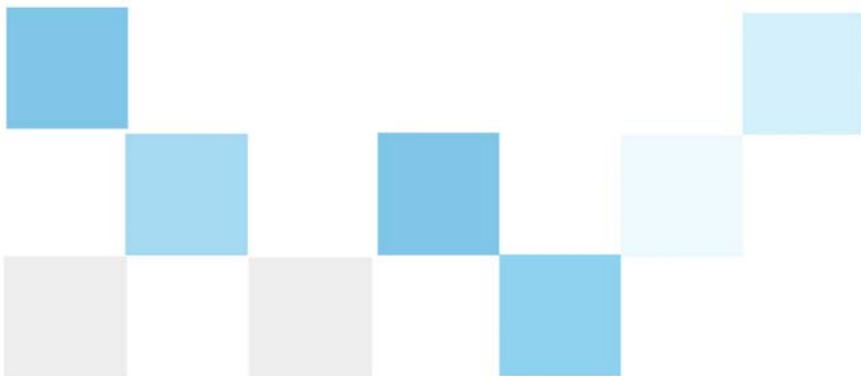




ᠠᠨᠢᠭᠤᠯᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ ᠤᠯᠤᠰ

内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院




科技助力蒙古文信息化发展

TECHNOLOGY BOOSTS MONGOLIAN INFORMATIZATION



实验室简介



内蒙古自治区蒙古文信息处理技术重点实验室于2007年经内蒙古自治区科技厅批准在内蒙古大学挂牌成立。实验室研究团队由内蒙古大学计算机学院智能与蒙古文信息处理团队的教师构成，实验室主任为高光来教授，现有成员17人，其中享受国务院政府特殊津贴专家1人、自治区有突出贡献的中青年专家2人、入选自治区“草原英才”计划2人、自治区“新世纪321人才”一层次2人；团队成员中高级职称人员8人，占团队总数的47%，有博士学位成员14人，占团队总数的82%。以本团队为主要力量组成的“蒙古文软件研究与开发团队”入选2012年度自治区草原英才产业创新人才团队。实验室长期专注于蒙古文智能信息处理研究和蒙古文软件平台开发工作，通过高水平的科研工作培养蒙古文信息处理领域的高层次人才，为少数民族语言信息化的发展提供了人才保障和技术支撑，推动了蒙古文信息处理技术的发展。

实验室于2003年和2007年先后承担了国家863计划课题和国家973计划专项课题，近五年共承担9项国家自然科学基金项目、6项工信部电子发展基金课题、2项自治区自然基金重大项目、20多项自治区横向项目，在国内外代表性期刊和学术会议发表学术论文200多篇，其中100多篇被SCI和EI收录，编写出版了《蒙汉文操作系统及办公软



内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院

内蒙古大学计算机学院

产学结合

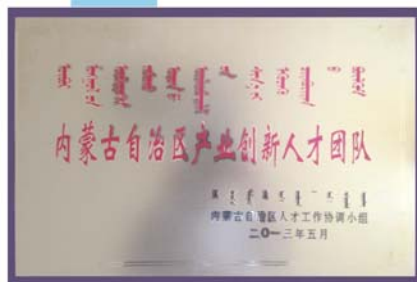
科技兴邦

顶天立地

件应用》教材1部, 获批国家发明专利3项, 获批计算机软件著作权30多项, 实验成果4次获得内蒙古自治区科技进步奖。

实验室经过多年的建设, 在系统平台和软件开发方面取得了很多成果, 自主开发了蒙古文信息检索系统、蒙古语大词汇量连续语音识别系统、蒙古语语音合成系统、蒙古文印刷体识别系统、蒙汉翻译系统、西里尔蒙古文与传统蒙古文相互转换系统、奥云蒙古文整词输入法、蒙汉英多语种翻译电子词典、蒙古文古籍检索系统等。

实验室本着资源共享和优势互补的原则, 与高新技术企业和科研院所积极开展产学研合作, 如与中标软件有限公司合作研发了具有完全自主知识产权且符合蒙古文国家标准的“中标麒麟蒙古文操作系统”和“中标普华蒙古文office”, 软件已上市实现产业化; 与科大讯飞合作建立了联合实验室, 研究蒙古语语音识别、关键词检测和蒙古语语音合成技术, 共同承担了“面向少数民族语言的智能语音技术及系统研发”科研课题; 与工信部电子四院共同承担了“蒙古文标准符合性检测平台”研究工作。



▶ 实验室主任简介

高光来，男，1964 年生，内蒙古扎赉特旗人，内蒙古大学计算机学院教授，博士生导师，内蒙古自治区蒙古文信息处理技术重点实验室主任，内蒙古大学研究生院常务副院长，是内蒙古自治区有突出贡献的中青年专家、享受国务院政府特殊津贴专家，入选自治区“草原英才”计划，担任中国人工智能学会机器感知与虚拟现实专委会委员、中国图形图像学会多媒体技术专委会委员、全国高等学校计算机教育研究会常务理事、教育部高等学校教学指导委员会计算机科学与技术专业分委员会委员、内蒙古计算机学会理事长。

长期从事蒙古文智能信息处理、人工智能和模式识别等方面的研究工作，主要研究包括文字识别、语音识别、图像识别、计算智能和数据挖掘等。主持了国家“863”计划1项、国家 973 计划前期研究专项课题1项、国家自然科学基金项目4项，内蒙古自治区级项目10多项，电子发展基金项目4项。主要科研成果获得内蒙古自治区科技进步二等奖3项、三等奖1项，获得国家发明专利授权2 项。在重要学术期刊和国内外学术会议上发表了60余篇论文，其中 40余篇被 SCI、EI、ISTP 收录。主编或参编的5部教材，获霍英东教育基金会颁发的第6届全国高等学校青年教师奖(教学类)三等奖。





内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院

DIRECTORY

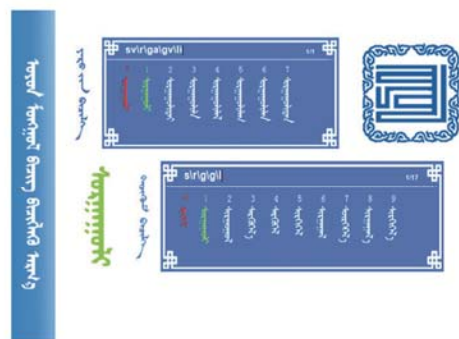
实验室成果目录

- 1 奥云蒙古文输入法
- 2 蒙古语语音识别系统
- 3 蒙古语语音合成系统
- 4 蒙古文网站舆情监控系统
- 5 西里尔蒙古文与传统蒙古文相互转换系统
- 6 蒙古文校正软件
- 7 蒙汉英多语种翻译电子词典
- 8 蒙古文编码转换系统
- 9 蒙古语语言文字资源库
- 10 蒙古文印刷体识别系统
- 11 蒙古文人名转换系统
- 12 蒙古文汉文双向翻译系统
- 13 蒙古文信息检索系统
- 14 蒙古文古籍图像检索系统
- 15 中标麒麟蒙古文操作系统
- 16 中标普华蒙古文Office

实验室成果

■ 奥云蒙古文输入法

- 实验室开发了符合蒙古文国家标准编码的整词输入法——奥云蒙古文输入法。
- 该输入法支持Windows XP、Win7、Win8、Win10等操作系统。
- 具有智能记忆、词库大、字体全、输入快捷、皮肤美观等特点。
- 用户可以免费下载使用，目前为止输入法下载量达到了2万以上。
- 下载网址：<http://oyun.mglip.com>



■ 蒙古语语音识别系统

- 蒙古语语音识别系统的目标是让机器能听懂蒙古语，自动将蒙古语转换成对应的蒙古文文字。
- 实验室研制了国内首款蒙古语大词汇量连续语音识别系统，该系统使用了大规模的蒙古语语音资源库和蒙古文文本语料库，并采用了最新的深度神经网络技术，识别正确率达到了90%以上。
- 蒙古文语音识别平台对外部开发者提供了云服务接口，开发者可以利用云平台接口将蒙古语语音识别功能嵌入到其它应用中。
- 这款系统的成功研制对蒙古文信息处理发展具有重要意义，表明蒙古文信息处理水平已经达到了新的高度，为深入研究蒙古文信息处理技术奠定了基础。
- 应用领域：智能家居、智能玩具、语音输入法、自动转录、舆情监控等领域。





内蒙古自治区蒙古文信息处理技术重点实验室
内蒙古大学计算机学院

产学结合

科技兴邦

顶天立地

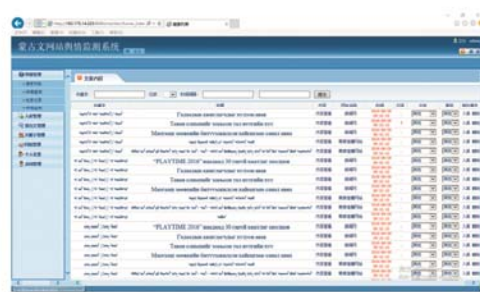
■ 蒙古语语音合成系统

- 实验室采用业界领先的技术开发了蒙古语语音合成系统，实现了将蒙古文文字转化成对应的蒙古语语音，在流畅度、自然度、可懂度等方面达到了较高的水平。
- 系统提供了云服务接口，开发者可以利用云服务接口将蒙古语语音合成功能嵌入到其他应用中。
- 产品演示网址：<http://mtts.mgclip.com>



■ 蒙古文网站舆情监控系统

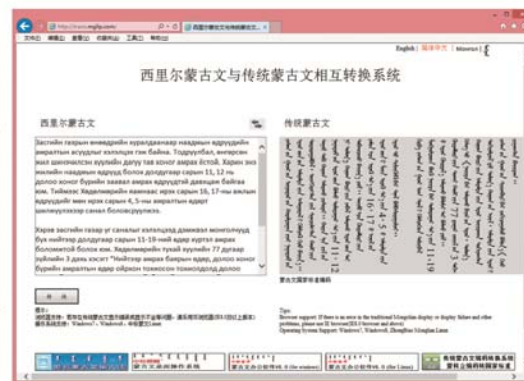
- 蒙古文舆情监控系统利用蒙古文信息检索、编码转换、错误校正、西里尔蒙古文与传统蒙古文相互转换等技术，通过对蒙古文网站信息自动抓取、自动分类、关键词检测等处理，满足用户的网络舆情监测和新闻专题追踪等信息需求，为客户全面掌握群众思想动态、做出正确舆论引导提供分析依据。
- 实时对网络、论坛、微博等进行监测，根据用户个性化需求，为用户定制专业的舆情监测服务平台，将用户关注的行业动态、同行动态、敏感信息等以简报、报告、图表的形式提供给用户，并及时作出预警。



实验室成果

■ 西里尔蒙古文和传统蒙古文相互转换系统

- 传统蒙古文和西里尔蒙古文分别是在中国和蒙古国使用的蒙古文，随着中国和蒙古国两国之间文化、教育和经济的交流与合作不断深入，两国之间的文字转换工作也变的极其重要。
- 实验室采用了基于规则和统计相结合的方法，研发了西里尔蒙古文与传统蒙古文相互转换系统，其中西里尔蒙古文到传统蒙古文的转换正确率达到了95%以上，传统蒙古文到西里尔蒙古文的转换正确率达到了90%以上。目前翻译量达到了50多万条。
- 实验室向广大软件开发者以Web服务的形式开放了西里尔蒙古文与传统蒙古文相互转换系统，开发者可直接调用Web服务进行软件开发，成功案例为“中国民族语文翻译局智能翻译系统”（网址：<http://www.mzywfy.org.cn>）。
- 产品链接：<http://trans.mglip.com>



■ 蒙古文校正软件

- 目前蒙古文文本语料中存在大量错误，包括显现形式正确但编码错误、格后缀使用错误、形同音不同单词的使用错误等，这对蒙古文统计与检索工作带来了很大困难。
- 为校正文本中的错误，实验室采用基于词典、规则和统计相结合的方法开发了蒙古文校对软件，校正准确率达到95%以上。
- 该软件为蒙古文文档编辑、分析、处理工作提供了基础服务，可以应用于新闻出版、教育、翻译等行业。



产学结合

科技兴邦

顶天立地

蒙汉英多语种翻译电子词典

- 实验室开发了面向Android操作系统的蒙汉英多语言翻译电子词典，词典包含7.8万条词条、500多常用对照句，以及词条和语句的语音。
- 每个词条都有对应的传统蒙古文、西里尔蒙古文、汉文、英文四种语言的解释，可以利用这四种语言中的任意一种作为关键字输入，查询到另外三种文字的解释。传统蒙古文按照书写习惯从左向右竖排显示，其它文字从左向右横排显示。
- 开发了基于Android的西里尔蒙古文输入法和传统蒙古文输入法，为词典提供了录入工具。
- 词典对新老蒙古文、汉语和英语的学习和翻译提供了高效快捷的服务，对蒙古文的文化和教育的发展具有重要意义。
- 下载网址：<http://www.mglip.com/Mdic.apk>



蒙古文编码转换系统

- 蒙古文存在多种编码方式，其中蒙古文国家标准是国家支持推广并得到国际编码组织一致认可的编码方式，蒙科立编码是目前非国标编码方式中使用比较广泛的蒙古文编码方式。
- 蒙古文编码转换系统用于实现蒙古文国家标准编码和蒙科立编码的相互转换，该功能有助于加快蒙古文标准编码的普及使用。
- 蒙古文编码转换系统建立在大规模的蒙古文语料、词典和专家制定的转换规则的基础上，转换效果达到非常高的精度。
- 系统网址：<http://mtg.mglip.com>



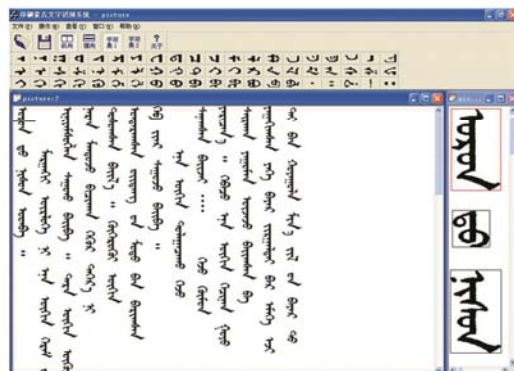
▶ 实验室成果

■ 蒙古语语言文字资源库

- 实验室先后建立了蒙古语语言文字资源库20余个，内容覆盖蒙古文信息处理的多个方面，包括蒙汉对照人名库(16万条)、蒙古文地名库(5万条)、汉蒙词典(121万条，包含30多个领域专有名词)、蒙汉词典(60万条)、传统蒙古文—西里尔蒙古文—汉文—英文对照词典(8万条)、西里尔蒙古文—汉文对照词典(30万条)、传统蒙古文—西里尔蒙古文对照词干库(5万条)、传统蒙古文文本语料库(8千万词)、西里尔蒙古文文本语料库(5千万词)、蒙古文命名实体标注语料(5万句)、蒙古语语音资源库(800小时)、蒙古语语音合成库(10小时)、蒙古文发音词典(7万条)等。
- 蒙古语语言文字资源库的建设对蒙古文信息处理相关技术的研究提供了基础保障，对蒙古文信息化的发展具有重要意义。

■ 蒙古文印刷体识别系统

- 蒙古文印刷体识别是利用模式识别、人工智能、数字图像处理等技术将图片格式中的蒙古文字转换成可编辑文本格式的过程。
- 实验室开发了支持Windows操作平台的蒙古文印刷体识别系统，功能包括倾斜检测与校正、版面分析、文字识别等。
- 蒙古文印刷体识别系统为蒙古文文档数据提供了一种快捷、方便、高效的录入方式。
- 蒙古文印刷体识别系统在字元级识别正确率达到98%以上，在单词级的识别正确率达到90%以上。





内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院

产学结合

科技兴邦

顶天立地

■ 蒙古文人名转换系统

- 蒙古文人名转换系统是将汉文书写的人名转换为蒙古文书写的人名，其中人名分为蒙古族人名、汉族人名和其他人名三大类。
- 实验室建立了大规模的蒙古文人名对照库，并采用基于词典和规则的方法开发了蒙古文人名转换系统。
- 该系统应用广泛，可以应用于医院、机场、车站等场所的人名播报系统和教育与政府部门的证书印制系统等，具有很高的应用价值。

■ 蒙古文汉文双向翻译系统

- 实验室建立了大量的蒙汉对照词典和语料库，并利用基于规则与统计相结合的方法建立了西里尔蒙古文和传统蒙古文翻译成汉文的机器翻译系统。
- 通过该系统可以在不同层次上对蒙古文的内容进行过滤和分析，帮助用户检索含有特定信息的蒙古文文档。通过进一步翻译分析，帮助用户理解相关文档的内容，辅助用户进一步进行翻译。

■ 蒙古文信息检索系统

- 实验室开发的蒙古文信息检索系统是一款针对蒙古文网页内容的搜索引擎。它运用网页爬虫从互联网上搜集蒙古文文字信息，在对信息进行组织和处理后，为用户提供检索服务。
- 蒙古文信息检索系统支持蒙古文国标和蒙古文其他编码标准，能够正确识别网页编码方式，将非国标编码的网页内容转换为国标编码的内容，并建立索引。
- 搜索平台包含网页抓取、网页内容预处理、索引建立和网页排序四大功能模块。

实验室成果

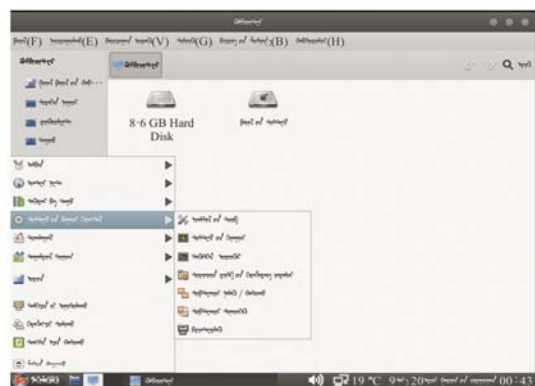
■ 蒙古文古籍图像检索系统

- 实验室利用图像处理、信息检索和模式识别等相关技术和方法，提出了一套适合于蒙古文古籍图像检索的技术路线，即：采用词定位技术，通过图像匹配的方式直接在古籍图像上定位出待检索单词。
蒙古文古籍图像检索系统的实现为蒙古历史研究人员提供了一种查阅蒙古文古籍文献的途径，从而提高了蒙古文古籍的利用效率，对传承和发展蒙古民族文化具有重要意义。
- 以具有代表性的蒙古文《甘珠尔经》为对象，当性能指标为R-Precision时，可达到50%；当性能指标为MAP时，可达49.97%。平均检索时间为10秒，与其他语言的古籍图像检索系统相比，无论检索性能还是效率都取得了较好的结果。



■ 中标麒麟蒙古文操作系统

- 中标麒麟蒙古文桌面操作系统是由内蒙古大学与中标软件针对蒙古文语言文字的特点，以满足蒙古文语言信息处理的用户需求而合作研发的面向桌面应用的操作系统产品。
- 中标麒麟蒙古文桌面操作系统是在中文桌面操作系统研究的成果基础上，落实国家少数民族相关政策，着重开发我国民族语言办公桌面系统，采用全新的图形界面环境，充分考虑使用习惯、文件和数据兼容问题，满足内蒙古地区用户需求，促进内蒙古地区的经济社会发展和信息化建设。
- 产品优势：符合蒙古文字形国家标准；提供方便用户使用的蒙古文输入法；界面支持蒙古文化；实现蒙古文与中英文的混合编辑与处理，形成符合语言特点与排版规范的处理规则。
- 产品网址：<http://www.mglip.com/profro02.htm>





内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院

内蒙古大学计算机学院

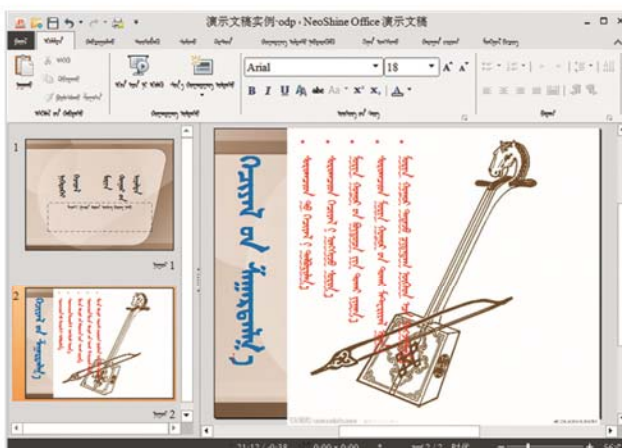
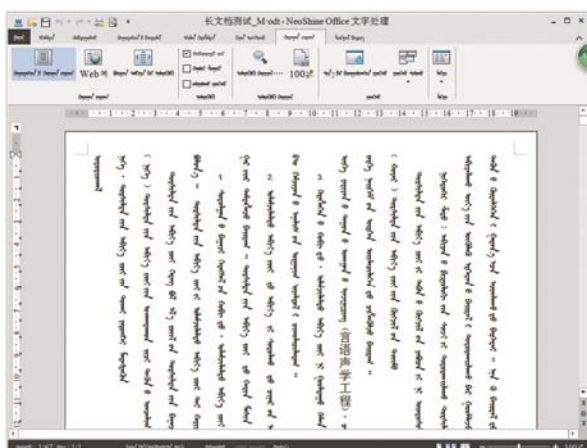
产学结合

科技兴邦

顶天立地

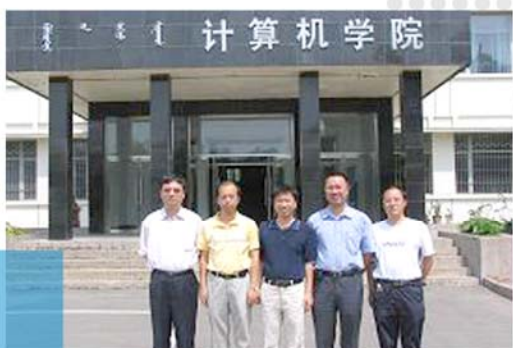
■ 中标普华蒙古文Office

- 中标普华蒙古文办公软件是由内蒙古大学和中标软件联合研制的一款针对蒙古文语言的办公软件产品，产品包含文字处理、电子表格、演示文稿三大组件，满足蒙古文办公和文档处理的需求，是国家工信部电子信息发展基金“蒙古文多语种跨平台办公套件”的重要项目，得到了工信部、内蒙古民委、上海市经信委、内蒙古经信委等单位的大力支持与认可。
- 产品特点：(1) 中标普华蒙古文办公软件符合蒙古文国家、国际字符编码标准；实现了蒙古文的从左到右竖排排版方式，支持蒙汉英多文种的混合编辑排版，符合各文种的排版规则和习惯；提供蒙古文和汉语两种软件界面，可以快捷的实现双向切换。(2) 中标普华蒙古文办公软件实现了蒙古文日期、蒙古文时间、蒙古文编号等蒙古文本地化需求；提供标准二次开发接口，可实现与各种电子政务系统以及企业信息化系统的无缝集成，实现蒙古文办公自动化。
- 下载网址：<http://www.mglip.com/profro01.htm>



合作交流

COOPERATION AND EXCHANGE





内蒙古自治区蒙古文信息处理技术重点实验室

内蒙古大学计算机学院

内蒙古大学计算机学院

越是民族的，就越是世界的

ᠠᠨᠠᠭᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ
ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ ᠤᠯᠤᠰᠤᠨ





内蒙古自治区蒙古文信息处理技术重点实验室

联系人: 高光来 教授, 飞龙 副教授

地 址: 内蒙古呼和浩特市大学西路235号内蒙古大学计算机学院

邮 编: 010021

电 话: 13947124377

网 址: <http://www.mglip.com>

E-mail: csggl@imu.edu.cn, csfeilong@imu.edu.cn