

**Agenda for WG2 Ad Hoc Meeting on Mongolian
Hohhot, Inner Mongolia
September 22-25, 2017**

Meeting location: Room 6 on F4 of HNA Grand Hotel Portman Hohhot

(see: <http://www.unicode.org/wg2/docs/n4799-HohotMeetingLogisticsUpdated2017-05-12.pdf> includes map)

Meeting times: Sat. and Sun. 9 a.m. – 12 noon, 2 p.m. – 5 p.m.

Monday 11 a.m. – 12 noon, 2 p.m. – 5 p.m.

DRAFT AGENDA

Friday, 22 September

- Evening: informal get-together of participants (HNA Grand Hotel Portman)

Saturday, 23 September

- Chair: Chen Zhuang
 - Morning: Welcome by host (Prof. Nasun-urt)
 - Brief presentation of meeting goals (Deborah Anderson)
 - Introductions of participants
 - Introduction to N4880 (=L2/16-309)
 - Afternoon Chair: Deborah Anderson
- General discussion
Discussion Facilitator: Peter Constable
Topic: Identification of current problems in Mongolian (see background document, page 3 of this doc)

Goal for Saturday: Get to know one another, intro to N4880, and identification of current problems

Evening: Social gathering of participants

Sunday, 24 September

- Chair: Deborah Anderson
 - Discussion
- Discussion Facilitator: Peter Constable
Topic: Investigate approaches to solve the identified problems
- Note 1: Participants are invited to give their input on why their approach is a good way to solve the identified problems
 - Note 2: N4880 (=L2/16-309) can be used to explain how improving the phonetic model will solve the issues identified

Goal for Sunday: Gather input from all participants on the strengths of their preferred approach to problems (outlined on Saturday).

Monday, 25 September

- Chair: Deborah Anderson
 - Discussion
- Discussion Facilitator: Peter Constable
Topics:
- Detailed exploration of models, with comparison and contrast of approaches
 - Examination of examples
 - Wrap-up and identification of next steps.

Goal for Monday: Discuss various models in detail and capture the consensus from the ad hoc for the ad hoc report

PARTICIPANTS

Participants (Chinese delegates nominated by the Chinese official Mongolian working group)

Mr. Chen Zhuang, China Electronics Standardization Institute
Prof. Quejingzhabu, Inner Mongolia University
Prof. Nashunwuritu (Nasun-urt), Inner Mongolia University
Prof. Menghejiya, Inner Mongolia University
Prof. Baohaishan, Inner Mongolia University of Finance and Economics (He is also working for Delxhii, whose CEO is Jirimutu.)
Mr. Baishuangcheng, Menksoft
Mr. Husile, Menksoft
Prof. Wu Jian, Software Insititute, Chinese Academy of Science
Mr. Liang Jinbao, Ethnic Affairs Commission of Inner Mongolia Autonomous Region
Mr. Liushisan, Inner Mongolian Publishing Group

Participants list from Debbie Anderson

Peter Constable, Unicode Consortium liaison to WG2
Michel Suignard, Editor ISO 10646/WG2 Convenor
Debbie Anderson, UC Berkeley
Lisa Moore, Unicode Technical Committee Chair
Andrew Glass, Microsoft, specialist in Uniscribe rendering engine/fonts, especially for complex writing systems
Greg Eck, Greyson Translation Services
Liang Hai, individual expert
Zheng Weizhe, individual expert (interested especially in Ali Gali and Todo/Manchu/Sibe)
Shen Yilei, individual expert
Wang Yihua, postdoctoral researcher, Department of Chinese, Fudan University
Badral Sanlig, Software architect, www.bolorsoft.com, Ulaanbaatar
Ms. B. Undraa, MASM, Ulanbaatar (NB rep from Mongolia)
B. Tsogt (Tsogoo), Ulanbaatar, Greyson Translation Services
Andrew West, UK NB rep
Jirimutu, Almas Inc in Japan

DOCUMENTS

N [4880](#) Proposed additions for Mongolian in 5th edition of UCS – Chen Zhuang et al. ([L2/16-309](#))

N [4882](#) Script Ad Hoc Group Recommendations on Mongolian Text Model (in English and Chinese)– Whistler et al. ([L2/17-328](#))

N [4889](#) A graphetic approach for the Mongolian encoding model - Script ad hoc group

N [4890](#) On migration issues of the graphetic model (in English and Chinese) – Script ad hoc group

N [4884](#) Comments on positional mismatches in Mongolian encoding – Shen Yilei (Note: A newly revised version, dated 18 September, will be posted shortly on the N4884 link; the current version posted is dated 11 August)

N [4883](#) A mixed encoding scheme for the Mongolian block – Weizhe Zheng (Note: A newly revised version will be posted shortly on the N4883 link, with a date of approximately 18 September; the current version posted is dated 11 September)

Background document for Saturday's discussion on current problems

(from page 4 of <http://www.unicode.org/L2/L2017/17328-mongolian-ad-hoc-rec.pdf>)

- Model is very complex, is underspecified, and has resulted in ambiguous representation of text.
- Assumes very complex contextual rules that have to be implemented in fonts and rendering engines, but without clear specification of those rules.
- Requires large set of variation sequences to force particular outcomes for gaps in the rules.
- Very difficult for implementers and font providers.
- Results in de facto non-interoperability of different implementations and fonts.
- Is subject to attempts to “fix” interoperability by an unstable stream of patches with unstable rationales.
- Will not be supported in internationalized domain names.
- For the average user, searching can be very difficult, because input text for matches is ambiguous.
- Has significant security concerns.
- Phonetic information in the text representation is not actually sufficient for text-to-speech or linguistic analysis applications.

- 模型非常复杂，详细规范不足，而且导致有歧义的文本标记。
- 认为非常复杂的上下文规则必须在字体和渲染引擎中实现，但对于那些规则又没有清楚的规范。
- 在规则的空白处要求很大一组变体序列来强行获得特定的结果。
- 对于实现方和字体提供方非常困难。
- 导致不同的实现和字体之间事实上不可互通。
- 会遭受“修正”互通性的尝试，不稳定的连串修补基于不稳定的理由。
- 在国际化域名中不会得到支持。
- 对于一般用户，搜索会很困难，因为用于匹配的输入文本是有歧义的。
- 显著的安全问题令人担忧。
- 对于文本至语音转换和语言学分析用途，文本标记中的语音信息其实并不充分。

WG2 蒙古文特别会议日程

内蒙古 呼和浩特

2017年9月22日-9月25日

会议地点：内蒙古博曼海航大酒店四层第6会议室

(见：<http://www.unicode.org/wg2/docs/n4799-HohotMeetingLogisticsUpdated2017-05-12.pdf> 含地图)

会议时间：周六上午9点到12点，下午2点到5点

周日上午9点到12点，下午2点到5点

周一上午11点到12点，下午2点到5点

会议日程

9月22日，星期五

- 晚上：非正式的参会代表聚会（内蒙古博曼海航大酒店）

9月23日，星期六

- 主席：陈壮
- 上午：主办人那顺乌日图教授开幕式致辞
 - 黛比·安德森对会议主题的简单介绍
 - 对参会代表的介绍
 - 介绍 [N4880](#) (=L2/16-309)
- 下午主席：黛比·安德森

广泛讨论

评议员：皮特·康斯特布尔

主题：识别蒙古文字数字化目前出现的未解决的问题（见背景介绍，第3页）

周六会议目标：互相了解，介绍 [N4880](#) 并了解并确定当前蒙古文数字化的问题。

- 晚上：与会代表交流

9月24日，星期天

- 主席：黛比·安德森
- 论坛

评议人：皮特·康斯特布尔

主题：就目前出现的问题调查解决方案

- 注 1：鼓励与会代表介绍为什么自己的方案有利于解决已识别出的问题
- 注 2：N4880 (=L2/16-309)可用于解释为什么改善语音模型可较好的解决现存问题。

会议目标：收集所有与会代表对于他们首选方案的想法（主要针对周六与会提出的问题）。

9月25日，星期一

- 主席：黛比·安德森
- 讨论

评议人：皮特·康斯特布尔

主题：

- 比较和对比各种方案以深入探讨各个模型；
- 检验各种示例
- 总结，并识别下一步工作。

会议目标：详细讨论不同模型，并从特别会议中记录一致意见，为特别会议报告做准备。

参会代表

中方参会代表（由中方蒙古文工作组提名）

陈壮：中国电子技术标准化研究所

确精扎布教授：内蒙古大学

那顺乌日图教授：内蒙古大学

孟和吉雅教授：内蒙古大学

包海山教授：内蒙古财经大学（内蒙古德力海信息技术有限公司经理，总经理吉日木图）

白双成：内蒙古蒙科立软件有限责任公司

呼斯勒：内蒙古蒙科立软件有限责任公司

吴健教授：中国科学院软件研究所

梁金宝：内蒙古自治区民族事务委员会

六十三：内蒙古出版集团

黛比·安德森所列与会代表

皮特·康斯特布尔 (Peter Constable)：Unicode 联盟在 WG2 的联络员

米歇尔·叙伊尼亚尔 (Michel Suignard)：ISO 10646/WG2 会议召集人

黛比·安德森 (Debbie Anderson)：加州大学伯克利分校语言系教授

丽萨·摩尔 (Lisa Moore)：Unicode 技术委员会主席

安德鲁·格拉斯 (Andrew Glass)：微软，Uniscribe 渲染引擎和字体专家，特别擅长于复杂书写系统

格雷格·艾克 (Greg Eck)：格雷森翻译服务 (Greyson Translation Services)

梁海：专家

郑维喆：专家 (主攻阿礼嘎礼以及托忒文、满文、锡伯文)

沈逸磊：专家

王奕桦：复旦大学中文系博士后

Badral Sanlig, Software architect, www.bolorsoft.com, Ulaanbaatar
Ms. B. Undraa, MASM, Ulanbaatar (NB rep from Mongolia)

B. Tsogt (Tsogoo), Ulanbaatar, Greyson Translation Services

安德鲁·韦斯特 (Andrew West)：英国国家代表

吉日木图：日本 Almas 公司

文档

- 陈壮等：N [4880](#)，建议在 UCS 第 5 版中对蒙古文新增的内容；
- 韦斯勒 (Whistler) 等 ([L2/17-328](#))：N [4882](#)，文字特别小组关于蒙古文文本模型的建议 (中英对照)；
- 文字特别小组：N [4889](#)，蒙古文编码的字形模型；
- 文字特别小组：N [4890](#)，关于字形模型的迁移问题 (中英对照)；
- 沈逸磊：N4884，蒙系文字编码中的位置失配
- 郑维喆：N4883，蒙古文区块的混合编码方案