<table>
<thead>
<tr>
<th>項目</th>
<th>内容</th>
</tr>
</thead>
<tbody>
<tr>
<td>項目</td>
<td>Adobe-Japan1-6とMJ文字図形名の対応</td>
</tr>
<tr>
<td>作者</td>
<td>安岡 孝一</td>
</tr>
<tr>
<td>項目</td>
<td>東洋学へのコンピュータ利用研究セミナー &quot;情報技術&quot;の役割</td>
</tr>
</tbody>
</table>
Adobe-Japan1-6 と MJ 文字図形名の対応
安岡孝一 *

1 はじめに
文字情報基盤促進協議会 (旧 IVS 技術促進協議会) のカラミで、Adobe-Japan1-6 と MJ 文字図形名の対応表を、作成せざるを得なくなった。ぶちまけて言えば、筆者個人としては Adobe-Japan1 の方が好き † だし、IVS (Ideographic Variation Sequence) についても、Adobe-Japan1 の IVS を好んで † 使用してきた。しかし、Hanyo-Denshi の IVS 追加 § に続いて、Moji_Joho の IVS 追加 ¶ もおこなわれ、そろそろ分が悪くなってきたようである。
文字情報基盤促進協議会としては、Adobe-Japan1 の IVS と、Moji_Joho の IVS との対応表が欲しいところのだが、これは直直かど難しい。Moji_Joho の IVS は、Hanyo-Denshi の IVS において一旦「包摂」されたはずの字体群を、一部「再分離」する形で構成しているため、どの字体が許容されるか、かなりわかりにくい。一方、Adobe-Japan1 の IVS が、どの程度の「デザイン差」を許しているのかについては、そもそも資料がない。
これらを要案した結果、筆者は、とりあえずの叩き台として、Adobe-Japan1-6 のオリジナル規格書 ‡ に用いられた小塚明朝 KozMinProVI-Regular の漢字と、MJ 文字情報基盤の IPAmj 明朝 Ver.004.01 の漢字を、字形レベルで比較することにした。以下に詳細を述べる。

2 小塚明朝と IPAmj 明朝の対応
小塚明朝 KozMinProVI-Regular の漢字 14667 字 ** と、IPAmj 明朝 Ver.004.01 の漢字 58861 字 †† を、筆者の目 (と鼻) で比較した。また、この比較データを、Adobe の Ken Lunde、情報処理推進機構 (IPA) の武藤主祐が作成した比較データと突合し、さらに精度を上げたところで、13856 組に関しては、小塚明朝と IPAmj 明朝で 1:1 対応が可能だという感触を得た。
これら 13856 組と、小塚明朝の残り 811 字について、次ページ以降に対応表を示す。小塚明朝には CID を、IPAmj 明朝には MJ を付した。例を挙げると、CID+1200 と MJ006294 は 1:1 対応可能だが、CID+14302 と MJ006369 は 1:1 対応とするには微妙な抵抗があり、CID+6480 は MJ006395 に近いが MJ006396 も気になる、ということを示している。

* 京都大学人文科学研究所附属東アジア人文学研究センター
† 安岡孝一：Adobe-Japan1-6 と Unicode — 異体字処理と文字コードの現実, 情報管理, Vol.48, No.8 (2005 年 11 月), pp.487-495。
‡ 安岡孝一：漢字 1 文字が最大 8 バイト、Unicode の「IVS」とは?, 日経 ITpro (2010 年 1 月 29 日)。
§ 安岡孝一：Unicode の IVS がもたらすメリットとデメリット, 日経 ITpro (2011 年 1 月 27 日)。
¶ 安岡孝一：行政情報処理用漢字コードの現状, 日経 ITpro (2014 年 6 月 30 日～7 月 4 日)。
Adobe-Japan1-6 Character Collection for CID-Keyed Fonts, Adobe Technical Note #5078 (2004 年 6 月 11 日)。
** CID+657 「々」 CID+658 「〆」 CID+12869 「注」 を含む。
†† AJ20300 「満」 など MJ 文字図形でない漢字を除く。
<p>| 你 | 你 | 你 | 你 |
| 位 | 位 | 位 | 位 |
| 住 | 住 | 住 | 住 |
| 住 | 住 | 住 | 住 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |
| 你 | 你 | 你 | 你 |</p>
<table>
<thead>
<tr>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
<th>磁</th>
</tr>
</thead>
<tbody>
<tr>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
<td>磁</td>
</tr>
</tbody>
</table>
3 考察

小塚明正 KozMinProVI-Regular の漢字 14667 字のうち、IPAmj 明朝 Ver.004.01 と 1:1 対応にならない 811 字に関して、どのような手当てが可能か。この章では、いくつかの典型例に対して、考察を試みる。

3.1 「次」を部分字体とする漢字

<table>
<thead>
<tr>
<th>次次</th>
<th>次</th>
<th>次次姿姿</th>
<th>茨茨</th>
<th>資資資資</th>
</tr>
</thead>
<tbody>
<tr>
<td>CID+13800 MJ014749</td>
<td>CID+2253</td>
<td>CID+13799 MJ014748</td>
<td>CID+1205 MJ021800</td>
<td>CID+13798 MJ025947</td>
</tr>
</tbody>
</table>

「次」を部分字体とする漢字は、小塚明朝では CID+2253「次」CID+2207「姿」CID+1205「茨」CID+2238「諮」CID+2239「資」の 5 字がある。IPAmj 明朝では MJ013852「次」MJ021800「茨」MJ032763「諮」MJ059139「資」の 4 字が発見されている。これらのうち、1:1 対応するのは「茨」だけであり、残りはパラパラである。ただ、IPAmj 明朝の MJ013852 は、元々は住基統一文字 6825 に由来しており、その意味では「次」に作るべきである。また、IPAmj 明朝の MJ059139 は、元々は戸籍統一文字 524500 に由来しており、やはり「次」に作るべきである。将来、もし MJ013852 や MJ059139 の字形を、修正（あるいは別字形追加）することがあるなら、その際に「次」「姿」「諮」「資」の 4 字も、IPAmj 明朝に追加すべきだろう。

3.2 「魚」を偏に持つ漢字
「亀」を偏に持つ漢字は、IPAmj 明朝には 20 字以上あるものの、小塚明朝には「鰤」「鮮」「鱒」の 3 字しかない。ならば、これら 3 字と CID+13178「亀」を、IPAmj 明朝に追加できれば、とりあえずは万々歳である。CID+20250・CID+20151・CID+20255は、戸籍統一文字にも住基統一文字にも含まれていないが、登記統一文字には01100160・01100170・01100510として含まれているようであり、あとは CID+13718を、どういうロジックでIPAmj明朝に追加するかだろう。

### 3.3 「兆」を部分字体とする漢字

「兆」を部分字体とする漢字のうち、左下画が突き抜けているのは、小塚明朝ではCID+13477「兆」CID+13530「姚」CID+14614「眺」CID+13484「桃」CID+13478「眺」CID+13573「竜」CID+13480「眺」CID+13485「逃」の8字がある。これに対し、IPAmj明朝ではMJ007218「兆」MJ009651「姚」MJ019241「竜」MJ025315「眺」の4字だけであり、残る4字は1:1対応しない。できれば、残る4字「眺」「桃」「眺」「逃」も、IPAmj明朝に追加したいところである。

### 3.4 「非」を部分字体とする漢字
「非」を部分字体とする漢字のうち、左下画が突き抜けているのは、小塚明朝では25字、IPAmj明朝では23字ある。これらの多くは1:1対応可能だが、CID+17320「劁」CID+20308「嚊」CID+14785「排」CID+7430「廓」の4字は、微妙な問題が残る。これら4字も、IPAmj明朝に追加したいところである。

3.5 「直」を部分字体とする漢字

CID+13934「直」CID+13464「埴」CID+13843「塩」CID+13845「植」CID+13846「殖」CID+13920「置」の7字については、できればこままの形でIPAmj明朝に追加したい。ただ、問題になるのがCID+13464とCID+13843で、これらを2字としてIPAmj明朝に追加するか、それとも1字にまとめて追加するか、判断が難しい。その結論によって、CID+13465やCID+13572の処遇も決定する方がいいだろう。

3.6 「耳」を部分字体とする漢字

CID+14382「咡」CID+13527「囁」CID+13541「揖」CID+4452「咡」CID+4928「囁」CID+2699「揖」CID-313
「耳」の下の画は、突き抜けない場合、突き抜ける場合、突き抜けてウロコのある場合、の大きさ 3 つに分かれる。この点に関して、小塚明潮と IPAmj 明朝は、それぞれかなり混乱しており、全貌を直感的に把握するのは難しい。たとえば「椕」は、戸籍統一文字 140810 も、住基統一文字 6442 も、ほぼ CID+2689「椕」の字形である。なぜ IPAmj 明朝では、これらを MJ012666「椕」に作っている†なのか、筆者には全く理解できない。あるいは「束」は、戸籍統一文字 196590 も、住基統一文字 6D31 も、CID+16932「束」の字形であり、なぜ IPAmj 明朝では、これらを MJ015309「束」に作っているのか理解できない。CID+22465「束」は、戸籍統一文字 386480 も、住基統一文字 8848 も、この字形であり、MJ023856「束」は理解できない。

とりあえずは、CID+14382「啣」CID+13527「嘔」CID+13541「慉」CID+2689「椕」CID+21667「齮」CID+13425「敢」CID+13551「椕」CID+14671「椕」CID+16932「束」CID+21880「満」CID+17995「摂」CID+21919「瀞」CID+15016「聾」CID+13580「聚」CID+13581「鍞」CID+13473「聾」CID+13517「聾」CID+13583「聾」CID+15401「聾」CID+13582「聾」CID+13518「聾」CID+13498「聾」CID+22465「聾」CID+22499「聾」CID+22533「聾」CID+2366「聾」CID+13456「聾」CID+22908「駭」CID+19108「駭」の 29 字を、全て IPAmj 明朝に収録してもらう、という形で治めるしかないだろうと思う。

3.7 「西」を部分字体とする漢字
「醤」を部分字体とする漢字の中でキーとなるのは、CID+13810「酌」と MJ026437「酌」だと考えられる。この2字を同一視できるなら、他も同一視可能だろう。一方、この2字を同一視できないなら、CID+13529「醇」CID+13901「尊」CID+19573「醯」CID+13590「酗」CID+13810「酌」CID+13610「釀」CID+13611「釀」CID+13612「酢」CID+13613「酢」CID+13615「醋」CID+13617「釀」CID+13616「釀」CID+13618「醯」CID+13619「釀」CID+6964「釀」の15字を、何らかの形でIPA mj明朝に引き取ってもらう必要があると考えられる。
3.8 「死」を部分字体とする漢字

<table>
<thead>
<tr>
<th>漢字</th>
<th>笔画数</th>
<th>笔画方向</th>
<th>笔画形状</th>
<th>笔画連続</th>
</tr>
</thead>
<tbody>
<tr>
<td>慨</td>
<td>5</td>
<td>左上右下</td>
<td>曲線</td>
<td>連続</td>
</tr>
<tr>
<td>慨</td>
<td>5</td>
<td>左上右下</td>
<td>曲線</td>
<td>連続</td>
</tr>
<tr>
<td>慨</td>
<td>5</td>
<td>左上右下</td>
<td>曲線</td>
<td>連続</td>
</tr>
</tbody>
</table>

「死」を部分字体とする漢字は、字形差が非常に多様なため、かなり混乱している。
筆者としては、CID+7964「歿」CID+13647「殤」CID+19255「嘅」CID+13678「慨」

- 317 -
3.9 「邊」と「邊」
IVSにおいて最大のグループをなす「邊」と「邊」は、小塚明州とIPA mj明朝の間においても、非常に対応が難しい。わずか11組が1:1対応可能であり、残りはバラバラな字形が散らばっている。筆者としては、CID+13407「邊」CID+14246「邊」CID+14250「邊」CID+14245「邊」CID+14251「邊」CID+14252「邊」CID+14243「邊」CID+14247「邊」CID+14235「邊」CID+14240「邊」CID+14239「邊」の11字を、IPA mj明朝に追加したいところだが、混乱に拍車をかける気がして、多少、躊躇せざるを得ない。

3.10 起筆の有無

「文」と「文」の字体差に代表される起筆の有無は、IPA mj明朝において、かなり大きな混乱を生んでいる部分である。汎用電子情報交換整備プログラム†において制作された平成明朝では、起筆を全て無くす方針だったのに対し、それを踏襲したはずの文字情報基盤整備事業において制作された IPA mj明朝では、起筆を一部「復活」しているからである。以下では、起筆の有無に関して、小塚明州とIPA mj明朝との間で1:1対応にならないものを、ざっと見ていくことにしよう。

3.10.1 「文」の起筆

†汎用電子情報交換整備プログラム成果報告書別冊、日本規格協会（2009 年3月1日）
3.10.2 えんにょうの起筆

3.10.3 「更」の起筆
3.10.4 「刃」の起筆

3.10.5 「交」と「父」の起筆
3.10.6 その他の起筆
ここまで起筆の有無を見てきた限りでは、IPA mj 明朝の制作過程において、一部の字形において起筆を戻し忘れた（あるいは戻すのを怠った）、というのが実情のようである。たとえば、MJ044449「絵」は、JIS X 0213 の 2-84-15 由来であり、JIS X 0213 規格票の平成明朝では起筆がある。そこから住基統一文字 AB50 に入ったが、やはり起筆がある。ところが、汎用電子情報交換プログラムにおいて制作された（新たな）平成明朝では、わずかに起筆を無くしている。それが IPA mj 明朝に流れ込み、起筆を無くしたままになっているのである。CID+18339「絵」の方が住基統一文字 AB50 に近いのであり、これを IPA mj 明朝に追加する方向で考えるべきだろう。

3.11 縦画のハネの有無
縦画のハネの有無に関して、問題のありそうな字形をざっと一覧にしてみたが、筆者としては、これらの CID を全て IPAmj 明朝に追加するべきか、多少、迷いがある。ただ、追加するとしても 23 字ほどなので、エイヤっと追加してしまってもいいだろう。

3.12 トメとハライ
小塚明朝も IPAmj 明朝も、トメとハライの違いを意識していないようであり、かなりの数の字形差が見られる。ただ、これらの CID を IPAmj 明朝に追加すると、運用上の不都合が想定されることから、ここは無理をせず、1:1 対応に準じて処理する方が良いように思える。

3.13 Adobe-Japan1 のみにある UCS
上に示した48字に関しては、字形レベルのみならず、UCSレベルにおいても、IPAmj明朝とは対応がおこなえない。つまり、これらに関しては、それこそ目をつぶってIPAmj明朝に追加するしかない、ということである。その際には、CID+19777「讃」に関しては、上に示したようにU+400Cに追加する方が、より安全だと思う。

3.14 Adobe-Japan1のIVSに難がある漢字

CID+20254「鰤」は、Adobe-Japan1のIVSでは<U+9C24 U+E0101>に紐づけられているものの、どう考えても別のUCSに配置しなおすべきである。CID+14146「槻」CID+20253「鰤」CID+13697「毘」も同様である。CID+14106「剣」CID+13763「五」CID+20174「畋」については議論のあるところだが、これら7字をIPAmj明朝に追加する場合には、UCSをどうするか再考する必要があるだろう。
4 おわりに

小塚明朝KozMinProVI-Regularの漢字14667字と、IPAmj明朝Ver.004.01の漢字58861字を比較した。この結果、13856組に関しては、小塚明朝とIPAmj明朝で1:1対応が可能だという感覚を得た。小塚明朝の残り811字に対しては、再度チェックの上でIPAmj明朝に追加するか、あるいは追加せずに「対応」するIPAmj明朝を示すか、いずれかの方策が必要だろう。

ただし、このような形でIPAmj明朝に追加する漢字においては、できればAdobe-Japan1のIVSを使い続けたい。現在、Moji_JohoのIVSは、大部分がHanyo-DenshiのIVSと重なっているが、それをAdobe-Japan1のIVSと重ねるわけである。たとえば、CID+2689「摂」をIPAmj明朝に追加する場合、Moji_JohoのIVSは、以下のようにすることが考えられる。

6442

E0102  Moji_Joho  MJ012666
E0103  Moji_Joho  MJ012667
E0100  Moji_Joho  MJxxxxxx

このような方策が可能かどうかは、情報処理推進機構と文字情報基盤促進協議会の「協議」次第だろう。追加すると言っても、高々811字程度のことだし、現実には、もっと少なくなると考えられる。日本の漢字情報処理の発展のため、情報処理推進機構と文字情報基盤促進協議会には、ぜひ頑張ってほしい。