

黄建年◎著

天華觅踪

—
古籍整理新论

Tianzhang Mizong
Guji Zhengli Xinlun



安徽师范大学出版社

黄建年◎著

天華賈踪

—
古籍整理新论

Tianzhang Mizong
Guji Zhengli Xinlun



安徽师范大学出版社

图书在版编目 (CIP) 数据

天章觅踪：古籍整理新论/黄建年著. —芜湖：安徽师范大学出版社，2011.12

ISBN 978-7-81141-679-4

I. ①天… II. ①黄… III. ①古籍整理—研究 IV. ①G256.1

中国版本图书馆 CIP 数据核字 (2011) 第 259383 号

天 章 觅 踪
——古籍整理新论
黄建年 著

出 版 人：张传开

责任编辑：潘 安

装帧设计：丁奕奕

出版发行：安徽师范大学出版社

芜湖市九华南路 189 号安徽师范大学花津校区 邮政编码：241002

发 行 部：0553-3883578 5910327 5910310 (传真) E-mail: asdcbsfxb@126.com

经 销：全国新华书店

印 刷：安徽芜湖新华印务有限责任公司

版 次：2011 年 12 月第 1 版

印 次：2011 年 12 月第 1 次印刷

规 格：787×1092 1/16

印 张：16.5

字 数：349 千

书 号：ISBN 978-7-81141-679-4

定 价：34.00 元

凡安徽师范大学出版社版图书有缺漏页、残破等质量问题，本社负责调换



序 —

黄建年 1990 年毕业于北京大学图书馆学系，毕业后长期从事图书馆技术工作。1999 年进入南京大学信息管理系攻读硕士学位，本人因为担任老师的缘故，与其多有接触，最后又直接指导其学位论文的写作，对其善用时间、努力学习、不畏艰难的精神留下了深刻印象。2006 年黄建年就学于侯汉清教授门下，攻读博士学位，视野遂逐渐转向文化典籍的整理与研究。学习期间成果丰硕，时见其文载于报刊。2009 年获得博士学位后，就任南京财经大学物资设备处副处长之职。虽然工作繁杂，但黄建年利用工作之暇，笔耕不辍、一门深入，撰写出古籍整理方面的论文 20 余篇。《天章觅踪——古籍整理新论》作为黄建年个人古籍整理相关论文的汇集，不仅记载了黄建年不断学习、探索古籍整理理论与技术的足迹，而且映射出当代图书馆学人对图书情报事业的无限热爱与挚诚。毫无疑问，该部论文集的出版，值得业界同行关注与祝贺！

该书的主题主要集中于古籍的整理与研究，但其研究内容却丰富多彩，从传统古籍整理方法到现代智能整理技术均有涉及。传统研究包括古籍标点、古籍索引等历史与现状的研究，现代技术则包括古籍智能整理与开发系统的构建、断句标点模式的研制、分词技术的应用、索引编制实验等内容。寺院文化是本书的另一个着力点，作者依据大量的古代典籍，研究了寺院藏书、土地、劳动力、救荒、科技等诸多方面的特点与成就。而这些成果的取得，离不开国家与政府对古籍整理工作的支持与赞助，离不开国家和政府对传统文化的继承与弘扬。这些论文的写作，既是申请支持与赞助的基础，同时也是相关基金项目的重要支撑。文集的最后一部分——“课题报告”专门搜集了 4 个有关古籍整理的基金项目申请报告，有助于全面了解作者关注古籍整理研究的心路历程，对于课题申报亦有一定的指导作用。

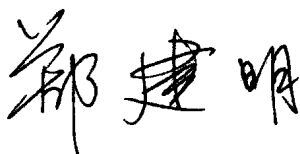
家谱族史是中华文明传承的重要记载，黄建年长期关注这一领域，并直接投身于家谱族史资料的搜集、整理与编撰。《圖阳当沙陶氏司分江都支谱序》即是其中的代表之作，详细叙述了圖阳当沙陶氏司分江都支族人世系传承之大略，不仅有功于陶氏族人慎终追远、返本还源，亦且有功于我炎黄子孙家族文化的传承与发展。

当然，该书毕竟只是一部个人文集，多个主题的划分在一定程度上或许淡化了研究的深度。比如《中国古代佛教寺院藏经楼研究》提出了中国古代佛教寺院藏经楼“已经具有现代公共图书馆的雏形”这一论点，很有见地，可惜未能进一步细化。对于古代佛教寺院藏经楼的开放性以及导读方法也作了较多的陈述，但尚缺少与现代图书馆的比较以及导致古代佛教寺院藏经楼衰落原因的进一步分析。衷心期待黄建年在未来的研究中能够进一步完善，不断有新的成果面世！

黄建年虽承担部门领导工作，事务众多，但在学术上仍孜孜以求、兢兢业业，成果斐然，且多次来电请我作序，余推辞再三，然不获其允，因而草此数语，聊以充之。

是为序。

南京大学信息管理系教授
博士生导师



2011 年 10 月 19 日



序 二

中国文化典籍是中华民族在数千年历史发展过程中创造的重要文明成果，蕴含着中华民族特有的精神价值、思维方式和想象力、创造力，是中华文明绵延数千年的历史见证，也是人类文明的瑰宝。对古籍的整理、保护与开发是炎黄子孙应尽的义务和职责。

中国的古籍整理，有着悠久的历史 and 传统。从孔子删定《诗经》、刘向父子校书，到清人编定《四库全书》、《古今图书集成》、《书目答问》，大规模的古籍整理工作持续不断，影响深远。新中国成立后，党和政府就已经注意到古籍整理这项工作的开展。现代信息技术的应用给古籍整理带来了巨大的动力，使古籍整理得以昂首阔步向前飞速发展。2007年年初《国务院办公厅关于进一步加强古籍保护工作的意见》颁布实施，国家实施的古籍保护计划进一步推动了各地区古籍整理与保护工作的开展。在国家古籍保护中心的组织下，第一批《国家珍贵古籍名录》和全国古籍重点保护单位的申报工作已经基本完成，同时积极制定古籍数字化标准，古籍整理与保护工作进入了一个新的历史阶段。

大量古籍整理论著的陆续面世，使我们对古籍以及古籍整理工作有了更加全面与深入的了解。古籍整理，有广义与狭义之分。现在一般人把有关古籍各方面的学术工作都统称之为古籍整理，这是广义的古籍整理。从这个意义上来说，它研究的范围包括古文献的种类、源流、形制、版本、功用、断代、辨伪、编目、整理、收藏、传播等。其中很多方面又可以形成独立的学问或分支学科，例如研究古书形制发展的“书史”，研究版本的古籍版本学，研究古书编目的古籍目录学，研究古书断代与辨伪的文献断代辨伪学，等等。这样的定义显然容易与古文献学相混。但实际上，古籍整理乃是着重研究古文献学中的“整理”这个方面，它

是古文献学的一个重要组成部分，二者还是有着相对明确的泾渭分野。

当然，严格意义上的古籍整理，刘琳、吴洪泽在其专著《古籍整理学》认为“是对古籍的原文进行某种形式的整理加工，以便于人们阅读与研究。比如校勘以厘正文字，标点以分清句读，注释以阐明文义，翻译以通连古今，辑佚以摭拾遗文，抄纂以采其菁华等等”。换句话说，古籍校勘、标点、注释、翻译、辑佚、编纂才是古籍整理的主要内容。

黄永年在《古籍整理概论》中认为，古籍整理“是对原有的古籍做种种加工，而这些加工的目的是使古籍更便于今人以及后人阅读利用，这就是古籍整理的涵义，或者可以说是古籍整理的领域”。这样的定义显然较为宽泛，但是“撰写讲述某种古籍的论文，以及撰写对于某种古籍的研究专著，尽管学术价值很高，也不算古籍整理而只能算古籍研究”，明确将古籍研究与古籍整理分开。为了进一步说明古籍整理的领域，黄氏还列举了10种基本工作内容，即选择底本、影印、校勘、辑佚、标点、注释、今译、索引、序跋、附录，所列内容要比刘琳、吴洪泽二氏所列要广泛一些。

来新夏在《古籍整理散论》和《古籍整理讲义》中，共论及分类、目录、版本、句读、工具书、校勘、考据、传注等八个方面的内容。时永乐先生的《古籍整理教程》分为古籍版本、校勘、标点、注释、辨伪与辑佚、古籍整理的其他工作（撰写序跋、编辑附录、编制索引）等8个部分进行论述。

程毅中在《古籍整理浅谈》则认为，古籍整理“除校勘文字之外，还包括了编纂（如新编的《全宋词》）、标点、注释、今译和辑佚（如果原书有佚文的话），还有写序跋、编索引、选择影印底本等任务”。

综上所述，古籍整理包括的具体内容，有关专家的观点也并不完全一致，但是，对于核心部分还是有着共同的看法。比如校勘、标点（其中1家标为句读）、辑佚、注译为上述5家共同主张，索引为4家共同主张（《古籍整理散论》工具书部分包括索引），可见校勘、标点、辑佚、

注译以及索引才是古籍整理的核心部分。采信这样的说法，并不是将辨伪、辑佚、序跋排除在古籍整理工作之外，只是并不作为古籍整理的重点。在本书中，古籍整理即采用这样的定义。

笔者长期留意古籍整理工作。本书的第一编、第二编均是这些成果的部分展示。它们陈述了古籍标点的历史与现状、古籍智能整理与开发的技术与方法。因为论述古籍索引的内容众多，所以单独列出，以搜罗专门研究古籍索引的相关历史、软件以及编制多文本索引编制的技术与方法。佛教文化博大精深，佛教典籍也内容丰富、卷册众多，是笔者同样关注的另一领域，从寺庙藏经楼、土地、劳动力、救荒到科技问题均有所涉猎。第三编“寺院文化”正是相关研究的集中展现。家谱族史是中华文明历史悠久、绵延不绝的记录证明，也是炎黄子孙慎终追远、报本还源文化与思想的传承工具。笔者有缘参加了黄氏家谱、陶氏家谱、龚氏家谱的编纂工作，也有幸参加了江苏省黄氏文化研究会的筹建活动，因而对家谱族史的意义有着更加深入的了解和发自自己五内的一种热爱。第四编“家谱族史”从中萃取部分以飨读者，希望能够引起炎黄子孙对生命的关注。第五编为各类项目的课题报告。近年来，国家和各级政府加大了对古籍整理研究工作的支持力度，设立了各种科研基金项目。笔者参与了几个项目的申请，并侥幸获得了相关资助，其申请报告或有益于同仁，或有助于社会，因而不揣浅陋，公示于众，以便就教于方家。

本书能够顺利成书，实离不开国家对古籍整理工作的支持以及有关基金项目的赞助，在此表示衷心的感谢！

中国的传世古籍浩如烟海，仅儒家经典，即非一人一世之力所可穷尽，遑论皇皇巨著之《道藏》、《佛教大藏经》。面对积屋盈栋之书，古今贤圣虽博采众长，成果卓著，然时有管窥蠡测之叹。庄子所谓“吾生也有涯，而知也无涯”，恰恰道出所有士子心声，因而引起千古共鸣。笔者草撰此稿，亦有庄子之慨。

“文章本天成，妙手偶得之。”传世古籍、经典文章，本上合天真、

· 4 · 天章 觅 珠

妙手偶得，实皆天章。宋真宗始建天章之阁，以奉御集御书。南宋复建于临安，以藏图籍、符瑞之物。自此，天章之阁遂成古代藏书之表征。天章所藏图书、符瑞、灵文，非王公大臣不能见，亦非有大智者不能信之、不能解之、不能行之、不能证之，吾人知识浅陋，虽难尽其妙，但汲汲以求、持之以恒，非不能觅其古圣先贤踪迹、寻其大道至理端绪。或许偶有一得，成之于文，以期嘉惠后来云耳。

是为序。



草于 2011 年菊月吉日



目 录

序 一	郑建明
序 二	黄建年

第一编 古籍整理

1 中国汉语古籍标点历史与现状	1
2 中国农业古籍标点历史与现状	25
3 古籍智能整理与开发系统构建研究	50
4 农业古籍断句标点模式研究	59
5 中国古籍文本分词的一次试验	69

第二编 古籍索引

1 汉文古籍索引自动化实践与研究概述	81
2 农业古籍索引编制历史与现状	90
3 基于 GB/T 22466—2008 的古籍索引编制技术要点	110
4 中国古籍索引编制软件概述	121
5 基于 VFP + WORD 的多文本古籍索引编制实验	128

第三编 寺院文化

1 中国古代佛教寺院藏经楼研究	136
2 中国古代佛教寺院藏经楼开放性研究	145
3 中国古代佛教寺院导读活动研究	157
4 中国古代佛教寺院土地、劳动力简述	167
5 中国古代佛教寺院救荒研究	180
6 中国古代佛教寺院科技概说	188

第四编 家谱族史

1 重修太昌黄氏宗谱序	210
2 贺重修太昌黄氏宗谱、祠堂序	212
3 贺重修太昌黄氏祠堂对联	213
4 太昌黄氏新增世次韵语	214
5 圖阳当沙陶氏司分江都支谱序	215

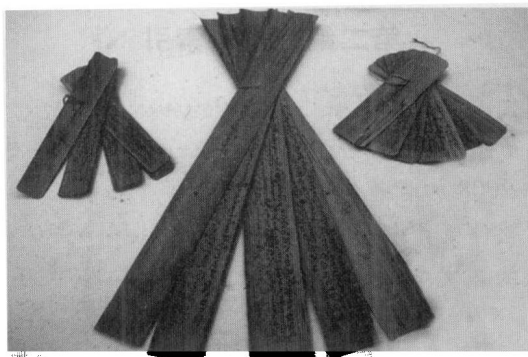
6 圖阳当沙陶氏司分江都支谱增补字辈·····	217
-------------------------	-----

第五编 课题报告

1 文化典籍整理与开发的智能技术研究·····	218
2 农业古籍自动分词及索引编制技术研究·····	223
3 中国地方志分类门目研究·····	231
4 佛教寺庙志分类体系研究·····	238

索 引 ·····	246
-----------	-----

后 记 ·····	251
-----------	-----



第一编 古籍整理

1 中国汉语古籍标点历史与现状

1.1 引言

所谓古籍的断句标点，就是给没有标点的古代汉语书籍断句并加上适当的标点符号的过程。尽管我国古代已经形成了若干种句读符号体系，但是，古代句读的使用在古籍中并不普及。胡适先生曾说，“古籍率无圈点，即有之矣，其所用符号，又不完备；或有圈而无点，有句而无读。其圈点又不依文法构造，但截长为短，以便口齿而已。”^{[1]281}虽然并不完全如此，但基本说出了 20 世纪以前中国古籍标点之大致情形。

1.2 旧式标点标注古籍

我国古代典籍汗牛充栋，有人认为有八万余种（一说十五万种），直到清末，绝大多数的古书仍然不加句读。即使到现在，经过标点整理的古书也不过四五千种而已。^{[2]2}可见，大多数古籍并没有加注新式标点符号。

1920 年，汪原放标点、分段并由上海亚东图书馆排印出版了《水浒传》，这是中国首次使用新式标点符号标点、出版古典书籍。^[3]

需要引起注意的是，尽管汪氏于古籍标点功勋卓著，但是，汪氏标点、出版之书在其版权页并没有采用“标点者”，而是“句读者”。可见，20 世纪 20 年代初“句读”一词的使用频率仍然远远高于“标点”。

汪氏的贡献主要在于利用新式标点符号标点中国古代小说，而由官方大规模利用旧式标点标注古籍却由来已久。

清代钱谦益认为，南宋年间刊刻的《淳熙九经》“点断句读皆精审”，可见南宋时官方已经对儒家经典加注了句读，带有句读的《淳熙九经》通行于全国，当时全国学子应该获益匪浅。然句读仅止于儒家经典，经外图书之句读加注仍然付阙。大规模由官方对经外之书加注句读可以追溯到明代《永乐大典》的编纂。

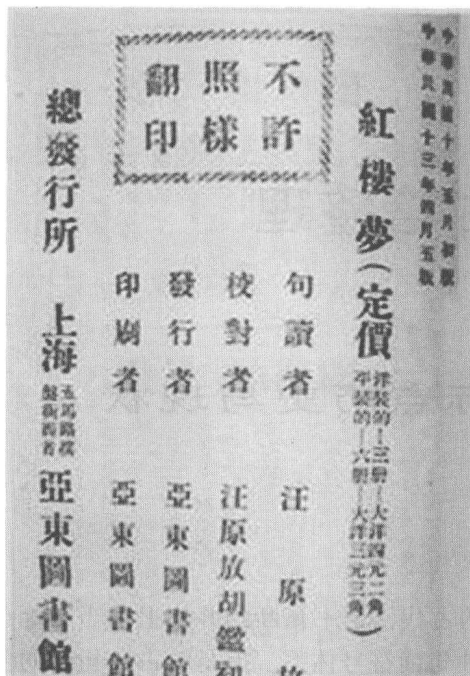


图 1-1 汪原放标点本《红楼梦》版权页

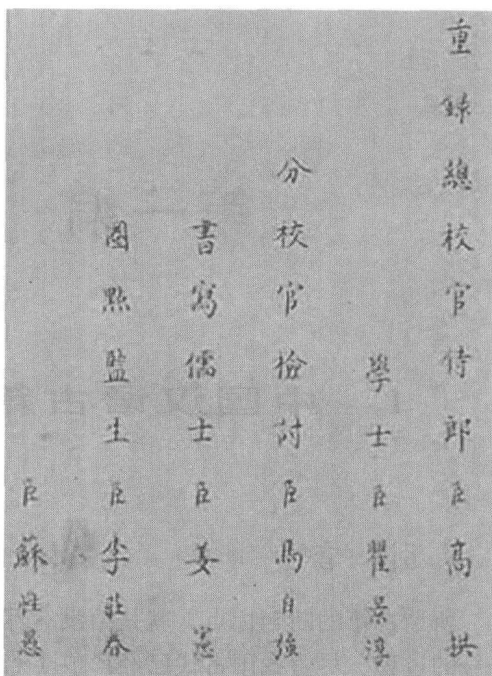


图 1-2 《永乐大典》卷末版权页

《永乐大典》正文 22 877 卷，凡例和目录 60 卷，装成 11 095 册，总字数约 3.7 亿字。书中保存了我国上自先秦的佚文秘典，下迄明初的各种典籍资料达 8 000 余种，是中国古代最大的类书。尽管该书规模巨大，但难能可贵的是，却全部加注了句读符号。该书“书名用硃书，句读用硃圈，工整匀洁”^{[4]序3}。“正文与注文都用小白圈（。）断句，置于字的右下侧，字距不变，小字长注文往往用空格的办法来分段。”^[5]

目前能够影印、出版的《永乐大典》各残卷均有句读人员的记载，这是非常难得的。比如日本东京大学图书馆藏本《永乐大典》910—912 卷末注明“圈点监生臣李庄春臣苏性愚”，说明此册的圈点工作是由李庄春、苏性愚二人共同完成的。而且圈点人员与总校官、分校官、书写人同时并列于卷末，说明圈点人具有一定的社会地位。

当然，明代加注句读的作品尚有嘉靖本《三国志通俗演义》、万历年间容与堂版《水浒传》等，这些古籍均实行全面句读，而大多数古籍刻本句读时有时无。

可惜的是，清代官方在印制或者抄录古籍时未能延续这一传统。比如清乾隆时官方编纂的《四库全书》为目前存世的最大一部古代丛书，共收录千余种图书。然与《永乐大典》加注标句读的做法相反，无论收入之书原来是否有断句标点，一经编入即用统一的体例不予断句标点。举总集为例，如宋吕祖谦《古文关键》、谢枋得《文章轨范》、王霆震《古文集成前集》、元方回《瀛奎律髓》、赖良《大雅集》等原本皆有著名的“批点”、“评点”，但收入《四库全书》后，其“圈点”均消失

无踪。^{[1]218}

尽管官方对于古籍句读重视不足，但于一般学者却多有提及，比如桐城派领袖方苞曾给皇帝上书，奏请重刻《十三经》、《廿一史》，“务期句读分明，使学者开卷了然”。他亲身实践，评点了不少经、史、子、集类著作，如《评点大戴记》、《左氏评点》、《史记评点》等数十部典籍。桐城派刘大櫆、姚鼐评点之书亦复不少，刘氏评点著作有《评点国语》、《评点庄子》、《评点扬子法言》、《评点诗经》、《评点楚词》、《评点文选》、《评点古诗选》等，而姚氏评点著作亦有《评点易经》、《评点毛诗故训传》、《评点周礼》、《评点礼记》、《评点左传》、《评点秦板九经》、《汉书评点》等。^{[1]218-238}

当然，清代学者并非对前代图书加注句读一律持赞同态度。比如清中兴名将曾国藩在其《经史百家简编序》中说道：^{[1]225}

后人不察，辄仿其法，以涂抹古书，大圈密点，狼藉行间。故章句者，古人治经之盛业也，而今专以施之时文；圈点者，科场时文之陋习也，而今反以施之古书。末流之迁变，何可胜道？

他认为圈点为“科场时文之陋习”，明确持否定态度。但是，曾氏并非全盘反对圈点，他只是反对“涂抹古书，大圈密点，狼藉行间”。实际上他自己做了大量的评点工作。他的评点著作就有《评点左传》、《评点孟子》、《评点文选》和《十八家诗钞》、《经史百家杂钞》等。他的评点对后人影响较大。

曾氏门下弟子俞樾也认为圈点、评语“皆明以来陋习”，所持观点与其师同出一辙，故此不需细论。

总之，有清一代，标注句读之书数量亦非常有限，尤其《四库全书》取消原文标注，对于清代图书刊刻影响甚巨。

1.3 新式标点标注古籍

新式标点标注古籍始于1919年以后。1920年，在陈独秀、胡适等人的支持下，汪原放标点、分段并由上海亚东图书馆排印出版了《水浒传》，这是中国第一次使用标点符号出版古典书籍。^[31]关于这一部古典小说的标点并出版，得到了极高的评价。当时胡适说：“我可预料汪君这部书将来一定要成为新式标点符号的实用教本，他在教育上的效能一定比教育部颁行的新式标点符号原案还要大得多。”实际上，汪原放先生标点的古代典籍并不止这一部，另外尚有《儒林外史》、《红楼梦》、《西游记》、《三国演义》等数十部。汪氏所有标点、出版图书其前有“句读符号说明”，向读者介绍本书所用标点符号的类型及其使用方法，并用本书中的原句作例证，对于新式标点符号的推广、教育的普及有不可估量的作用。其后均有“校读后记”，详细说明版本的选择、文字的增删补改等校勘的过程，为不可多得的学术论文。鲁迅先生曾说过：“标点只能让汪原放，做序只能推胡适之，出版只能由亚东图书馆，

刘半农、李小峰、我，皆非其选也。”即是对汪氏标点工作的赞赏。《光明日报》发文称汪氏为“标点校勘古典小说第一人”，良有以也。^[6]

中国古籍流落于世界各国，因而中国古籍整理也为世界很多国家所重视。不仅中国重视古籍标点，其他各国亦然。下文只对中国古籍整理加以介绍。

1.4 中国

1920年—1949年，古籍整理工作以影印古籍或者提供校勘精良的版本为主，取得了不少成绩。当然，其间也颇有些出版机构对古籍重新进行排印并加注标点。这些出版机构主要有中华书局、商务印书馆、世界书局、大东书局、开明书店等。

1.4.1 丛书

商务印书馆于1922年起陆续开始印售《四部丛刊》，先后出版了初编、续编、三编，共收录古籍502种，分装成3100多册，规模巨大，可以说是20世纪我国新出的规模最大的丛书。该书全部采用石印法影印，因而完整地保存了古籍原貌。也正因为如此，全书没有加注标点。1992年，北京书同文数字化技术有限公司根据北京大学图书馆所藏涵芬楼影印版《四部丛刊》制作了电子版，功能大大增强。

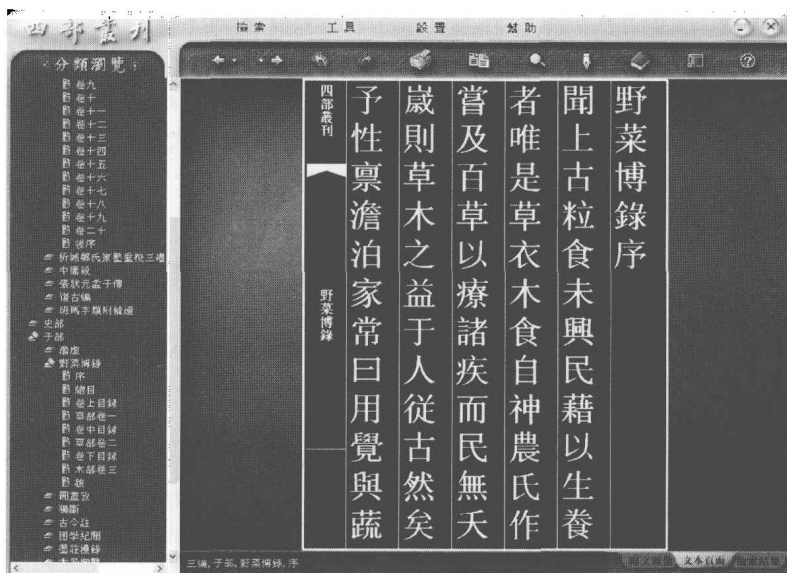


图 1-3 《四部丛刊》电子版界面

《万有文库》是商务印书馆于1929年—1935年出版的另一部大型丛书，共收录图书1721种，分为2集4000册，其中，《国学基本丛书》初集100种，《国学基本丛书》二集300种，《学生国学丛书》60种。另附大本参考书《十通》、《佩文韵府》共11种28册。这些图书基本上采用影印方法，因而标点符号付之阙如。另外，该套丛书中有相当数量的图书并不是“古籍”。比如《汉译世界名著》、《自然科学小丛书》等所收各书均为当时学者著译之作。主编王云五先生亲自草拟了《万有文

库》第一集目录——《千种丛刊目录》。通过它，我们可以了解《万有文库》的大致情形。

为了与商务印书馆抗衡，中华书局于1924年开始，至1931年，陆续编纂出版了《四部备要》，先后共出五集，收书336种11305卷，分装为2500册。该书初期出版与《四部丛刊》相同，没有标点。为了弥补这一缺陷，1935年，中华书局又重新印行洋装16开点句本，使该书进一步流通。但洋装点句本只收录初版部分内容，共收图书126种，加点句读，分订119册。因为收录图书数量有限，因而不能替代初版。需要说明的是，该书重印点句本采用的仍然是旧式句读符号，而非新式标点符号。

与此同时，中华书局于1935年—1937年陆续编辑出版了《丛书集成初编》，该书原计划收录宋、元、明、清四代的丛书100部，分订为4000册，因抗日战争爆发，出书中断，实际出版3467册。《丛书集成》有排印和影印之分，如是影印，基本上无句读；如是排印本，则有句读。此套丛书虽为“近代出版史上尤其是丛书史上的经典之作”，“唯断句、排校不无讹误”。^[7]

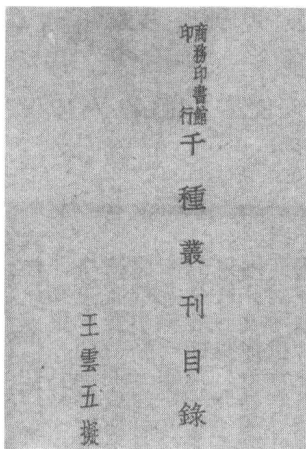


图1-4 《万有文库》第一集目录页



图1-5 上海大东书局《四库全书总目》书脊

1935年，世界书局编辑出版的《诸子集成》，辑录先秦到南北朝儒家、道家等诸子代表作28种。这套丛书为排印本，惜其排校不够精严。针对世界书局版《诸子集成》中“断句、排印尚有不少错误”，中华书局决定自1982年起重新编辑出版新版《诸子集成》丛书——《新编诸子集成》。

《新编诸子集成》收录子书之精注精校本，尤重集解本，以清人及近人注本为主，反映了诸子研究的最新成果，且“各书正文、注文一律加以新式标点”^[8]，因而后出转精。相对于世界书局版《诸子集成》来说，新编本标点校勘均较世界书局本为良。

大东书局创办于1916年，由吕子泉、王幼堂、沈骏声和王均卿等人在上海创办，曾经出版《四库全书总目》、《曲海总目提要》、《中国医学大成总目提要》等

