

文章编号:1005-3085(2005)08-0134-05

它山之石, 可以攻玉

—美国大学生建模竞赛之评审与评述*

胡代强

(暨南大学数学系, 广州 510632)

摘 要: 本文介绍美国大学生数学建模竞赛(MCM)和交叉学科建模竞赛(ICM)的评审办法, 总结美国赛历年优秀论文的评述, 对比美国赛与全国赛的异同, 指出美国赛中可供借鉴的方方面面。

关键词: 美国大学生数学建模竞赛; 交叉学科建模竞赛; 评审

分类号: AMS(2000) 00A71; 97C30

中图分类号: O424.31

文献标识码: A

1 引言

美国大学生建模竞赛由数学建模竞赛(MCM)和交叉学科建模竞赛(ICM)两部分组成。1985年举行了第一届 MCM, 赛题是来源于实际的应用数学问题, 竞赛每年举行一次, 迄今为止已经举行了21届。MCM 是为激励和改善学生的数学方面的问题解决和写作技能而设计的, 该项竞赛的一般介绍参见[1]。ICM 是从1999年开始的, 每年与 MCM 同时举行, 主要是为了发展和提高参赛者的交叉学科问题解决能力和通讯写作能力。

我国大学生最早是在1989年开始参加该项比赛(根据叶其孝教授与 MCM 首任竞赛委员会主任 Ben Fusaro 教授的商定), 第一次参赛时只有北京大学、清华大学、北京理工大学的共4支队伍, 此后参赛队数逐年增加, 到2005年两项竞赛的非美国参赛队数已大大超过本土参加队数, 而其中又多数是中国大学生。

为了有更多的人参加该项竞赛并从中获益, 培养出更多的掌握运用数学建模方法分析和解决问题技能的人才, 我们收集了近二十年来美国赛的大量特等奖论文及其评述文章并加以总结, 结合我国数学建模竞赛的现状对美国赛的评审与评述进行一些描述。

2 美国参赛论文的评审

现任竞赛组委会主任 Frank Giordano 和 Chris Arney 每年都会指出, 参赛论文在美国数学及其应用协会(COMAP)总部编号以避免作者名字和隶属关系被评委知道。评审分成初评(triage judge)和终评(final judge)两步。初评的前两轮由两个初评评委分别完成, 评分的依据主要是摘要的质量与内容的总体组织(2004年的 ICM 还首次强调会考虑模型的描述)。终评将由另外一批专家完成, 最后将论文分成特等奖(Outstanding)、

收稿日期: 2005-11-22. 作者简介: 胡代强(1968年2月生), 男, 博士, 副教授, 研究方向: 数学建模、管理系统仿真、图论。

*基金项目: 广东省自然科学基金(031904).

一等奖 (Meritorious)、二等奖 (Honorable Mention)、成功参赛奖 (Successful Participant) 和未获奖共五类。从1996年到2005年 MCM 获特等奖、一等奖和二等奖的比例平均为1.88%、14.33%和27.73%，总计为43.94%。从1999年到2005年 ICM 获特等奖、一等奖和二等奖的比例平均为3.02%、15.88%和35.03%，总计为53.93%且波动较大。

从历年 MCM 和 ICM 的竞赛组委会主任公布竞赛结果的文章和2003年评委 William P. Fox 的评述文章^[2]可以看出：虽然经过了八年之久，MCM 的评审方法却并无大的改变。从历年 ICM 组委会主任公布竞赛结果的文章大致可以看出，ICM 也采用了相似的评审方法，只是获奖比例有所不同而已。因此文献^[3,4]中关于评审办法的描述仍然非常有借鉴意义。下面根据1996年 Frank Giordano 的评述文章^[3]和 Steve Harper 的评述文章^[4]，对 MCM 的评审情况进行具体的描述。

评审的初评和终评分为三个阶段：浏览 (screening)、评分 (grading) 和评定 (judging)。浏览阶段一共分为三轮，均采用7分制 (1分最差，7分最好)，前两轮由初评评委完成。评分阶段共分为至少四轮，采用百分制，评分阶段与第三轮浏览均由终评评委完成。评定阶段是由各评委讨论决定特等奖的最终获得者。

浏览阶段的根本目的是确定那些不能进入最后的43% (即不能获二等以上奖励) 的论文。初评评委在初步阅读几篇论文后一起讨论制定一个7分制的评分标准，然后据此对论文进行评分，并据此将论文分成三类：质量优秀的、保留并继续评审的、去掉不再评审的。每篇论文由两个评委评审，当他们的评分差距超过两分时就需要进行讨论，如果两位评委不能达成一致，则安排第三个评委进行评审以决定该论文的去留。每个评委浏览一篇论文的时间约为10-15分钟。经过前两轮的浏览后约有25%论文被排除在外。

第三轮浏览与评分阶段由另外一批评委完成。组委会首先对前两轮的评分进行标准化，然后根据总分对留下的论文进行排序并据此对论文进行分组，以保证每个评委刚好对应一组论文，并且每组论文中均有较好的论文和较差的论文，这是随机分配论文所无法达到的效果。评委首先阅读这些论文，但不评分，然后讨论制定一个7分制的评分标准并据此进行评审以完成第三轮的浏览。第三轮浏览结束后约有43%的论文保留下来，保留下来的论文可获得二等以上奖励，被淘汰的论文多数获成功参赛奖，极少数不符合竞赛规则的将不能获奖。

三轮浏览结束后，评委们再讨论制定一个百分制的评分标准，并进行至少四轮评分。每个评委用在一篇论文评分上的时间约为30-45分钟。除了要求评委给出指定论文的一个百分制评分外，还要求对所评论进行排序，比如排序为“2/7”表示该论文在所评的7篇论文中是第二好的。另外还要求评委就每篇论文给出独立于其分数、排序之外的获奖等级。

四轮评分结束后进入评定阶段，每个评委再仔细阅读保留下来的论文中其未评审过的部分，然后评委们一起讨论每篇论文的优缺点并最终根据多数人的意见选出获特等奖的论文。

从第三轮浏览开始，评委会主任、副主任和两个题目的评委组长会一起讨论确定什么样的论文无需再进行评审，他们主要根据标准化的分数的总排名、每轮的轮排序和评委给出的论文获奖等级做出决定。如果某论文评分较低但排序非常靠前则仍有可能继续保留下来以参加下一轮的评审。每轮淘汰的论文并无限额，但第三轮浏览结束后通常有43%的论文会保留下来，这部分论文将至少获得二等奖，而第三轮评分结束后通常会有18%的论文被保留下来，这部分论文将获得一等奖或特等奖。

虽然所选的评委的公正和诚实受到认同，但人类的偏好是不可避免的，因此组委会通过一些措施在技术上尽量减少这些偏好对论文评审的不利影响，以便对论文进行公正的评价。

措施之一是保证每个评委都能看到优秀的、好的和一般水平的论文，这无法通过随机分配论文的方式来实现。组委会通过之前各轮的评分对论文进行排序，然后根据此排序分配论文给

每个评委。

措施之二是为了避免部分评委偏好给高分和偏好给低分的习惯,对每个评委根据其评分习惯不断修正其权数。一个由 MCM 的首位主任 Fusaro 传给现任主任 Frank Giordano 的加权公式如下:

$$\text{加权评分} = \begin{cases} \frac{\text{总平均分}}{\text{该评委平均分}} \times \text{原始评分}, & \text{若总平均分} \leq \text{该评委平均分}, \\ \frac{100 - \text{总平均分}}{100 - \text{该评委平均分}} \times (\text{原始评分} - 100) + 100, & \text{否则}. \end{cases}$$

该公式意在使各评委的平均分与总平均分保持一致。在计算权数时只考虑目前仍保留下来的论文的各轮评分,并以此权数修正之前的各轮原始分以得出一个最新的标准化的总分。

措施之三是各论文其他评委的评分和排序不被正在评分的评委所知以免影响其评审。

措施之四是综合考虑标准化得分、每轮排序和评委对论文的分类来决定论文的取舍,并允许竞赛委员会主任有一定的酌情权。

从以上的评审办法可以看出,美国赛的评审规则是相当完善的,并且保持了长时间的稳定,有值得我国竞赛的评审工作借鉴之处,其突出特点主要有三个方面。其一是不同阶段采用不同的评分标准(浏览轮采用7分制,评分轮采用百分制);其二是对评委评分进行加权修正以部分排除评委偏好造成的影响;其三是根据标准化的总分来分配论文以使每个评委都能看到不同质量的论文。

3 美国赛历年评述精选

我们搜集了1996年至2004年MCM和ICM的全部发表的评述文章和部分1985年到1994年MCM的发表的评述文章并进行总结,以期得出这个比赛的评审原则,并希望以此来促进我国的大学生数学建模竞赛的评价体系的完善,提升教练指导大学生参加数模竞赛的能力,使我国的这项规模最大的大学生科技竞赛活动能更加兴旺发达,为教育事业的发展做出应有的贡献。

通过对全国大学生数学建模竞赛与 MCM/ICM 评述文章的对比,我们发现国内外的这些评述文章内容就有比较大的差别。国内的不少评述文章把主要精力集中在给出一种建模解决方案,事实上就是参考答案,一般较少篇幅去对优秀论文及其它论文进行评述,有的评述文章甚至根本没有评述,因此多数评述文章不是真正的评述。美国赛的评述文章则较大的不同,在其中根本找不到评述人所建立的模型,也极少有评述指出该如何具体的建模,评述人把大量的精力集中到对所发表的特等奖论文的评价,对共同的优缺点加以描述和点评,对每篇优秀论文的创新之处加以介绍,对如何完成建模的各个步骤以及如何写好论文等给出建议,所有这些都对参赛者和后来者有很好的提示与参考价值,对学生真正学会这种分析解决问题的方法有比较大的好处。

由于美国赛的评述文章分成了评委的评述和实际工作者的评述(后者不是每年都有),加起来有几十篇之多,内容非常丰富并极有启发意义,我们在此精选部分内容与大家分享,并与全国赛的情况进行比较。

初评(浏览的第一、二轮)时主要根据论文的摘要和内容的总体组织来进行评分。关于如何写好摘要,Patrick J. Driscoll 在2002年 MCM 的评述文章[5]中对此有非常详尽的描述,他认为:摘要应该简要的陈述问题、描述建模方法、叙述读者应该记住的重要结果和结论、提出直接针对问题的建议。一个可以用来评价摘要质量的好方法是:“如果某人只读了摘要而未读报告的其余部分,他能大概知道问题是什么、我们做了什么、我们的结论是什么以及我们的建议是什么吗?”他还提请参赛者注意,多数方程、代码、推导都应该出现在摘要之外的地方。

全国赛章程中规定，“竞赛评奖以假设的合理性、建模的创造性、结果的正确性和文字表述的清晰程度为主要标准”。美国赛虽未这样具体明确地指出评审原则，但从其大量评述文章中可以看出假设和合理性、建模的创造性和文字表述的清晰程度也是其重要标准，只是没有一篇评述文章强调结果的正确性，反而有评委强调：由于所作的假设和建立的模型各不相同甚至相差较大，因此结果亦可能有相当大的差别。在其特等奖的论文中某些结果相差50%、100%甚至更大都有可能出现，这与我国的不少模竞赛题有比较接近的结果、甚至部分题目的结果具有唯一性相比有很大的不同。国内赛结果的这种确定性在不少赛区的评审中更被上升到最重要的标准，一旦不符合极有可能无法参加全国的评审，这在客观上将不少有创新思想的论文挡在了全国奖之外，如果能在出题时避免可能出现某些比较固定的结果，为赛题提供更大的灵活性，更进一步，如将章程中的“结果的正确性”修改为“结果的合理性”甚至删掉该原则，这样势必在评审中更加注重创造性，这对我国数模竞赛的作用会有不小的影响。

对于“假设的合理性”的重视程度在国内部分评审中亦显得重视不够，这从部分发表出来的优秀论文中仍有部分论文的假设不合理甚至错误就可见一斑。其实假设的合理性是后面所有工作的基础，一旦未加以仔细分析就有可能导致整个解决方案偏离原问题，从而使其不具有实用性。美国赛的评述认为假设的主要作用是简化问题以便在有限的时间内能应用参赛者所学知识产生解决方案、提供某些必须而在竞赛时间内又无法采集到的信息，需要对假设的合理、必要和实际影响进行清晰的描述，还应该对其进行灵敏度分析，对假设引起的模型的优缺点进行描述。

对于“文字表述的清晰程度”，美国赛和国内赛都非常重视，甚至可以说有过之而无不及，MCM 与 ICM 都以提升参赛者的写作能力为根本目的之一（另一个目的是为了提升参赛者的问题的解决能力）。

对于“建模的创造性”，美国赛中也特别强调，而且其它方面的创造性也被重视，在某些方面表现出很好创造性的论文即使出现比较大的错误亦有可能获得一等奖。其评审办法亦有利于创造性论文的发现（评分轮的评委有30-45分钟的时间阅读论文）。国内赛由于在部分赛区评审中过分强调结果的正确性，再加上赛区评审时每个评委用在一篇论文上的时间有的甚至不到5分钟，这非常不利于创造性思想的发掘。

从全国赛发表的优秀论文以及每年的评述文章可以看出国内赛对模型的建立非常重视，但对问题的分析、模型的检验、结果的分析强调不够，但这却是美国赛最强调的部分。当然这有多方面的原因，比如国内赛题的建模难度相对较高、竞赛时间相对较短（比美国赛少了一天）、传统教学中对检验和分析的强调不够等等。由于数模竞赛赛题多数是现实问题简化而得，建模难度相对较大，出错可能极高，如果忽略了检验环节非常容易出现问题，美国赛评述文章认为模型的检验越多越好。

关于模型的检验，Patrick J. Driscoll 在[5]中指出，可以采用证明的方法，但更多的是对某些感兴趣的情形进行计算并分析结果、对重要参数的高中低水平进行计算并分析，考虑放松某些假设等。这方面的工作在全国赛发表的优秀论文中分量比较轻，像误差分析之类对实际问题有重要意义的工作更是可称为罕见了。

美国赛的每篇论文中几乎都有专门的“结论”一节，在其中给出前面的建模、分析所获得的有关的结论，但这在全国赛论文中却罕有踪影。其实对实际问题的解决而言，中间的建模、求解、检验、分析等等仅仅是分析过程的再现而已，这对问题的解决来说固然重要，但对不少实际问题提出者来说要理解与掌握的难度会比较大，他们最关心的是自己问题的解决方案和相关的结论，这些都是需要在结论一节中以其熟悉的方式而非过分技术化地表达出来的。

从近年的获奖数据可以看出，美国赛的获奖比例比全国赛明显要高，只是其英文题目和英

文论文限制了部分学校的参赛热情,但从更好的锻炼学生的角度出发应该鼓励更多的学生去参加该项竞赛。

根据美国赛的众多评述文章,要想在美国赛中获得二等奖,需要特别注意摘要和内容的组织,还要有真正的建模,另外就是要重视论文的写作。再进一步,要想获得一等奖除了前面的要求外,更需要特别重视各种分析工作(模型的检验也是分析的一种),需在某些方面有特别之处。如果要获得特等奖那就必须要全面完整的解决问题、详尽细致的分析、清晰简明的表述,即使是未能恰当的引用文献、术语符号的说明不够清楚等细微方面做得不好都会将参赛者挡在特等奖的大门之外。获得特等奖的国内参赛院校屈指可数。

4 结论

美国赛和国内赛既有区别也有很多共同之处,美国赛有如下方面可供国内赛借鉴:

- 赛题具有更多的灵活性和开放性,给学生提供了较大的发挥空间;
- 有一套完善、稳定的评审办法。该办法使越是质量好的论文有越多的时间加以评审,并从客观上承认并有具体措施尽量减少评委的主观偏向;
- 评审方面不强调结果的正确性而只是强调其合理性;
- 特别强调推理和分析的合理性,重视模型的检验和结果的分析;
- 重视模型的实用性和有效性而非模型的精巧与解法的精妙。

参考文献:

- [1] 叶其孝. 学生数学建模竞赛辅导教材(一)[M]. 长沙: 湖南教育出版社, 1993
- [2] Fox William P. Judge's commentary: the outstanding student person papers[J]. The UMAP Journal, 2003, 24(3): 317-322
- [3] Giordano Frank. Contest director's commentary: judging the MCM[J]. The UMAP Journal, 1996, 17(3): 341-344
- [4] Harper Steve. Practitioner's commentary: computer support for the MCM[J]. The UMAP Journal, 1996, 17(3): 345-348
- [5] Driscoll Patrick J. Judge's commentary: the outstanding wind and waterspray papers[J]. The UMAP Journal, 2002, 23(3): 267-271

The Judge and Commentary of American Undergraduate Contests in Modeling

HU Dai-qi

(Department of Mathematics, Jinan University, Guangzhou 510632)

Abstract: This paper introduces the judging method of American undergraduate mathematical contest in modeling (MCM) and the interdisciplinary contest in modeling (ICM), summarizes judge's commentaries of these contests, and compares MCM/ICM with China undergraduate mathematical contests in modeling, we point out some American methods which could be used for our reference.

Keywords: American undergraduate mathematical contest in modeling; the interdisciplinary contest in modeling; judging