

문자코드(JTC1/SC2) 국제 표준화 동향

정보표준과 공업연구원 남택주

02) 509-7333 ijnam@ats.go.kr

1. ISO/IEC JTC1/SC2(글자 부호계, Coded Character Set) 분야 추진 전략

북한 글자 부호계를 주로 다루기 위하여 2000년에 SC2/WG2 내에 만들어 졌다.

1.1 ISO/IEC JTC1/SC2(글자 부호계, Coded Character Set) 분야 현황

1.2 ISO/IEC JTC1/SC2(문자코드) 회원 국 현황

SC2는 글자(character)의 부호화, 다시 말하여 글자 부호계에 관한 표준을 다루는 Sub-committee로 간략하게 소개하고자 한다.

SC2는 1947년도에 설립되어 정회원(P-member)으로 우리나라를 비롯하여 북한(DPRK), 미국, 캐나다, 노르웨이, 일본 등 35개 국가와 O-member로는 베트남을 비롯하여 13개 국가들이 가입하고 있다.

1차 회의에 대한 구체적인 정보는 찾을 수 없으나, SC2 2차 총회가 1990년 4월에 미국에서 처음 개최되었다.

- 국제의장 : Prof. K. Shibano (Japan) (국내위원장 : 부산대 김경석)

- 국제간사 : Ms. T. Kimura (Japan, JISC)

현재 SC2 아래에는 WG2 (여러 바이트 글자 부호계)와 WG3 (한 바이트 부호계, 부호계 확장법) 2개의 WG으로 구성되어 있으며, WG1은 활동하지 않고 있는 실정이다.

동 SC2 총회는 매년 한번씩 개최되는데 참가인원 수는 30여명이다. WG2 회의는 그 동안 1년에 보통 2회씩 열렸으며 전문가가 보통 40여명 참석하였으나, 앞으로 2003년부터는 매년 1회씩 개최될 예정이다.

또한, WG2 아래에는 IRG (국제한자특별위원회)와 Korean Script ad hoc group (한글 특별위원회)이 있다. 한글특별위원회는 ISO/IEC 10646에 들어갈 남

현재 WG3는 정기적인 모임은 없지만, 필요한 토론은 1차적으로 전자 우편으로 처리하며, 특별히 회의를

소집해야 할 안전이 있을 경우에는 회의를 소집하기로 2000년에 결정한 바 있다. 따라서 사실상 활동이 거의 없는 상태이다.

1.3 국가별 WG2 회의 참석 현황

SC2의 기술적인 사항은 모두 WG2 회의에서 다루어지기 때문에 WG2에 한하여 최근의 자료를 참고로 하여 하면 참석 현황은 다음과 같다.

(제43차(2002.12. 일본), 제42차 (2002.05, 아일랜드), 제41차 (2001.10, 싱가포르), 제40차 (2001.04, 샌프란시스코) WG2 회의 참석자를 기준으로 한 것임)

가장 많이 참석하는 국가는 미국으로 다국적 기업을 중심으로 회사 이익을 위하여 가장 열심히 참여하고 있으며, 두 번째로 많이 참가하는 국가는 일본인데, 자국과 관련된 현안은 별로 없지만 다국적 기업 직원이 많이 참석하고 있다. 세번째로는 중국으로 한자 및 중국 안의 소수 민족 글자계 문제로 많이 참석하고, 네 번째로 많이 참여하는 국가는 북쪽 (DPRK)인데, 한글 문제 때문에 열심히 참여하고 있으며, 북한 다음으로 우리나라가 많이 참가하고 있으며, 한글 문제 및 남북 문제와 관련하여 참여하고 있다.

WG2 제40~43차 Meeting 참석자 현황

국 가	43th	42th	41th	40th	합 계 (%)
미 국	5	6	3	8	22(21%)
한 국	2	2	1	3	8(8%)
북 한	2	2	4	4	12(12%)
중 국	7	6	1	2	16(16%)
일 본	7	5	3	3	18(17%)
아일랜드	1	4	0	2	7(7%)
캐나다	1	1	2	2	9(6%)
싱가포르	1	1	2	0	4(4%)
스웨덴	0	1	2	1	4(4%)
핀란드	1	2	0	0	3(3%)
이 란	2	0	0	0	2(2%)
리트아니아	0	0	0	1	1(1%)
합 계	29	30	18	26	103명 (100%)

2. ISO/IEC JTC1/SC2 세부 기술분야 (WG) 현황

SC2 산하에 WG2 (여러 바이트 부호계)와 WG3 (한 바이트 부호계 및 부호계 확장)의 2개 WG을 두고 있다.

2.1 세부 기술 분야(WG)별 국제 표준화 프로젝트 현황

1) WG2

ISO/IEC 10646 국제규격은 현재 Part 1와 Part 2로 나누어져 있지만 하나로 통합하는 작업을 수행중에 있다. 또 한 규격인 ISO/IEC 15285(Character Glyph model)는 이미 통합 작업이 완료되어, WG2에서 다루는 과제는 사실상 ISO/IEC 10646 한 개뿐이라고 볼 수도 있다.

또한, ISO/IEC 10646은 전 세계 국가의 부호계를 다루고 있어 규격은 하나지만, 실제 내용상 40여개의 개별규격이라고 보아도 된다.

2) WG3

이 WG에는 43 개 규격이 출판되어 있으나, 현재 WG3의 활동은 거의 없으며, 앞으로도 활동은 많지 않을 것으로 예상하고 있다.

3. ISO/IEC JTC1/SC2/WG2의 표준화 프로젝트 세부 분석

국제규격 ISO/IEC 10646의 Part 1과 Part 2의 통합작업은 2003년도에 추진할 계획으로 있으며, 이에 따라 우리나라의 국가표준인 KS X 1005-1과 KS X 10646-2의 규격도 KS X 1005로 만드는 작업을

앞으로 추진해야 할 과제이다. 그러나 현재에는 ISO/IEC 10646-1, ISO/IEC 10646-2에 계속하여 글자계를 추가하는 작업을 하고 있는 것이 다른 국제규격과 다른 점이라 할 수가 있다.

3.1 ISO/IEC 10646-1

국제규격 ISO/IEC 10646-1에 대응되는 국가표준은 KS X 1005-1이다.

1) 적용범위

이 규격은 국제문자 부호계 (UCS: Universal Multiple-Octet Coded Character Set)를 규정하고 전 세계 언어 및 기호의 필기(쓰여진) 꼴 (written form)의 표현, 전송, 교환, 처리, 저장, 입력과 출력에 적용할 수 있다.

이 규격의 제1부에서는 전체적인 구조 및 다음 사항을 규정한다.

- (1) 이 규격에서 쓴 용어를 정의한다.
- (2) 이 부호계의 일반적인 구조를 기술한다.
- (3) 국제 문자 부호계의 기본 다국어 평면(BMP: Basic Multilingual Plane)을 규정하고 전 세계 언어의 글자계(script)와 필기 꼴에 쓰는 모양 문자(graphic character)의 집합을 정의한다.
- (4) 기본 다국어 평면에 있는 도형 문자의 이름과 부호화된 표기법을 규정한다
- (5) 국제 문자 부호계의 4 Octet(32 Bit) 정규형 (canonical form) 인 UCS-4 를 규정한다.
- (6) 국제 문자 부호계의 2 Octet(16 Bit) 기본 다국어 평면 꼴 (BMP form) 인 UCS-2 를 규정한다.
- (7) 제어 기능 (control function)에 대한 부호화된 표기를 규정한다.

(8) 이 문자 부호계에 앞으로의 추가에 대한 관리를 규정한다.

특히, 이 규격에서 규정하는 국제문자 부호계는 ISO/IEC 2022에 규정되어 있는 것과는 다른 부호계이다.

2) 주요 내용(우리 나라와 관련된 사항)

북한(DPRK)에서 제안한 글자 70자 가운데 39자가 추가될 것으로 예상되며, ISO/IEC 10646-1(= KSX 1005-1)에는 한글과 관련하여 첫가끝 조합형 240 글자가 U+1100~U+11FF에, 완성형 글자마다 11,172 개가 U+AC00~U+D7A3에 들어가 있다. 그 밖에도 호환형 한글 글자, 반각 한글 글자 등이 들어가 있다.

3) 주도 국가 및 관련 핵심 전문가

이 국제규격에는 미국, 일본, 중국, 한국 등 아주 많은 나라가 주도적으로 참여하고 있다.

- (1) 의장 : WG2 Mike Ksar
- (2) Editor (Project Leader) : 쉬냐드 (미국)
- (3) 주도 전문가 : 미국 (유니코드 회사 소속으로 미국 대표로 참석) 및 그 밖의 여러 나라 전문가들이 있음.
- 4) 국내 표준화 추진 현황 (수행 주체별로 분류 - 있을 경우)

ISO/IEC 10646-1:2000은 2001년에 KSX 1005-1:2002 (제2판)를 KS규격으로 제정할 계획이다.

5) 국내 기술 현황, 구축 사례, 시장 및 기업 현황

- (1) 전 세계적으로 ASCII 다음에 쓰게 될 글자 부호계
- (2) 보기를 들어 아래 아 한글의 경우에 이 부호계 (한글 부분 포함)를 지원

(3) 마이크로소프트에서도 이 부호계(한글 부분 포함) 지원

(4) 앞으로 ISO/IEC 10646의 한글 관련 사항에 대하여 집중적으로 연구 할 필요가 있다.

3.2 ISO/IEC 10646-2:2001

국제규격 ISO/IEC 10646-2에 대응되는 국가표준은 KSX ISO/IEC 10646-2이다. 2003년부터 ISO/IEC 10646의 Part 1과 Part 2의 통합 작업이 추진되고 있으며, 이에 따라 KSX 1005-1과 KSX ISO/IEC 10646-2도 KSX 1005 하나로 통합하는 작업을 해야 한다.

1) 적용 범위

이 규격은 국제 문자 부호계(UCS: Universal Multiple-Octet Coded Character Set)를 규정하며, 추가적인 기호뿐만 아니라 전 세계 언어의 필기(쓰여진) 꼴(written form)의 표현, 전송, 교환, 처리, 저장, 입력과 출력에 적용할 수 있다. 이 규격의 Part 1에서는 전체적인 구조 및 국제 문자 부호계(UCS)의 기본 다국어 평면 (BMP)과 Part 2에서는 다음과 같이 보충 평면(Supplementary Planes) 세 개를 규정한다.

- (1) 다국어 보충 평면(Supplementary Multilingual Plane)을 규정한다.
- (2) 국제 문자 부호계(UCS)의 한자 보충 평면(SIP: Supplementary Ideographic Plane)을 규정하고, 동 아시아에서 쓰는 모양 글자의 집합을 정의한다. 이것은 한중일 한자(동 아시아 한자: East Asian Ideographs)로 정의되어 있다.
- (3) 특별 목적 보충 평면(SSP: Supplementary Special-purpose Plane)을 규정하고, 특별한 목적에 쓸 모양 글자 집합을 정의한다.



- 이 여러 평면의 모양 글자를 위한 이름을 규정하고 국제 문자 부호계의 4 Octet(32 Bit) 정규형 (canonical form)을 써서 부호화된 표현을 규정한다.
- Part 1에서 이미 부호화된 모양 글자는 보충 평면에 되풀이하지 않는다. 더욱이 어떤 글자에, 이 규격의 이 부에 규정된 보충 평면 집합 안에 부호 자리 한 개만 할당된다.

2) 주요 내용

우리 나라와 관련된 사항으로는 아주 드물게 쓰는 한자가 여기에 들어 있으며, ISO/IEC 10646-2에는 ISO/IEC 10646-1에 산입되지 못한 글자를 별도 추가하여 규정하고 있다.

- 예를들어 한자 Ext. B가 이 규격에 해당됨.
- 수정판 1에 들어가는 글자는 애제 글자등 아주 드물게 쓰는 글자임.

3) 주도 국가 및 관련 핵심 전문가

ISO/IEC 10646-1과 동일함

4) 국내 표준화 추진 현황

ISO/IEC 10646-2:2001은 2002년에 KSX ISO/IEC 10646-2:2002(제1판) 로 KS규격으로 제정하였다.

5) 국내 기술 현황, 구축 사례, 시장 및 기업 현황

ISO/IEC 10646-2을 UTF-16 방식으로 지원하려면 2/4 (2/4) Byte 체계를 사용하여야 하나, ISO/IEC 10646-1 (BMP)는 보통 2 Byte 체계로 처리를 해야 하며, ISO/IEC 10646-2는 보통 4 Byte 체계로 처리해야 한다.

따라서 Part 1과 Part 2를 동시에 지원하려면 2/4 Byte 체계를 써야 한다. 이러한 이유 때문에 아직 ISO/IEC 10646-2를 지원하는 프로그램은 많지 않다.

3.3 이미 발간된 WG2 규격의 간략한 소개 및 대응 전략

1) ISO/IEC TR 15285:1998

ISO/IE TR 15285의 규격명은 An Operational Model for Characters and Glyphs (Technical Report Typ 3)로 한글의 경우 첫가끝 조합형의 각 글자(U+11xx)를 조합하여 글자마디 (가, 각, 간, ...)을 만들 수 있는데, 이 때 부호값 어떤 것이 조합하면 글자마디 어떤 것이 나오는지에 대한 풀이를 할 수 있는 것이 이 기술 보고서의 주요 내용이다.

우리의 한글이 동 규격과 관련성이 매우 깊으며, 이 기술 보고서를 만들 때 WG2 국제회의에서 SC2 국내 위원장(부산대 김경석교수)이 검토 의견을 제출한 바 있다. 앞으로 이 기술 보고서를 KS규격으로 제정할 것인지에 검토를 하여야 한다.

4. ISO/IEC JTC1/SC2(WG2) 향후 추진방향

JTC1/SC2(문자코드)는 한글과 관련 있는 국제 표준화 기구로서 북한과 우리가 국제규격으로 추진하는 바가 달라 지속적인 협의가 요구되고북한의 제안내용 등 국제동향을 수시로 파악하여 대처해야 하는 매우 중요한 분야이다.

ISO/IEC 10646은 ASCII를 대신할 “차세대 부호계”이므로 아주 중요하며, 우리 나라에서는 이미 Part 1을 KS로 제정하였으나, Part 2의 KS 제정은 10646의 Part 1과 Part 2를 하나로 통합하는 과정을 지켜 보면서 KS규격으로 반영하여야 하고, 나아가 남북한이 한글 부호계 단일화를 위하여 이를 공동규격으로 채택할 그 날을 손꼽아 기다려 본다.

