Opinions received at WG2 #31 in Quebec on Ideographic Composition

ISO/IEC JTC1/SC2/WG2/IRG **N392**

Date: 1996/08/16

N 1430 (Ideographic Structure Characters)

Note: In the following, a structure character is referred to as a structure operator, and the arguments to the operator as structure operands.

(1) What is the scope of a structure operator? That is, how does one determine the operands of an operator? Does it apply only to individual characters or may it apply to an entire CC-data-element (a sequence of characters)?

(2) May a structure operator apply to only Unified CJK Ideographs or may it apply to any character?

(3) If it can apply to a CC-data-element, may that CC-data-element contain a structure operator? In other words, is *nesting* allowed?

(4) If nesting is allowed, is there a mechanism for grouping a CC-data-element so as to interpret it as a single operand to a structure operator?

(5) If nesting is allowed, are structure operators prioritized so as to determine operand scope in the absence of grouping?

(6) Is a CC-data-element consisting of a structure operator along with its operands constitute a composite character sequence, or does it represent something else? If not, then is there a proposed designation for such a sequence? Perhaps *structured ideograph sequence*?

Note: In the following, a CC-data-element consisting of a structure operator and its operands is referred to as a structured ideograph sequence.

(7) May a structured ideograph sequence be used to represent a coded ideograph (i.e., an ideograph already coded in the standard)?

(8) If coded ideographs may be alternately represented in this fashion, are alternate representation(s) considered to represent the same information (i.e., construed as *equivalent*)? Is more than one alternate representation possible? Will equivalence tables be provided for coded ideographs in order to indicate their alternate equivalent representation(s)?

(9) Is another implementation level needed which admits structured ideograph sequences?

(10) Some structured ideograph sequences may be graphically impossible according to the conventions of ideograph formation; for example, structure operators #5 through #B require that the operand representing the *outer* component be of a certain type. If the outer component is not of the required type, then does such a structured ideograph sequence constitute an illegal CC-data-element? How should applications respond to such illegal CC-data-elements?

(11) Given the apparent restrictions on the possible operands of each structure operator, is a table provided enumerating each permissible operand for each operator?

(12) It appears that structure operators #5 through #B may be unified into a single *outside-to-inside* operator? This is strongly indicated given the implicit graphical properties of possible outside components. If they are not unified only due to a small number of exceptions, are there other ways of representing the exceptions or may the exceptions simply be excluded from representation?

N 1431 (Ideographic Variation Mark)

Note: In the following, a variation mark character is referred to as a variation operator, and the argument to the operator as the variation operand.

(1) Similarity of ideographs may occur along one or more axes: X (meaning), Y (abstract shape), or Z (glyph image). To which one or more of these axes does this variation mark apply?

(2) Is there a requirement to distinguish between different axes of variation? If so, should more than one variation mark be proposed?

(3) May this variaton operator apply only to ideographs or may it apply to other characters? May it apply to CC-data-elements, such as a structured ideograph sequence (see above comments on N1430)?

(4) Is a CC-data-element consisting of a variation operator and its operand designated in a special manner? Perhaps *structured ideograph sequence*?

Note: In the following, a CC-data-element consisting of a variation operator and its operand is referred to as a variation ideograph sequence.

(5) What properties (sorting, lexical, etc.) obtain for a variation ideograph sequence? Is it to be identical to the properties of the variation operand?

N 1432 (Ideographic Radical Supplement)

(1) This request addresses only a portion of the full radical set, relying on previously coded ideographs to serve as the remainder of the radical set. Given that a request is outstanding for a full KangXi radical set (N1182), is it better to encode a complete radical set, including the radicals proposed by this supplement, as a separate set of characters independent from the Unified CJK Ideograph repertoire?

(2) If a separate, complete radical set is not to be encoded, what is the rationale for encoding some radicals as Unified CJK Ideographs and others as radical characters which are considered to be distinct from the unified ideograph repertoire?

(3) What unification principles were applied in selecting these radicals?

(4) May radicals proposed by this supplement be used with ideograph structure characters or with ideograph variation mark?

N 1433 (Ideographic Component Supplement)

(1) This request addresses only a portion of a potentially larger component set, relying on previously coded ideographs to serve as the remainder of the component set. Is it better to encode a complete component set, including the components proposed by this supplement, as a separate set of characters independent from the Unified CJK Ideograph repertoire?

(2) If a separate, complete component set is not to be encoded, what is the rationale for encoding some components as Unified CJK Ideographs and others as component characters which are considered to be distinct from the unified ideograph repertoire?

(3) What unification principles were applied in selecting these components?

(4) May components proposed by this supplement be used *only* with ideograph structure characters or may they be used independently, apart from structure characters?