

2009-04-06

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation internationale de normalisation
Международная организация по стандартизации

Doc Type: Working Group Document**Title: Towards an encoding of symbol characters used as emoji****Source: Irish and German National Bodies****Status: National Body Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2009-04-06**

Documents N3582 “Proposal for Encoding Emoji Symbols” and N3583 “Emoji Symbols Proposed for New Encoding” from the Unicode Consortium propose to add 674 characters whose ultimate source is three separate coded character sets used by three mobile telephone carriers in Japan: NTT DoCoMo, KDDI, and Softbank. We recognize the interoperability concerns raised by the Unicode Consortium, and understand that this is the primary concern they have in terms of adding additional characters to the UCS.

We have also been aware that there have been many critics of the proposal to encode “emoji symbols” at all, due to the somewhat whimsical nature of a subset of the characters being proposed. Ireland and Germany take the view that the encoding of symbol sets in the UCS is legitimate and not to be avoided. We do, however, also take the view that symbol sets should be encoded in terms of the semantics of the characters. In the case of the characters proposed in N3583, we have identified certain gaps which should be filled for the encoded characters to “make sense” in the UCS.

It is our view that these characters are not “emoji symbols” despite their origin in coded character sets used in Japanese telephony. “Emoji” is a usage environment, not a semantic domain, and once these characters are encoded in the UCS they are nothing more than “symbols” available for use in any environment.

In this document we propose a number of modifications to the proposal. These modifications entail:

- Character glyph changes. In general, the reference glyphs for most of the characters in N3583 are either too “pictorial” or far too “cute” to be suitable for generic standardization. We are convinced that a character named PIG or BEAR should be represented as generically as possible in the UCS, even if a Japanese telephone implementation may represent them as smiling disembodied heads. Whilst Japanese *anime* drawings may be quite suitable for some implementation environments, that style is too specific for generically-named characters. We are of course open to discussion of the improvements or replacements we have suggested. Note also that there are many glyphs in the tables below which also ought to be replaced on the same basis, though we did not have the time to devise new glyphs for everything.
- Character name changes. We have tried to improve the names where more suitable names could be found. We have not been able (in the short time available to us) to supply a list of name deltas, but one might be generated on the basis of our mapping table.

- Character re-arrangement into more logical subsections, with a small amount of built-in room for expansion of subsections which have semantic domains which are likely to attract additional character additions.
- Character additions. Of the 674 characters proposed in N3583 we have serious concerns about encoding only 14 of them. However, as will be seen below we suggest a way of mitigating that concern. We propose about a 20% increase in the number of graphic characters added.

We will discuss the changes block by block. Code points will be given as in our revised code tables. A mapping table from N3583 to this document is given at the end of the document. In those tables, characters we propose to add to the repertoire are given in yellow; characters from N3583 which are moved to different blocks are coloured grey (characters re-arranged within a block are not coloured).

2300 Miscellaneous Technical. N3583 proposes to add 5 characters to this block, four double-triangles and one clock. We propose to move 6 user-interface characters from elsewhere in N3583 to this block, and to add additional user interface characters commonly found on hardware and software telephones, digital audio players, and similar contexts. These additions provide a more complete set of symbols needed for the composition of user instructional text. For example, U+231A WATCH and the new U+23F0 ALARM CLOCK proposed in N3583 are well-complemented by our proposed U+23F0 STOPWATCH and U+23F0 TIMER CLOCK. Some additional double triangle characters are added, as well as some user-interface symbols having to do with sound and muting, and a variety of arrows associated with playlist navigation.

2700 Dingbats. N3583 proposes to add 3 characters to this block, a check mark and two handshakes. We propose to move 8 punctuation-like characters and one star sparkle from elsewhere in N3583 to this block.

1F0A0 Playing Card Pips. N3583 proposes to add a single Joker character, but this makes no sense in the absence of an encoded set of cards as proposed in N2760 “Proposal to encode dominoes and other game symbols in the UCS”. The source documents—as well as the reference glyph in N3583—indicate that a playing-card Joker is intended, and the repertoire of playing card characters is well-known and canonical. Accordingly we propose to add one Joker which can be mapped to the Joker used in the emoji environment, a second Joker as found in many playing card decks, a generic card back, and a set of suited cards for both the 52-card deck most commonly used today and the 56-card decks used in some European games. (The difference is that the latter include a Knight in addition to Jack, Queen, and King.) These cards map completely to the Minor Arcana of the Western Tarot (from which they derive), and unification is appropriate. The 58 characters proposed here comprise about 38% of the graphic characters we propose for addition.

1F100. Enclosed Alphanumeric Supplement. We concur with the Irish NB’s comments on FPDAM 6 with regard to filling out the repertoire here. We have not counted these in our percentage tally of the graphic characters we propose for addition.

1F300 Miscellaneous Pictographic Symbols. We have divided this block into a number of notional sections. The first three columns contain Geographical, Astronomical, and Moon, Sun, and Star symbols. The next five columns contain Plant, Fruit, Food, and Beverage symbols. The next five columns contain

Celebration, Entertainment, Game, Music, and Sport Symbols. The last three columns are earmarked for Building and Map symbols. 20 of the 190 symbols proposed in this block are newly-proposed by us.

1F400 Miscellaneous Pictographic Symbols-A. The first four columns here contain Animal symbols. Of these, the first 25 are mappable to the animals used by various traditions to represent the different signs of the Asian Zodiac. 72% of these were already proposed in N3583; we have added 7 (the RAT, BULL, OX, CROCODILE, GOAT, SHEEP, and ROOSTER).

The remaining seven columns contain Facial Parts, Hand Signal, Clothing and Accessory, Portrait and Role, Fairy Tale, Role, Personal Care, Medical, Romance, Heart, and Comic-book style symbols. 17 of the 167 symbols proposed in this block are newly-proposed by us.

1F4B0 Letterlike Symbols Extended-A. We have moved most of the more obvious letterlike symbols in N3583 to this block, and allocated a total of five columns for extensions. None of the 21 symbols proposed in this block were proposed by us.

1F500 Miscellaneous Pictographic Symbols-B. This half block is nearly full with Money, Office, Communication, Audio and Video, Tool, Geometric Shape, and User-Interface symbols. 9 of the symbols proposed in this block are newly-proposed by us.

1F580 Transport and Map Symbols. This half block is about half full with Vehicle, Traffic Signs, and generic Signage symbols. 35 of the 68 symbols proposed in this block are newly-proposed by us.

1F600 Emoticons. Four columns are proposed here, though in reasonable expectation of expansion it might conceivably be wise to allocate seven. 3 of the 62 symbols proposed in this block are newly-proposed by us.

1F700 Alchemical Symbols. One character, the Zodiacal character OPHIUCHUS, used in emoji context. has been allocated to the Alchemical symbols block, because Western Zodiac symbols are used in alchemical contexts and this seems a better fit than a generic block for pictographs. If it were possible to move this character to the Miscellaneous Symbols block (by moving characters on the ballot there to positions in a different block), that would be preferable as it would allow all of the Western Zodiac symbols to be in the same block.

E0200 Symbols for Two-letter Codes. N3583 proposes to add 10 “Emoji regional indicator symbols” which clearly represent the People’s Republic of China (CN), Germany (DE), Spain (ES), France (FR), the UK (GB), Italy (IT), Japan (JP), Korea (KR), Russia (RU), and the US (US). In the context of telephony, these “regional indicator symbols” are represented by national flags. Germany (one of the 10 countries “lucky” enough to be included in emoji glyph sets) concurs with Ireland (one of the 193 countries “unlucky” enough *not* to be included in emoji glyph sets) that this is an inappropriate use of the UCS. Politically, the proposal made in N3583 is both naïve (no one is fooled) and untenable (being prejudicial against 95% of the world’s sovereign states). However, there is a simple way forward that will solve this otherwise insoluble problem.

We propose to encode a new block of 676 Symbols for Two-letter Codes, exhaustively representing the letter combinations from AA to ZZ and encoding them in Plane 14. These can be used for any purpose requiring graphic characters that might represent in any way two-letter codes, regardless of the meaning of those codes. ISO 639-1 language codes and ISO 3166-1 country codes are of course two of the obvious uses of these characters.

E0500 Emoji Compatibility Symbols. Both Germany and Ireland have very, very strong reservations about this block of characters. It is clear that what it amounts to is “stealth encoding” of characters which would never be permitted to be encoded otherwise. In an emoji context, these are obviously graphic symbols, but N3583 says only:

“Special, rarely used, carrier-specific symbols are proposed for encoding in the Emoji compatibility symbols block. They are needed to complete the set for interoperability but are only identified by their source mappings (N3585), not specific glyphs and names.”

We think that “stealth encoding” describes this pretty accurately.

We understand that some members of the Unicode Consortium may consider the allocation of these codepoints for this purpose to be essential to interoperability. Nevertheless, we consider it quite problematic to “encode” characters in the way that has been proposed. Our minimum requirement is to move this block to Plane 14, since the characters have no glyphs and are not intended to be true “graphic” characters.

In any case, we are hopeful that the Unicode Consortium will consider our request for a relatively few additional symbol characters as given here to be as compelling as they consider their own requirements to be.

Mapping Table. The column on the right gives the N3583 code positions; the right column gives the code positions proposed here. A tab-delimited text file is available.

23E9	23F0	1F307	1F307	1F32B	1F342	1F34F	1F466
23EA	23E9	1F308	1F308	1F32C	1F343	1F350	1F467
23EB	23EA	1F309	1F309	1F32D	1F33A	1F351	1F468
23EC	23EB	1F30A	1F311	1F32E	1F33B	1F352	1F469
23ED	23EC	1F30B	1F314	1F32F	1F334	1F353	1F46A
2705	2705	1F30C	1F313	1F330	1F335	1F354	1F46B
270A	270A	1F30D	1F319	1F331	1F33E	1F355	1F46C
270B	270B	1F30E	1F315	1F332	1F33D	1F356	1F46D
275F	275F	1F30F	1F31B	1F333	1F344	1F357	1F46E
2760	2760	1F310	1F4BA	1F334	1F330	1F358	1F470
2E32	2E32	1F311	1F4BB	1F335	1F33C	1F359	1F471
1F12E	1F12E	1F312	1F4BC	1F336	1F33F	1F35A	1F472
1F170	1F170	1F313	23F3	1F337	1F352	1F35B	1F473
1F171	1F171	1F314	1F572	1F338	1F34C	1F35C	1F474
1F17E	1F17E	1F315	1F573	1F339	1F34E	1F35D	1F475
1F18E	1F18E	1F316	1F574	1F33A	1F34A	1F35E	1F476
1F121	1F121	1F317	1F575	1F33B	1F353	1F35F	1F477
1F122	1F122	1F318	1F576	1F33C	1F349	1F360	1F478
1F231	1F231	1F319	1F577	1F33D	1F345	1F361	1F479
1F232	1F232	1F31A	1F578	1F33E	1F346	1F362	1F47A
1F233	1F233	1F31B	1F579	1F33F	1F348	1F363	1F47B
1F234	1F234	1F31C	1F57A	1F340	1F34D	1F364	1F47C
1F235	1F235	1F31D	1F57B	1F341	1F347	1F365	1F47D
1F236	1F236	1F31E	1F57C	1F342	1F351	1F366	1F627
1F237	1F237	1F31F	1F57D	1F343	1F34F	1F367	1F47F
1F238	1F238	1F320	1F774	1F344	1F440	1F368	1F480
1F239	1F239	1F321	1F30A	1F345	1F442	1F369	1F481
1F250	1F250	1F322	1F30D	1F346	1F443	1F36A	1F482
1F251	1F251	1F323	1F30B	1F347	1F444	1F36B	1F415
1F300	1F300	1F324	1F30C	1F348	1F445	1F36C	1F408
1F301	1F301	1F325	1F340	1F349	1F483	1F36D	1F40A
1F302	1F302	1F326	1F337	1F34A	1F484	1F36E	1F425
1F303	1F303	1F327	1F331	1F34B	1F485	1F36F	1F426
1F304	1F304	1F328	1F341	1F34C	1F486	1F370	1F424
1F305	1F305	1F329	1F338	1F34D	1F487	1F371	1F428
1F306	1F306	1F32A	1F339	1F34E	1F464	1F372	1F41F

1F373	1F40E	1F3B0	1F619	1F3ED	1F462	1F42A	1F384
1F374	1F416	1F3B1	1F61A	1F3EE	1F453	1F42B	1F385
1F375	1F405	1F3B2	1F61B	1F3EF	1F455	1F42C	1F38C
1F376	1F42C	1F3B3	1F61C	1F3F0	1F456	1F42D	1F386
1F377	1F401	1F3B4	1F61D	1F3F1	1F451	1F42E	1F388
1F378	1F40C	1F3B5	1F61E	1F3F2	1F454	1F42F	1F389
1F379	1F435	1F3B6	1F61F	1F3F3	1F452	1F430	1F38D
1F37A	1F419	1F3B7	1F620	1F3F4	1F457	1F431	1F38E
1F37B	1F41A	1F3B8	1F621	1F3F5	1F458	1F432	1F393
1F37C	1F434	1F3B9	1F622	1F3F6	1F459	1F433	1F392
1F37D	1F427	1F3BA	1F623	1F3F7	1F45A	1F434	1F38F
1F37E	1F420	1F3BB	1F624	1F3F8	1F45B	1F435	1F387
1F37F	1F42B	1F3BC	1F625	1F3F9	1F500	1F436	1F390
1F380	1F41B	1F3BD	1F62A	1F3FA	1F501	1F437	1F383
1F381	1F418	1F3BE	1F62B	1F3FB	1F509	1F438	1F38A
1F382	1F429	1F3BF	1F62C	1F3FC	1F502	1F439	1F38B
1F383	1F412	1F3C0	1F62D	1F3FD	1F503	1F43A	1F391
1F384	1F40F	1F3C1	1F62E	1F3FE	1F504	1F43B	1F52F
1F385	1F42E	1F3C2	1F62F	1F3FF	1F505	1F43C	1F52E
1F386	1F404	1F3C3	1F630	1F400	1F400	1F43D	1F540
1F387	1F407	1F3C4	1F631	1F401	E0241	1F43E	1F541
1F388	1F40D	1F3C5	1F632	1F402	E0252	1F43F	1F52D
1F389	1F414	1F3C6	1F633	1F403	E027A	1F440	1F530
1F38A	1F417	1F3C7	1F634	1F404	E0293	1F441	1F539
1F38B	1F430	1F3C8	1F635	1F405	E029D	1F442	1F53A
1F38C	1F422	1F3C9	1F636	1F406	E02E3	1F443	1F53B
1F38D	1F42F	1F3CA	1F637	1F407	E02F9	1F444	1F53C
1F38E	1F421	1F3CB	1F638	1F408	E0315	1F445	1F53D
1F38F	1F41C	1F3CC	1F639	1F409	E03CE	1F446	1F533
1F390	1F436	1F3CD	1F63A	1F40A	E041A	1F447	1F534
1F391	1F423	1F3CE	1F63B	1F40B	1F547	1F448	1F531
1F392	1F409	1F3CF	1F63C	1F40C	1F45C	1F449	1F4AC
1F393	1F42D	1F3D0	1F63D	1F40D	1F45D	1F44A	1F535
1F394	1F435	1F3D1	1F3D0	1F40E	23F4	1F44B	1F536
1F395	1F41D	1F3D2	1F3D1	1F40F	1F5AA	1F44C	1F537
1F396	1F41E	1F3D3	1F3D2	1F410	1F4A9	1F44D	1F50A
1F397	1F600	1F3D4	1F3D3	1F411	1F55F	1F44E	1F50B
1F398	1F601	1F3D5	1F3D5	1F412	1F558	1F44F	1F51E
1F399	1F602	1F3D6	1F3D6	1F413	1F562	1F450	1F50C
1F39A	1F603	1F3D7	1F3D7	1F414	1F563	1F451	1F50D
1F39B	1F604	1F3D8	1F3D8	1F415	1F548	1F452	1F50E
1F39C	1F605	1F3D9	1F3D9	1F416	1F55E	1F453	1F51D
1F39D	1F606	1F3DA	1F3DA	1F417	1F55A	1F454	1F513
1F39E	1F607	1F3DB	1F3DB	1F418	1F54E	1F455	1F514
1F39F	1F608	1F3DC	1F3DC	1F419	1F52C	1F456	1F515
1F3A0	1F609	1F3DD	1F3DF	1F41A	1F54F	1F457	1F511
1F3A1	1F60A	1F3DE	1F3E0	1F41B	1F527	1F458	1F512
1F3A2	1F60B	1F3DF	1F3DD	1F41C	1F528	1F459	1F523
1F3A3	1F60C	1F3E0	1F3DE	1F41D	1F529	1F45A	1F526
1F3A4	1F60D	1F3E1	E0542	1F41E	1F525	1F45B	1F524
1F3A5	1F60E	1F3E2	1F532	1F41F	1F52A	1F45C	1F51B
1F3A6	1F60F	1F3E3	E0543	1F420	1F52B	1F45D	1F516
1F3A7	1F610	1F3E4	E0544	1F421	1F5BC	1F45E	1F51A
1F3A8	1F611	1F3E5	E0545	1F422	1F5BB	1F45F	1F518
1F3A9	1F612	1F3E6	1F55B	1F423	1F5BF	1F460	1F519
1F3AA	1F613	1F3E7	1F55C	1F424	1F5C0	1F461	1F517
1F3AB	1F614	1F3E8	1F55D	1F425	1F488	1F462	1F51C
1F3AC	1F615	1F3E9	1F45E	1F426	1F489	1F463	1F522
1F3AD	1F616	1F3EA	1F45F	1F427	1F380	1F464	1F51F
1F3AE	1F617	1F3EB	1F460	1F428	1F381	1F465	1F520
1F3AF	1F618	1F3EC	1F461	1F429	1F382	1F466	1F521

1F467	1F463	1F4A4	1F3B9	1F4E1	1F37B	1F51E	1F566
1F468	1F3BD	1F4A5	1F3BA	1F4E2	1F379	1F51F	1F567
1F469	1F3BE	1F4A6	1F3BB	1F4E3	1F56E	1F520	1F31F
1F46A	1F3BF	1F4A7	1F3BC	1F4E4	1F56F	1F521	1F320
1F46B	1F3C0	1F4A8	1F549	1F4E5	1F570	1F522	1F56A
1F46C	1F3C1	1F4A9	1F50F	1F4E6	1F571	1F523	1F56B
1F46D	1F3C2	1F4AA	1F510	1F4E7	2753	1F524	1F56C
1F46E	1F3C3	1F4AB	1F54A	1F4E8	2754	1F525	1F56D
1F46F	1F3C4	1F4AC	1F54B	1F4E9	2755	1F526	1F4AE
1F470	1F3C6	1F4AD	23F9	1F4EA	1F492	1F527	1F4AF
1F471	1F3C7	1F4AE	1F53F	1F4EB	1F493	1F528	1F4C3
1F472	1F3C8	1F4AF	1F48A	1F4EC	1F494	1F529	1F4C4
1F473	1F3CA	1F4B0	1F48B	1F4ED	1F495	1F52A	1F550
1F474	1F586	1F4B1	1F48C	1F4EE	1F496	1F52B	1F551
1F475	1F587	1F4B2	1F48D	1F4EF	1F497	1F52C	23FE
1F476	1F584	1F4B3	1F48E	1F4F0	1F498	1F52D	23FF
1F477	1F585	1F4B4	1F48F	1F4F1	1F499	1F52E	1F552
1F478	1F597	1F4B5	1F490	1F4F2	1F49A	1F52F	1F553
1F479	1F599	1F4B6	1F491	1F4F3	1F49B	1F530	1F554
1F47A	1F58D	1F4B7	1F546	1F4F4	1F49C	1F531	1F555
1F47B	1F58F	1F4B8	1F545	1F4F5	1F49D	1F532	1F4BD
1F47C	1F5A2	1F4B9	1F542	1F4F6	1F49E	1F533	1F556
1F47D	1F5B2	1F4BA	1F543	1F4F7	1F5AC	1F534	23FA
1F47E	1F589	1F4BB	1F354	1F4F8	1F5AD	1F535	1F538
1F47F	1F580	1F4BC	1F359	1F4F9	1F564	1F536	1F446
1F480	1F5A4	1F4BD	1F370	1F4FA	1F565	1F537	1F447
1F481	1F595	1F4BE	1F35C	1F4FB	1F5A9	1F538	1F448
1F482	1F5B6	1F4BF	1F35E	1F4FC	1F5B9	1F539	1F448
1F483	1F59A	1F4C0	1F373	1F4FD	1F5BA	1F53A	1F44A
1F484	1F592	1F4C1	1F366	1F4FE	1F5BC	1F53B	1F44B
1F485	1F591	1F4C2	1F35F	1F4FF	274C	1F53C	1F44C
1F486	1F593	1F4C3	1F361	1F500	274E	1F53D	1F44D
1F487	1F5A5	1F4C4	1F358	1F501	1F5AB	1F53E	1F44E
1F488	1F5A7	1F4C5	1F35A	1F502	1F557	1F53F	1F44F
1F489	1F3A8	1F4C6	1F35D	1F503	1F4B0	1F540	1F450
1F48A	1F3A0	1F4C7	1F35B	1F504	1F4B1	1F541	1F4BE
1F48B	1F3A1	1F4C8	1F362	1F505	1F4B2	1F542	1F4BF
1F48C	1F3A2	1F4C9	1F363	1F506	1F4B3	1F543	1F4C0
1F48D	1F3A3	1F4CA	1F371	1F507	1F4B4	1F544	1F4C1
1F48E	1F3A4	1F4CB	1F372	1F508	1F4B5	1F545	1F4C2
1F48F	1F3A5	1F4CC	1F367	1F509	1F4B6	1F600	E0500
1F490	1F3A6	1F4CD	1F356	1F50A	1F4B7	1F601	E0501
1F491	1F3A7	1F4CE	1F365	1F50B	1F4B8	1F602	E0502
1F492	1F3A8	1F4CF	1F360	1F50C	1F4B9	1F603	E0503
1F493	1F3A9	1F4D0	1F355	1F50D	2795	1F604	E0504
1F494	1F3AA	1F4D1	1F357	1F50E	2796	1F605	E0505
1F495	1F3AB	1F4D2	1F368	1F50F	2797	1F606	E0506
1F496	1F3AC	1F4D3	1F369	1F510	1F4A0	1F607	E0507
1F497	1F3AD	1F4D4	1F36A	1F511	1F4A1	1F608	E0508
1F498	1F3AE	1F4D5	1F36B	1F512	1F4A2	1F609	E0509
1F499	1F3AF	1F4D6	1F36C	1F513	1F4A3	1F60A	E050A
1F49A	1F3B0	1F4D7	1F36D	1F514	1F4A4	1F60B	E050B
1F49B	1F3B1	1F4D8	1F36E	1F515	1F4A5	1F60C	E050C
1F49C	1F3B2	1F4D9	1F36F	1F516	1F4A6	1F60D	E050D
1F49D	1F3B3	1F4DA	1F364	1F517	1F4A7	1F60E	E050E
1F49E	1F3B4	1F4DB	1F374	1F518	1F4A8	1F60F	E050F
1F49F	1F0CF	1F4DC	1F378	1F519	1F4AA	1F610	E0510
1F4A0	1F3B5	1F4DD	1F37A	1F51A	1F4AB	1F611	E0511
1F4A1	1F3B6	1F4DE	1F375	1F51B	2728	1F612	E0512
1F4A2	1F3B7	1F4DF	1F376	1F51C	1F568	1F613	E0513
1F4A3	1F3B8	1F4E0	1F377	1F51D	1F568	1F614	E0514

1F615	E0515	-----	1F0A3	-----	1F135	-----	1F187
1F616	E0516	-----	1F0A4	-----	1F136	-----	1F188
1F617	E0517	-----	1F0A5	-----	1F137	-----	1F189
1F618	E0518	-----	1F0A6	-----	1F138	-----	1F18E
1F619	E0519	-----	1F0A7	-----	1F139	-----	1F18F
1F61A	E051A	-----	1F0A8	-----	1F13A	-----	1F190
1F61B	E051B	-----	1F0A9	-----	1F13B	-----	1F30E
1F61C	E051C	-----	1F0AA	-----	1F13C	-----	1F30F
1F61D	E051D	-----	1F0AB	-----	1F13E	-----	1F310
1F61E	E051E	-----	1F0AC	-----	1F140	-----	1F312
1F61F	E051F	-----	1F0AD	-----	1F141	-----	1F316
1F620	E0520	-----	1F0AE	-----	1F143	-----	1F317
1F621	E0521	-----	1F0B1	-----	1F144	-----	1F318
1F622	E0522	-----	1F0B2	-----	1F145	-----	1F31A
1F623	E0523	-----	1F0B3	-----	1F147	-----	1F31C
1F624	E0524	-----	1F0B4	-----	1F148	-----	1F31D
1F625	E0525	-----	1F0B5	-----	1F149	-----	1F31E
1F626	E0526	-----	1F0B6	-----	1F14F	-----	1F332
1F627	E0527	-----	1F0B7	-----	1F150	-----	1F333
1F628	E0528	-----	1F0B8	-----	1F151	-----	1F336
1F629	E0529	-----	1F0B9	-----	1F152	-----	1F34B
1F62A	E052A	-----	1F0BA	-----	1F153	-----	1F350
1F62B	E052B	-----	1F0BB	-----	1F154	-----	1F37C
1F62C	E052C	-----	1F0BC	-----	1F155	-----	1F3C5
1F62D	E052D	-----	1F0BD	-----	1F156	-----	1F3C9
1F62E	E052E	-----	1F0BE	-----	1F158	-----	1F3D4
1F62F	E052F	-----	1F0C1	-----	1F159	-----	1F400
1F630	E0530	-----	1F0C2	-----	1F15A	-----	1F402
1F631	E0531	-----	1F0C3	-----	1F15B	-----	1F403
1F632	E0532	-----	1F0C4	-----	1F15C	-----	1F406
1F633	E0533	-----	1F0C5	-----	1F15D	-----	1F40B
1F634	E0534	-----	1F0C6	-----	1F15E	-----	1F410
1F635	E0535	-----	1F0C7	-----	1F160	-----	1F411
1F636	E0536	-----	1F0C8	-----	1F161	-----	1F413
1F637	E0537	-----	1F0C9	-----	1F162	-----	1F42A
1F638	E0538	-----	1F0CA	-----	1F163	-----	1F431
1F639	E0539	-----	1F0CB	-----	1F164	-----	1F432
1F63A	E053A	-----	1F0CC	-----	1F165	-----	1F433
1F63B	E053B	-----	1F0CD	-----	1F166	-----	1F441
1F63C	E053C	-----	1F0CE	-----	1F167	-----	1F465
1F63D	E053D	-----	1F0CF	-----	1F168	-----	1F46F
1F63E	E053E	-----	1F0D1	-----	1F169	-----	1F47E
1F63F	E053F	-----	1F0D2	-----	1F170	-----	1F4AD
1F640	E0540	-----	1F0D3	-----	1F171	-----	1F506
1F641	E0541	-----	1F0D4	-----	1F172	-----	1F507
-----	23ED	-----	1F0D5	-----	1F173	-----	1F53E
-----	23EE	-----	1F0D6	-----	1F174	-----	1F544
-----	23EF	-----	1F0D7	-----	1F175	-----	1F54C
-----	23F1	-----	1F0D8	-----	1F176	-----	1F54D
-----	23F2	-----	1F0D9	-----	1F177	-----	1F559
-----	23F5	-----	1F0DA	-----	1F178	-----	1F560
-----	23F6	-----	1F0DB	-----	1F17A	-----	1F561
-----	23F7	-----	1F0DC	-----	1F17D	-----	1F581
-----	23F8	-----	1F0DD	-----	1F17E	-----	1F582
-----	23FB	-----	1F0DE	-----	1F180	-----	1F583
-----	23FC	-----	1F0DF	-----	1F181	-----	1F588
-----	23FD	-----	1F12E	-----	1F182	-----	1F58A
-----	1F0A0	-----	1F130	-----	1F183	-----	1F58B
-----	1F0A1	-----	1F132	-----	1F184	-----	1F58C
-----	1F0A2	-----	1F133	-----	1F185	-----	1F58E
-----		-----	1F134	-----	1F186	-----	1F590

----- 1F594	----- 1F5A0	----- 1F5B3	----- 1F5C3
----- 1F596	----- 1F5A1	----- 1F5B4	----- 1F5C4
----- 1F598	----- 1F5A3	----- 1F5B5	----- 1F626
----- 1F59B	----- 1F5A6	----- 1F5B7	----- 1F628
----- 1F59C	----- 1F5AE	----- 1F5B8	----- 1F629
----- 1F59D	----- 1F5AF	----- 1F5BE	
----- 1F59E	----- 1F5B0	----- 1F5C1	
----- 1F59F	----- 1F5B1	----- 1F5C2	

A. Administrative

1. Title

Towards an encoding of symbol characters used as emoji.

2. Requester's name

Irish and German National Bodies

3. Requester type (Member body/Liaison/Individual contribution)

Member Body contribution.

4. Submission date

2009-04-06

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical -- General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

Proposed name of script

Playing Card Pips, Miscellaneous Pictographic Symbols, Miscellaneous Pictographic Symbols-A, Letterlike Symbols Extended-A, Miscellaneous Pictographic Symbols-B, Emoticons, Alchemical Symbols, Symbols for Tw0-letter Codes, Emoji Compatibility Symbols.

1b. The proposal is for addition of character(s) to an existing block

1b. Name of the existing block

Miscellaneous Technical, Dingbats, Supplemental Punctuation, Enclosed Alphanumeric Supplement, Enclosed Ideographic Supplement.

2. Number of characters in proposal

1544 (1a: 1424 = (59 + 190 + 167 + 21 + 178 + 62 + 1 + 676 + 70)) + (1b: 120 = (23 + 12 + 1 + 71 + 13)). To give an estimate to put this into perspective, 1544 less 666 of the Two-Letter-Code characters is 878; 878 less the 674 characters asked for in N3583 is 204; 204 less the 67 Enclosed Alphanumerics asked for by Ireland in ballot comments on FPDAM 6 is 137, the real number of additions proposed by us (about a 20% increase on the 674 asked for in N3583).

3. Proposed category (see section II, Character Categories)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the character naming guidelines in Annex L of ISO/IEC 10646-1: 2000?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson and others.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

No.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

No.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc.,

Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

The characters should have the same properties as other generic dingbat symbols.

C. Technical -- Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No.

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

No.

4a. The context of use for the proposed characters (type of use; common or rare)

Generic symbol use.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In Japan and elsewhere.

6a. After giving due considerations to the principles in Principles and Procedures document (a WG 2 standing document) must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

12a. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

12b. If YES, reference

13a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

13b. If YES, describe in detail (include attachment if necessary)

14a. Does the proposal contain any Ideographic compatibility character(s)?

No.

14b. If YES, is the equivalent corresponding unified ideographic character(s) identified?