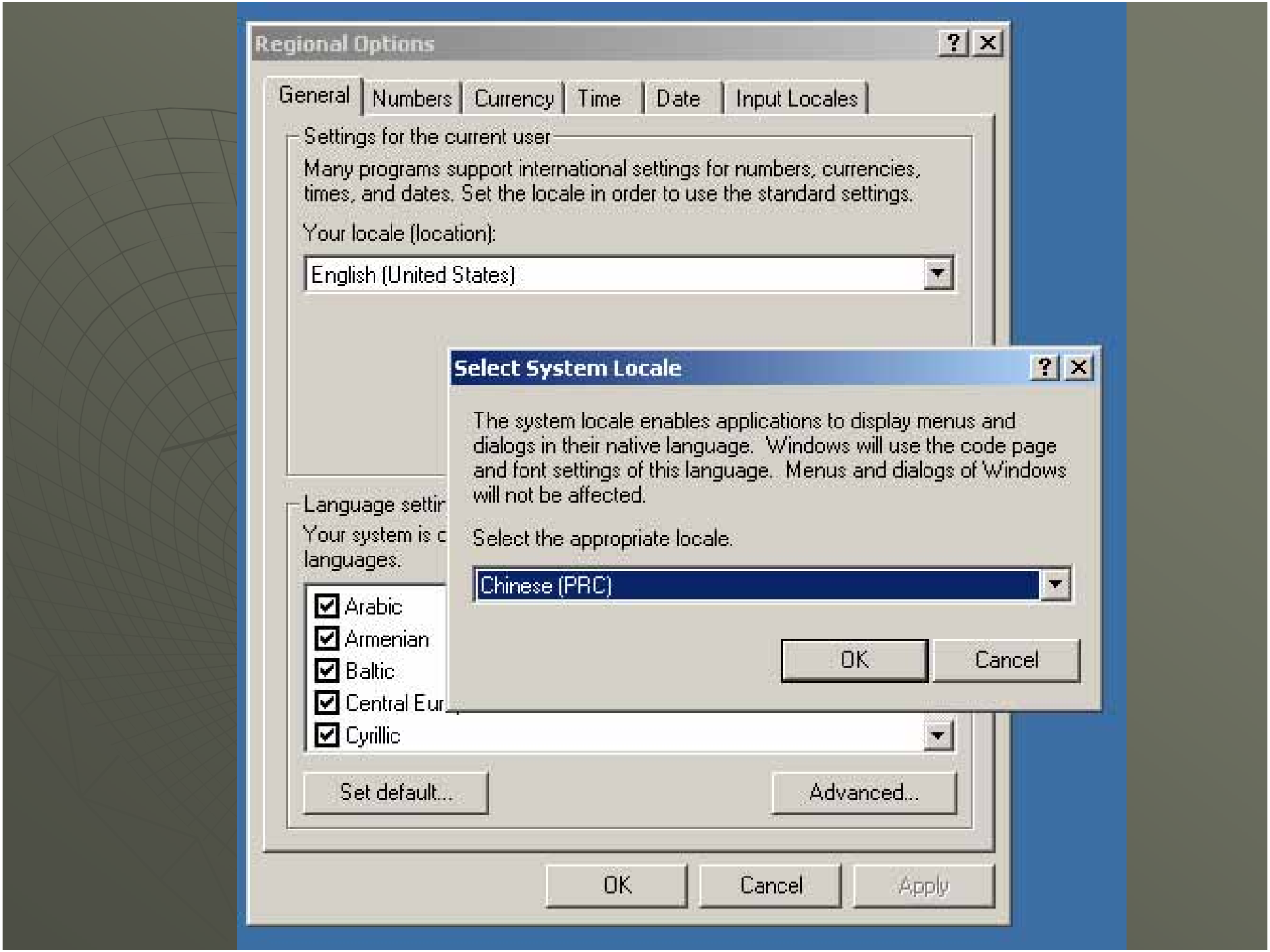
Download global IME

http://office.microsoft.com/assistance/preview.aspx?AssetID=HA010347361033&CTT=6&Origin=EC010553071033

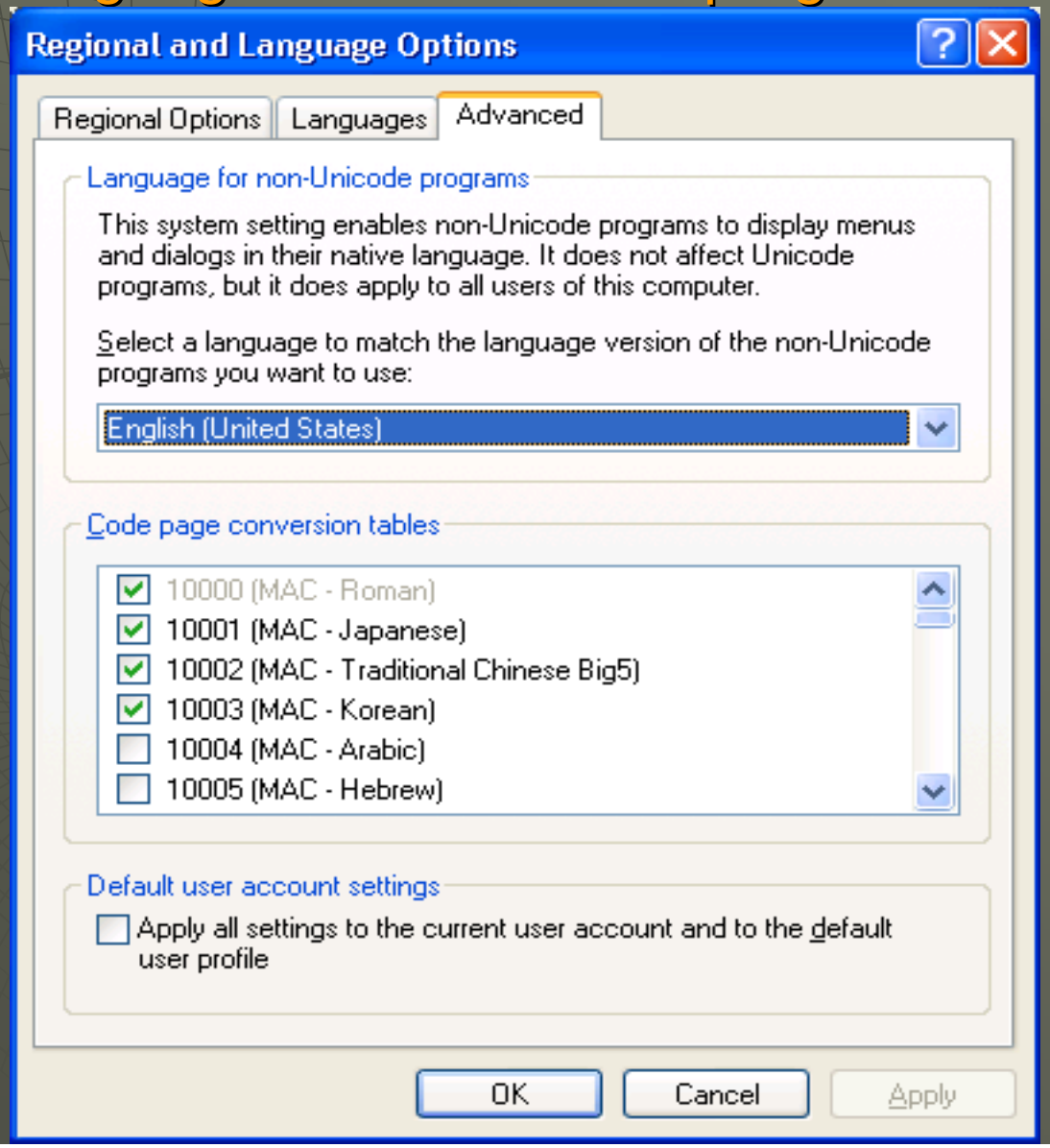
4. Non-Unicode setting

Windows 2000

Start → Settings → Control Panel → Regional Options →
General/ Language Settings for the system


Regional Options
General
Numbers
Currency
Time
Date
Input Locales
Settings for the current user
Many programs support international settings for numbers, currencies, times, and dates. Set the locale in order to use the standard settings.
Your locale (location):
English (United States)
Language settir
Your system is c
languages.
Arabic
Armenian
Baltic
Central Eur
Cyrillic
Set default...
Advanced...
OK
Cancel
Apply
Select System Locale
The system locale enables applications to display menus and dialogs in their native language. Windows will use the code page and font settings of this language. Menus and dialogs of Windows will not be affected.
Select the appropriate locale.
Chinese (PRC)
OK
Cancel

XP: start-control panel-Date, Time, Language, and Regional Options-regional and language options-advanced-language for nonunicode programs



5. PDF

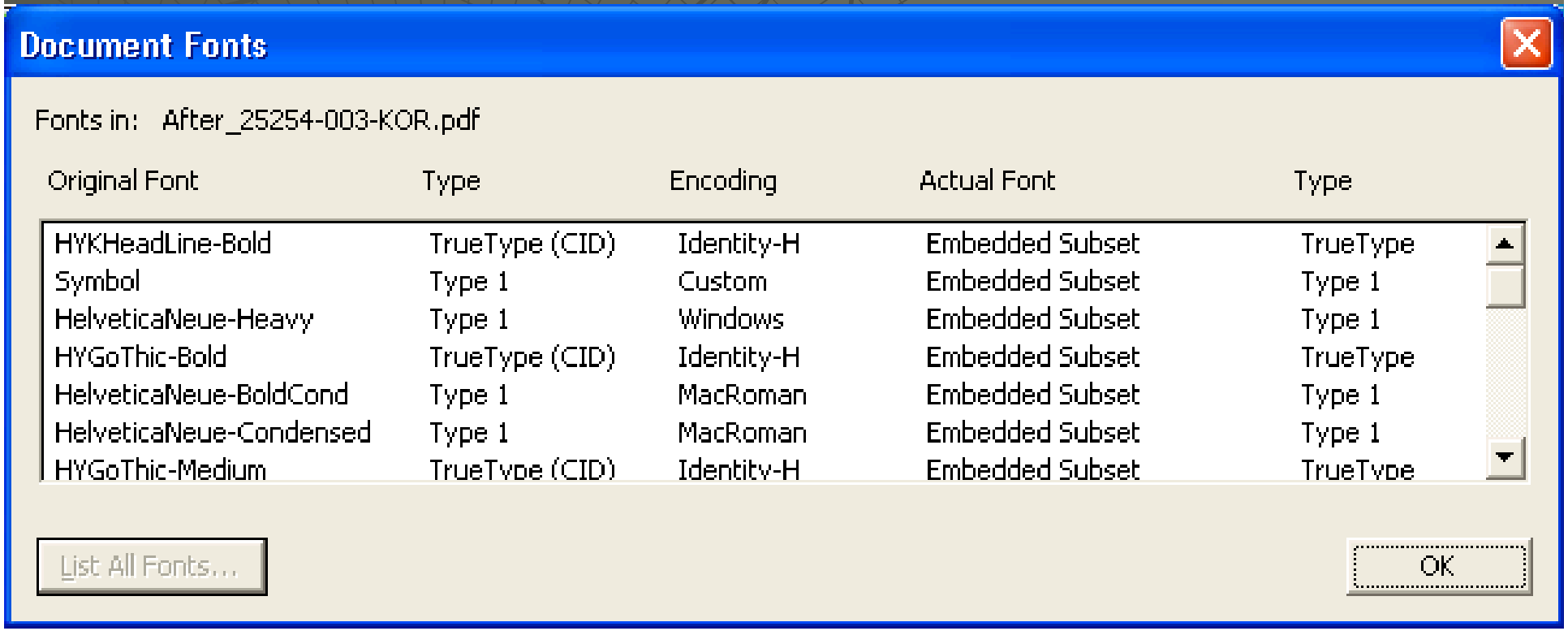
4.0 and above has no problem in dealing with CJK

Asian font packs for Adobe Reader

http://www.adobe.com/products/acrobat/acrrasianfontpack.html

Checking for font embedded

File-document property-fonts



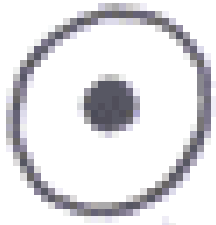
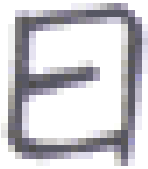
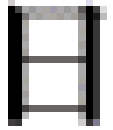
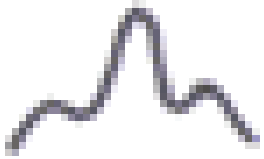
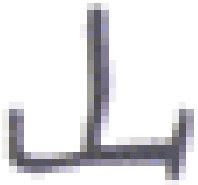
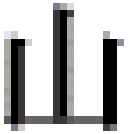
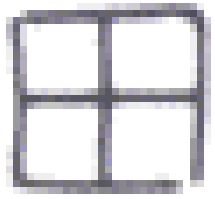
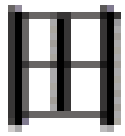
6. Outlined vs. Fonts

- **Font** **outline**

Save from MS Word- place in Illustrator/InDesign- type- create outline

- coding TrueType Vector Image
- encoding Text graphic path)
- size small big
- change easy hard
- printing thin thicker
- font file needed unneeded

ABC of Chinese

	Ancient	Modern	Print
Sun			
Mountain			
Farmland			

- No Alphabet
- Not used only in China or Chinese spoken countries
- No word! Each character has it’s own meaning and sound
- No specific Line break rules (Mostly)

Simplified/Traditional vs Madarin/Cantonese

Target Market	Spoken (Interpreting)	Written (Translation)
Mainland China	Mandarin	Simplified
Taiwan	Mandarin or Taiwanese*	Traditional
Hong Kong	Cantonese**	Traditional
Singapore	Mandarin	Simplified
Chinese Community within USA or other foreign countries	Mandarin/ Cantonese***	Traditional

* Mandarin is the official language in school and broadcasting, but Southern Min--often called “Taiwanese”--is commonly used in conversation. So, if an interpreting client is a businessperson going to Taiwan, he or she will need a Mandarin interpreter, but if your client is an elderly person from Taiwan, you should confirm whether he/she speaks Mandarin or Taiwanese.

** Mandarin is increasingly prevalent, subsequent to Hong Kong’s return to Chinese control.

*** A large percentage of Chinese immigrants came from a region where Cantonese is the spoken dialect. Still, Mandarin is regarded as “standard.” However, one should not send materials for the American market to translators in China, since they will not understand the idioms and common expressions that are used only in the USA. This is particularly true in the context of insurance, social security and education.

Dialects of Chinese

Dialect	Chinese Name	% Population*
Mandarin	普通话	74.8
Wu	吴语	8.0
Yue (Cantonese)	广东话	4.8
Xiang	湖南话	4.0
Southern Min	闽南话	2.7
Northern Min	闽北话	1.0
Hakka	客家话	2.7
Gan	江西话	2.0

- Mandarin:

1　普通话 (Used in China, Referring only to spoken language)

2　国语 (Used in Taiwan. In Taiwan, this term refers not only to the spoken language but also to the written Traditional Chinese language)

3　北京官话 (Old translation that was in use about 100 years ago.)

- Chinese:
- 中文 (The Chinese Language, The language of 中国)
- 汉语 (Language of the Han Nationality. 95% of Chinese are of Han Nationality, then there are 55 more other minority nationality.)

- Since the founding of the PRC in 1949, the Chinese government has been quite active in Chinese character reform. On February 2, 1956, the government published a document called A Scheme for the Simplification of Chinese Characters (汉字简化方案). This scheme consists of three tables. Table 1 lists 352 simplified characters that cannot be used as radicals. Table 2 lists 132 simplified characters that can be used as radicals, plus 14 simplified radicals. Table 3 contains 1,754 characters that were derived using the 132 radical-capable simplified characters and the 14 simplified radicals. The three tables simplified a total number of 2,236 characters.

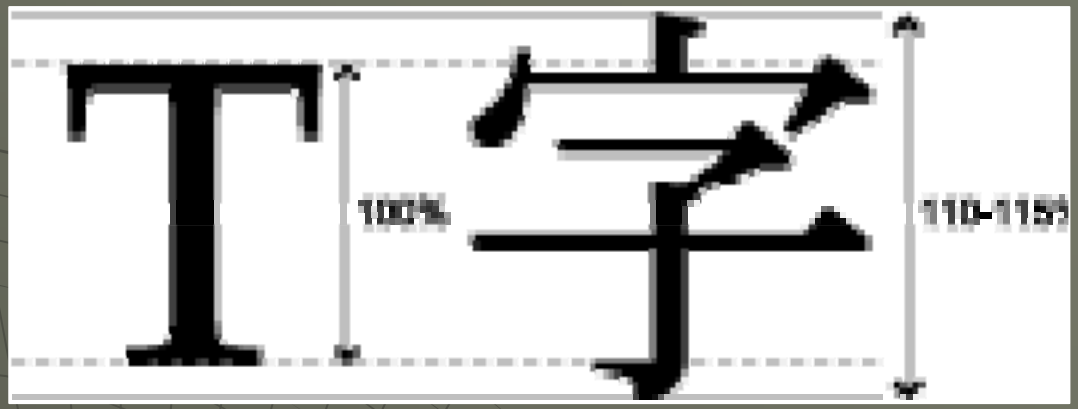
- In 1977, the Chinese government published the Second Scheme for the Simplification of Chinese Characters. Table 1 contained 248 characters. Table 2 included 605 simplified characters and 61 simplified radicals. However, since the simplification was so extreme, it met with strong resistance from the society. On June 24, 1986, that second scheme was rescinded. Later in 1986, the first scheme was republished with a few words in the tables adjusted. As a result, the total number of simplified characters now stands at 2,235.

WHERE THE MAIN CHINESE 'DIALECTS' ARE SPOKEN
Cantonese
Gan
Hakka
Xiang
Mandarin
Min Bei
Min Nan
Wu
RUSSIA
Lake Baykal
MONGOLIA
INNER MONGOLIA
XINJIANG
HEILONGJIANG
JILIN
LIAONING
NORTH KOREA
SOUTH KOREA
Beijing
TIANJIN
HEBEI
SHANXI
SHANDONG
NINGXIA
CHINA
QINGHAI
GANSU
SHAANXI
Yellow River
HENAN
ANHUI
JIANGSU
YELLOW SEA
TIBET
SICHUAN
HUBEI
ZHEJIANG
JIANGXI
EAST CHINA SEA
Yangtze
HUNAN
FUJIAN
NEPAL
BHUTAN
GUIZHOU
TAIWAN
INDIA
BANGLA-DESH
YUNAN
GUANGXI
GUANGDONG
Hong Kong
Macau
MYANMAR
VIETNAM
PHILIPPINES
BAY OF BENGAL
LAOS
HAINAN
THAILAND

- **2 Byte and L10N**
  - Encoding Support-2byte enabling
  - Space in memory, windows: One Chinese character = 2
- **In the computer realm, Simplified Chinese uses GB2312 encoding and Traditional Chinese uses the so-called Big5 encoding.**

Pricing

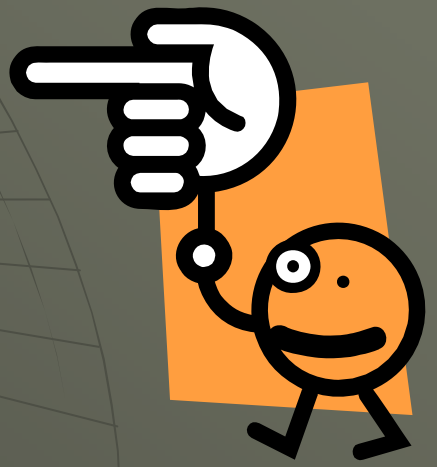
- **Source/Target *:***
  - **In China, all translation prices are by Chinese characters, not English words.**
  - **They count not only Chinese characters but also numbers, English characters and punctuations**
- **Golden number**
  - **1000 chinese characters are roughly 1200 to 1500 English words**
- **Hard copy word counting**
- **Changes in DTP**
  - ***Less spaces***
  - ***The fonts are always look bigger***



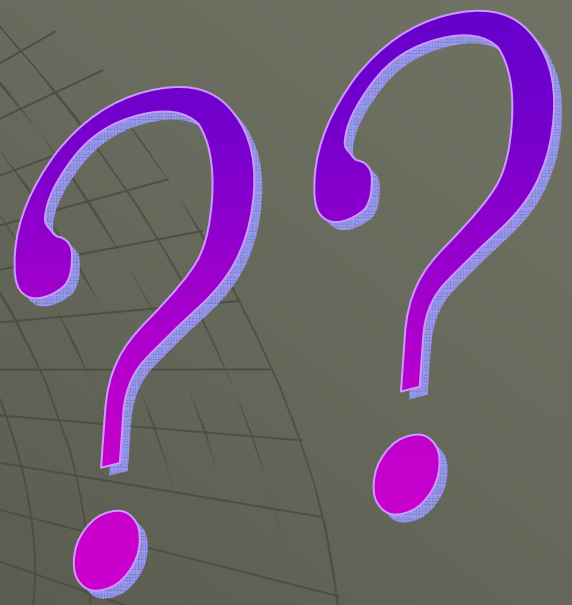
- **Formatting rates**
  - **MS-Word**
  - **Tables**
  - **PPT**
  - **FM/PM/Quark/Indesign**

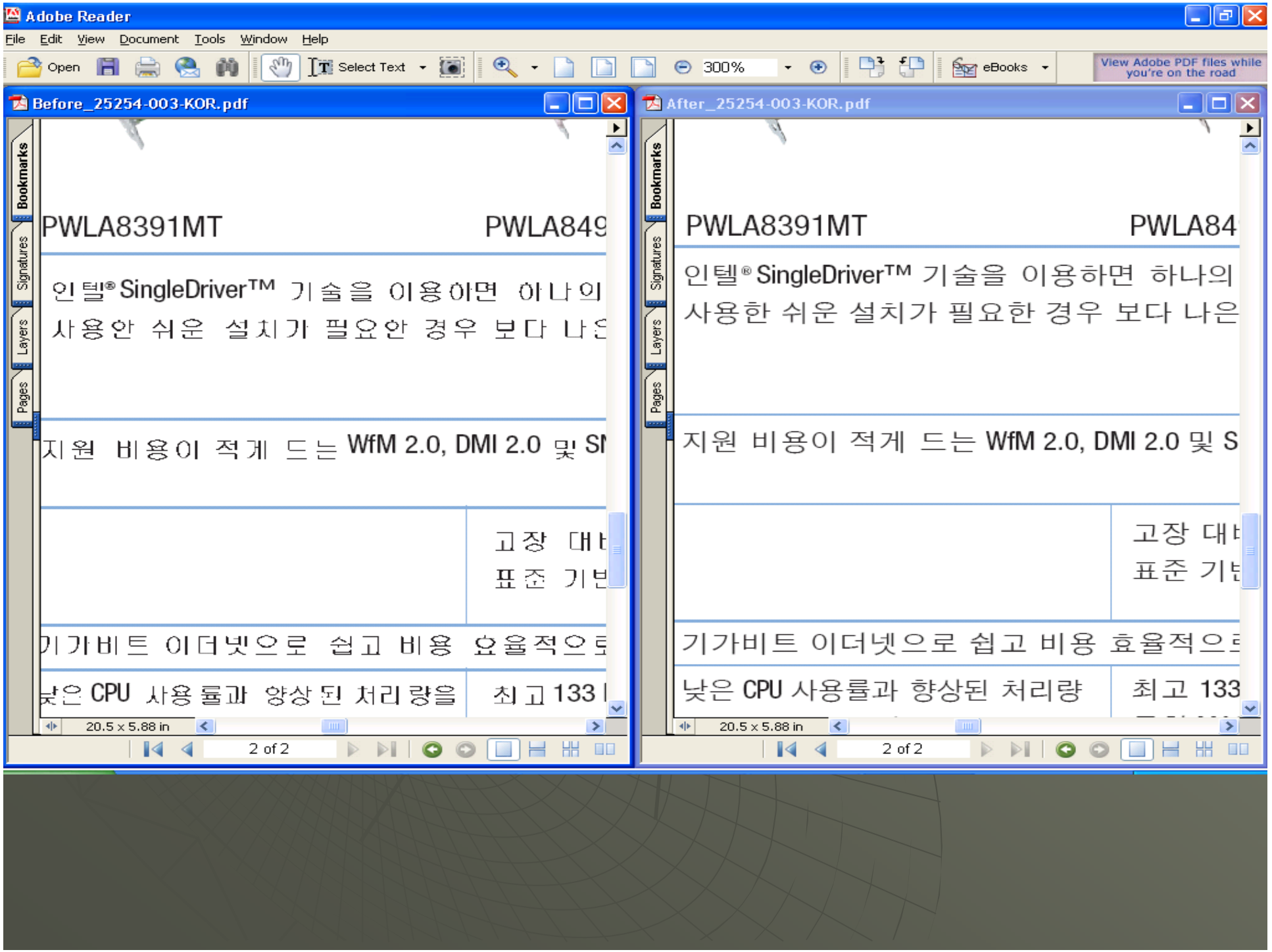
Process and details for Chinese (and other Asian languages)

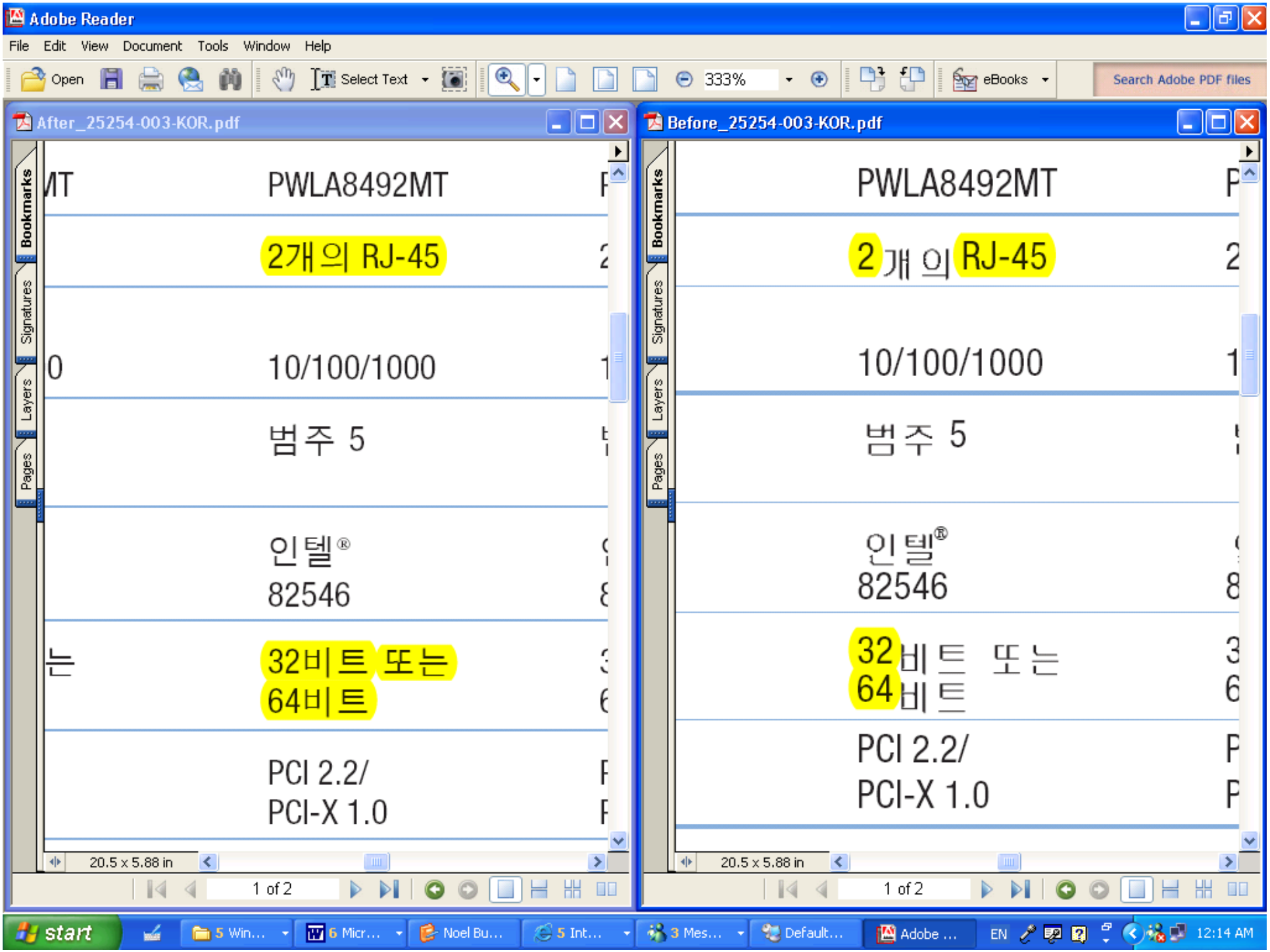
- DTP

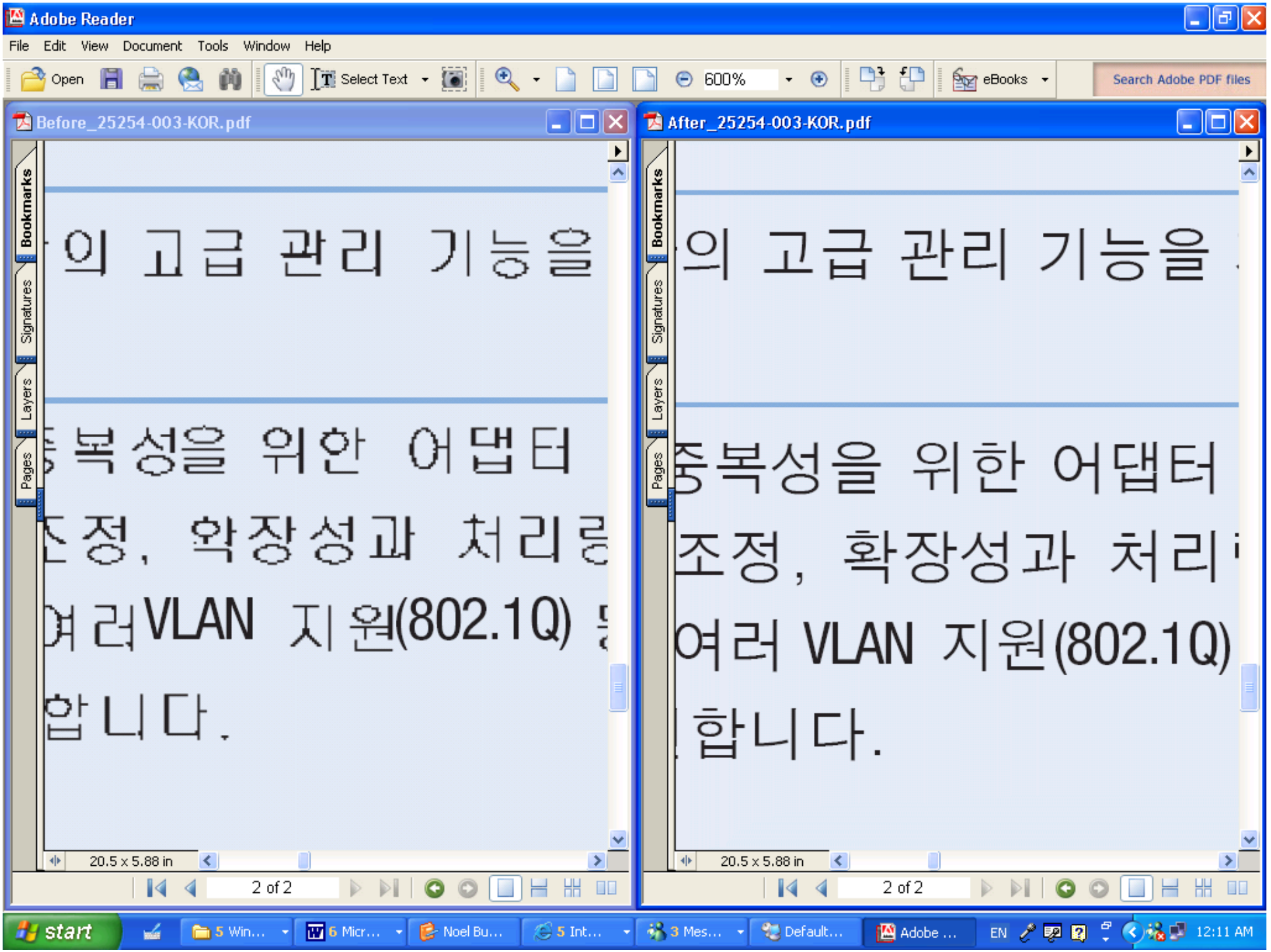


- QUIZ 2- finding differences




Adobe Reader
File Edit View Document Tools Window Help
Open
Select Text
300%
eBooks
View Adobe PDF files while you're on the road
Before_25254-003-KOR.pdf
Bookmarks
Signatures
Layers
Pages
PWLA8391MT
PWLA849
인텔® SingleDriver™ 기술을 이용하면 하나의
사용한 쉬운 설치가 필요한 경우 보다 나은
지원 비용이 적게 드는 WfM 2.0, DMI 2.0 및 SN
고장 대비
표준 기반
기가비트 이더넷으로 쉽고 비용 효율적으로
낮은 CPU 사용률과 향상된 처리량을
최고 133
20.5 x 5.88 in
2 of 2
After_25254-003-KOR.pdf
Bookmarks
Signatures
Layers
Pages
PWLA8391MT
PWLA84
인텔® SingleDriver™ 기술을 이용하면 하나의
사용한 쉬운 설치가 필요한 경우 보다 나은
지원 비용이 적게 드는 WfM 2.0, DMI 2.0 및 S
고장 대비
표준 기반
기가비트 이더넷으로 쉽고 비용 효율적으로
낮은 CPU 사용률과 향상된 처리량
최고 133
20.5 x 5.88 in
2 of 2


Adobe Reader
File Edit View Document Tools Window Help
Open
Select Text
333%
eBooks
Search Adobe PDF files
After_25254-003-KOR.pdf
Bookmarks
Signatures
Layers
Pages
MT
PWLA8492MT
2개의 RJ-45
0
10/100/1000
범주 5
인텔®
82546
는
32비트 또는
64비트
PCI 2.2/
PCI-X 1.0
20.5 x 5.88 in
1 of 2
Before_25254-003-KOR.pdf
Bookmarks
Signatures
Layers
Pages
PWLA8492MT
2개의 RJ-45
10/100/1000
범주 5
인텔®
82546
32비트 또는
64비트
PCI 2.2/
PCI-X 1.0
20.5 x 5.88 in
1 of 2
start
5 Win...
6 Micr...
Noel Bu...
5 Int...
3 Mes...
Default...
Adobe ...
EN
12:14 AM


Adobe Reader
File Edit View Document Tools Window Help
Open
Select Text
600%
eBooks
Search Adobe PDF files
Before_25254-003-KOR.pdf
의 고급 관리 기능을
복성을 위한 어댑터
정, 확장성과 처리
여러VLAN 지원(802.1Q)
합니다.
20.5 x 5.88 in
2 of 2
After_25254-003-KOR.pdf
의 고급 관리 기능을
중복성을 위한 어댑터
조정, 확장성과 처리
여러 VLAN 지원(802.1Q)
합니다.
20.5 x 5.88 in
2 of 2
Bookmarks
Signatures
Layers
Pages
start
5 Win...
6 Micr...
Noel Bu...
5 Int...
3 Mes...
Default...
Adobe ...
EN
12:11 AM

PDF

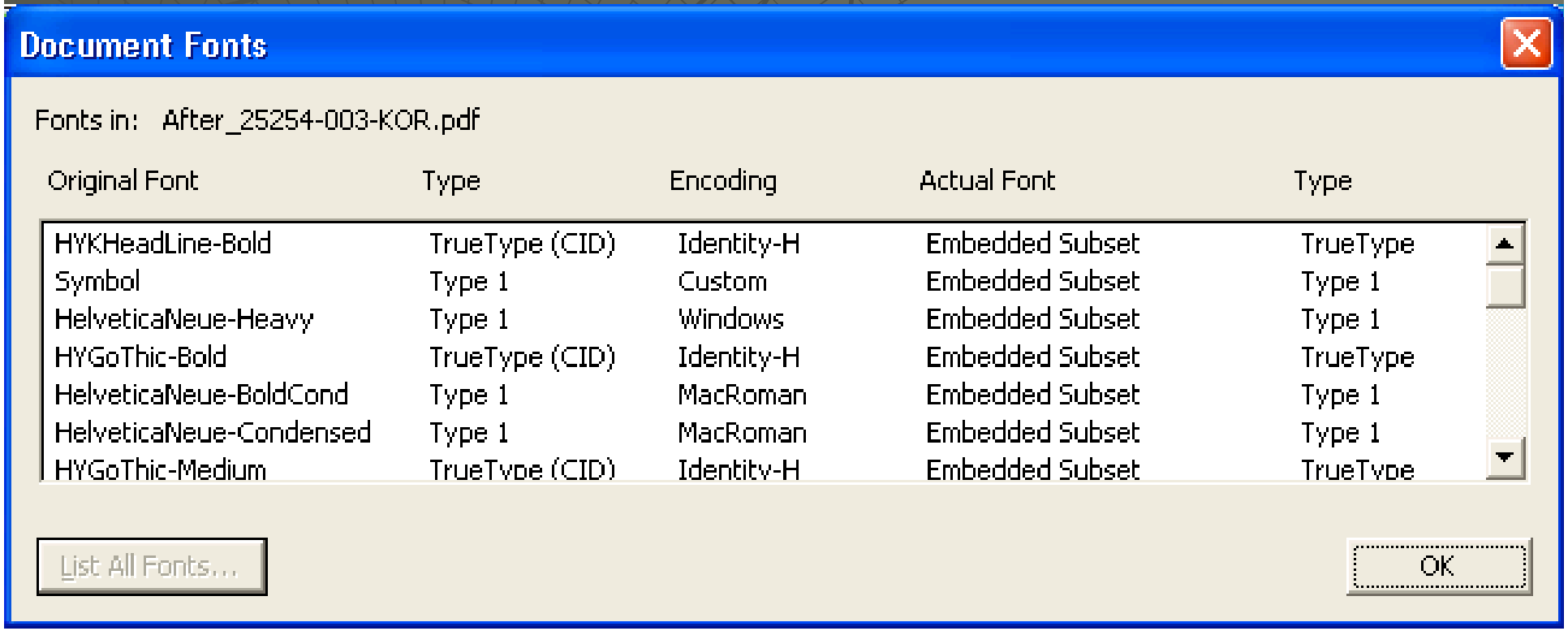
4.0 and above has no problem in dealing with CJK

Asian font packs for Adobe Reader

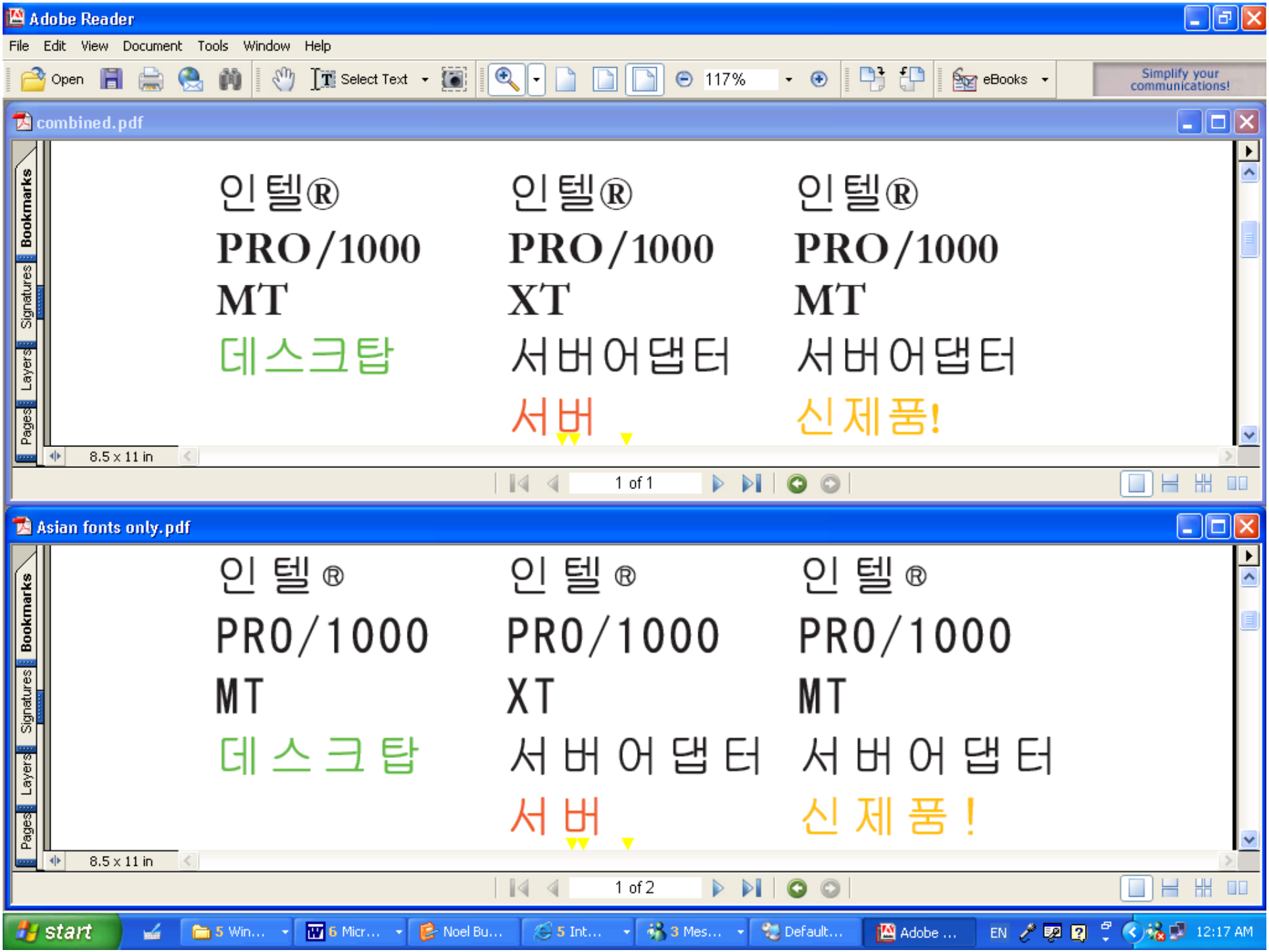
http://www.adobe.com/products/acrobat/acrrasianfontpack.html

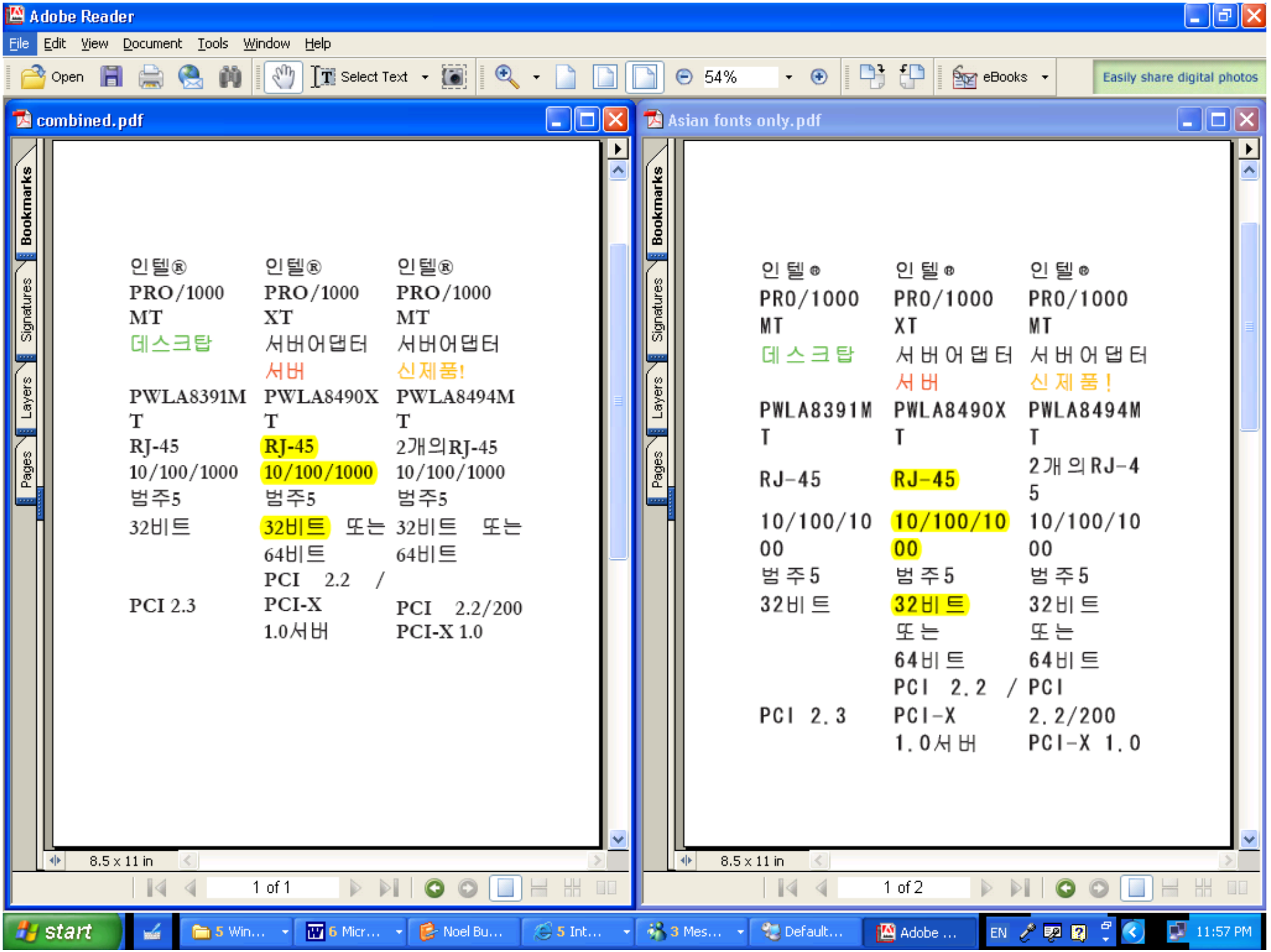
Checking for font embedded

File-document property-fonts



- QUIZ 3


Adobe Reader
File Edit View Document Tools Window Help
Open
Select Text
117%
eBooks
Simplify your communications!
combined.pdf
Bookmarks
Signatures
Layers
Pages
인텔®
PRO/1000
MT
데스크탑
인텔®
PRO/1000
XT
서버어댑터
서버
인텔®
PRO/1000
MT
서버어댑터
신제품!
8.5 x 11 in
1 of 1
Asian fonts only.pdf
Bookmarks
Signatures
Layers
Pages
인 텔 ®
PRO/1000
MT
데 스 크 탑
인 텔 ®
PRO/1000
XT
서 버 어 댑 터
서 버
인 텔 ®
PRO/1000
MT
서 버 어 댑 터
신 제 품 !
8.5 x 11 in
1 of 2
start
5 Win...
6 Micr...
Noel Bu...
5 Int...
3 Mes...
Default...
Adobe ...
EN
12:17 AM


Adobe Reader
File Edit View Document Tools Window Help
Open
Select Text
54%
eBooks
Easily share digital photos
combined.pdf
Bookmarks
Signatures
Layers
Pages
인텔®
PRO/1000
MT
데스크탑
PWLA8391M
T
RJ-45
10/100/1000
범주5
32비트
PCI 2.3
인텔®
PRO/1000
XT
서버어댑터
서버
PWLA8490X
T
RJ-45
10/100/1000
범주5
32비트 또는
64비트
PCI 2.2 /
PCI-X
1.0서버
인텔®
PRO/1000
MT
서버어댑터
신제품!
PWLA8494M
T
2개의RJ-45
10/100/1000
범주5
32비트 또는
64비트
PCI 2.2/200
PCI-X 1.0
8.5 x 11 in
1 of 1
Asian fonts only.pdf
Bookmarks
Signatures
Layers
Pages
인텔®
PRO/1000
MT
데스크탑
PWLA8391M
T
RJ-45
10/100/10
00
범주5
32비트
PCI 2.3
인텔®
PRO/1000
XT
서버어댑터
서버
PWLA8490X
T
RJ-45
10/100/10
00
범주5
32비트
또는
64비트
PCI 2.2 /
PCI-X
1.0서버
인텔®
PRO/1000
MT
서버어댑터
신제품!
PWLA8494M
T
2개의RJ-4
5
10/100/10
00
범주5
32비트
또는
64비트
PCI
2.2/200
PCI-X 1.0
8.5 x 11 in
1 of 2
start
5 Win...
6 Micr...
Noel Bu...
5 Int...
3 Mes...
Default...
Adobe ...
EN
11:57 PM

4 points to check Asian DTP Quality

- Baseline
- Font embedded or bitmap?
- Justified Alignment
- Used original English characters for English Alphabet and numbers

Other things remember to Quote CJK DTP jobs

- Need English Fonts and graphics
- Space-Mostly smaller
- Get translation approved
- Charge for Changes
- Final Target format

Quark Process : Outlined eps vs PDF

Western Languages

- Quark >> (Extract translatable text with Trados> RTF> (Translator)>automatic Import with Trados> Quark>Save PDF>> Done!!

Asian Languages

- Quark>Extract Text box per box>Combine each text files into one> Edit for translator to understand which goes where> Doc file>(Translator)> Manually copy and paste each text line> Quark> Save each page in EPS > Delete Every back ground and graphic from EPS files> Outline each 'Text only eps files'> Delete every Text from Quark files> Combined these Quark and 'outlined text only EPS files' together>>Now create PDF .

- *And for the last, each time client change one thing, we have to go through the red part again.*

- Word /Interpreting
  - Proper Name
    - E>C
    - C>E
      - *CT and CS difference in reading names*
- Business Card
  - Front/Back
  - Logo file and color printing
  - No address, e-mail
  - Color settings

More to read

- ***CJKV Information Processing***
  Chinese, Japanese, Korean & Vietnamese Computing
- By Ken Lunde
  1st Edition December 1998
  ISBN: 1-56592-224-7
  1125 pages, $69.95 US, $108.95 CA