写好英语科技论文的诀窍：
主动迎合读者期望，预先回答专家可能质疑

周耀旗
印地安那大学信息学院
印地安那大学医学院计算生物学和生物信息中心

以此文献给母校中国科技大学五十周年校庆

前言

我的第一篇英语科技论文写作是把在科大的学士毕业论文翻译成英文。当我一九九零年从纽约州立大学博士毕业时，发表了20多篇英语论文。 但是，我对怎样写高质量科技论文的理解仍旧处于初级阶段，仅知道尽量减少语法错误。之所以如此，是因为大多数时间我都欣然接受我的博士指导老师Dr. George Stell和Dr. Harold Friedman的修改，而不知道为什么要那样改，也没有主动去问。这种情况一直持续到我去北卡州立大学做博士后。我的博士后指导老师Dr. Carol Hall建议我到邻近的杜克大学去参加一个为期两天的写作短训班。这堂由Gopen教授主办的短训班真使我茅塞顿开。第一次，我知道了读者在阅读中有他们的期望，要想写好科技论文，最有效的方法是要迎合他们的期望。这堂写作课帮我成功地完成了我的第一个博士后基金申请，有机会进入哈佛大学Dr. Martin Karplus组。在哈佛大学的五年期间，在Karplus教授的指导下，我认识到一篇好的论文需要从深度广度进行里里外外自我审查。目前，我自己当了教授，有了自己的科研组，也常常审稿。我觉得有必要让我的博士生和博士后学好写作。 我不认为我自己是写作专家。我的论文也常常因为这样或那样的原因被退稿。但是我认为和大家共享我对写作的理解和我写作的经验教训，也许大家会少走一些我走过的弯路。由于多年未用中文写作，请大家多多指正。来信请寄：yqzhou@iupui.edu。 欢迎访问我的网站：http://sparks.informatics.iupui.edu。

导言

通常来讲，研究生和博士后从他的导师那儿得到研究方向。经过多次反复试验，得到一些好的结果。接下来他们需要对得到的数据进行总结和分析，写成论文。一篇精写的论文更容易被高档杂志接受。而写得不好的论文很可能被退稿。论文的数量和质量是学生和导师事业发展的敲门砖。不成文，便成仁，是学术生涯的写照。

很多学生以为当结果到手的时候研究就结束了。他们写的草稿，常常把原始数据放在一起，没有对方法和数据进行详细分析，没有对当今论文的评述。事实上，写作是研究不可分割的一部分。此刻是弄懂方法的成功与失败，寻找结果的解释及其隐含的意义，以及与其他相关研究进行比较的时候。

我们为什么需要在写作上如此认真努力？原因很简单。一个研究结果只有在被别人使用时才有意义。而想被别人使用，文章必须能引起其他科学家的兴趣，而且得保证其他人能看懂并可以重复和再现你的结果。只有可以被理解的研究才会被重复，也只有可以被再现的工作才能导致别人的引用和跟踪。而你的论文被引用的数量常常用来衡量研究的影响力。从某种角度看，写作就象是把你的工作成果推销给其他的科学家。

为了更好的推销，科学论文必须满足它独特的顾客：由聪明能干的科学家组成的尖端读者。它必须能先说服（通常也是竞争对手的）同行们，因为他们的评审是文章在发表前的第一道关口。同时，它也必须满足一般读者的要求。为了达到这个目标，我们首先要理解他们需要什么？

读者需要什么

你的文章的潜在读者可能有刚进入这领域的新手，大学生和研究生，也有专家（潜在审稿人），他们对你的领

域会有不同程度的了解。因此，写文章的时候应该力求简单到可以被新手理解，同时深刻到可以引起专家的兴趣。

所有的科学家（不论是学生还是他们的导师)往往都很忙。大量期刊杂志使他们不可能仔细阅读每一篇论文。他们通常希望能在最短时间内找到文章最重要的信息。典型的情况是如果文章标题不吸引人，他们或许就会跳过这篇论文，如果文章的摘要没有包含重要的新方法或新结果，他们不会去读这篇文章。即使已经决定要读的论文，他们也会跳过很多段落直接去找自己最感兴趣的地方。因此，保证文章的结构能使读者很快找到所需的信息非常重要。文章的关键在于结构，不在于语法。语法错误易改，结构错误则往往让人无从下手，不知所云。我审过一些国内同行的论文，结构问题很常见。

总之，一篇文章只有在不需太多努力就可以理解的情况下才会被广泛地引用。文章清晰的关键就是使读者能在他们想找的地方找到他们需要的东西。这也就是说，**要想让读者不费力理解你的论文，你必须费力去满足他们的期望**。

读者期望什么

读者对句子的期望

1. 读者希望在句子的开始看到熟悉的信息。句子是文章的最小功能单元。最容易理解的句子是整句都在说读者知道的东西。但这对科技论文是不可能的，因为只有新的东西才会被发表。事实上科技论文通常会包含很多新术语。所以一个容易理解的句子应该从读者熟悉的信息（或刚刚提过的）开始而以新信息结束，并在它们之间平滑地过渡。好文章的所有句子都应该这样从旧到新地平滑过渡。帮助你写好一句开头的金科玉律是问问你自己：“我以前有没有提过这个概念？”大多数文章很难读是因为很多新概念在没有被介绍之前就使用了。例如：

> Samples for 2-dimensional projection of kinetic trajectories are shown in Figure 7. The coil states are loosely gathered while the native states can form a black cluster with extreme high density in 2-dimensional projection plane.

这里从第一句到第二句信息无法流动。“The coil states”不知道是从何而来的。读者会发现下面改动后的句子更容易明白。

> Kinetic trajectories are projected onto xx and yy variables in Figure 7. This figure shows two populated states. One corresponds to loosely gathered coil states while the other is the native state with a high density.

在这个新段里，新插入的第二句使每句均能从旧信息出发到新信息结束。第一句与第二句之间以“Figure”相连而第二句与第三句之间以“two states”相连。而新信息“coil states”则出现在第三句的最后。整段环环相连，成为一个整体。再看一个例子：

> The accuracy of the model structures is given by TM-score. In case of a perfect match to experimental structure, TM-score would be 1.

在第二个句子里，旧信息“TM-score”被埋在中间，被新信息“a perfect match to experimental structure”打断了。这里建议修改如下：

> The accuracy of the model structures is measured by TM-score, which is equal to 1 if there is a perfect match to the experimental structure.

科技写作中的最大问题就是新旧信息顺序颠倒。新信息和旧信息对作者来说可能不是很好区分，因为他非常熟悉所

有的信息。为了避免这种问题，不管什么时候，每当你开始写新句，你应该问问自己，这些词前面有没有被提到过。**一定要把提到过的放前面，没提过的放后面**。

2. 读者想在主语之后立刻看到行为动词。对一个说明谁在做什么的句子，读者需要找到动词才能理解。如果动词和主语之间相隔太远，阅读就会被寻找动词打断。而打断阅读就会使句子难以理解。这里有个例子：

> The smallest URFs (URFA6L), a 207-nucleotide (nt) reading frame overlapping out of phase the NH2-terminal portion of the adenosinetrip hosphatase (ATPase) subinit 6 gene has been identified as the animal equivalent of the recently discovered yeast H+-ATPase subunit 8 gene.

同样的句子，将动词放在主语之后：

> The smallest of the URFs is URFA6L, a 207-nucleotide (nt) reading frame overlapping out of phase the NH2-terminal portion of the adenosinetriphosphatase (ATPase) subinit 6 Gene; it has been identified as the animal equivalent of the recently discovered yeast H+-ATPase subunit 8 gene.

这样新的句子就更加平衡了。尽量避免过长的主语和过短的宾语。这就像头重脚轻的人很难站稳。**短的主语紧跟着动词加上长的宾语效果会更好**。

3. 读者期望每句只有一个重点，这个重点通常在句尾。比较下面两个句子，我们可以感觉到他们着重强调不同的东西。

> URFA6L has been identified as the animal equivalent of the recently discovered yeast H+-ATPase subunit 8 gene.
>
> Recently discovered yeast H+-ATPase subunit 8 gene has a corresponding animal equivalent gene URFA6L.

很明显，前面的句子是关于一个最近发现的酵母基因，而第二句则着重强调了它有一个和动物一致的基因。另外一个例子：

> The enthalpy of hydrogen bond formation between the nucleoside bases 2-deoxyguanosine (dG) and 2-deoxycytidine (dC) has been determined by direct measurement.

这个句子看起来好像是在强调“direct measurement”。 这不太像是原作者的目的。颠倒一下会使句子更加平衡。

> We have directly measured the enthalpy of hydrogen bond formation between the nucleoside bases 2-deoxyguanosine (dG) and 2-deoxycytidine (dC).

新的句子更简单而且更短，同时避免了头重脚轻的症状。总之，句尾是读者对该句最后的印象。**把最好的，最重要的，和想要读者记住的东西放在句尾**。

<u>读者对段落的期望</u>

每一个段落都应该只讲一个故事。在一段里表述多个观点会使读者很难知道该记住什么、这段想表达什么。一段的第一句要告诉读者这一段是讲什么的。这样读者想跳过这段就可以跳过。一段的最后一句应该是这段的结论或者告诉读者下一段是什么。段落中的句子应该由始到终通过逻辑关系连接，实现由旧信息到新信息的流动。比如这一段：

The enthalpy of hydrogen bond formation between the nucleoside bases 2-deoxyguanosine (dG) and 2-deoxycytidine (dC) has been determined by direct measurement. dG and dC were derivatized at the 5 and 3 hydroxyls with triisopropylsilyl groups to obtain solubility of the nucleosides in non-aqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. From isoperibolic titration measurements, the enthalpy of dC:dG base pair formation is -6.650.32 kcal/mol.

很难知道作者在这段里想表达什么。从这段的起始和结束看来，焓(enthalpy)应该是他想表达的重点。下面是重新组合后的段落。

We have directly measured the enthalpy of hydrogen bond formation between the nucleoside bases 2-deoxyguanosine (dG) and 2-deoxycytidine (dC). dG and dC were derivatized at the 5 and 3 hydroxyls with triisopropylsilyl *groups*; *these groups* serve both to solubilize the nucleosides in non-aqueous solvents and to prevent the ribose hydroxyls from forming hydrogen bonds. The enthalpy of dC:dG base pair formation is -6.650.32 kcal/mol according to isoperibolic titration measurements,

首句描述了整段的主题。原段里的第一句颠倒是为了1）使新信息“dG”和“dC”在句子最后并强调他们。2）更好地跟下面一句衔接。 原段里的第二句被分成两部分，这样每一部分只表达了一个观点。最后一句时总结整段。 再看另一个例子：

Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates at which tectonic plates move and accumulate strain at their boundaries are approximately uniform. Therefore, in first approximation, one may expect that large ruptures of the same fault segment will occur at approximately constant time intervals. If subsequent main shocks have different amounts of slip across the fault, then the recurrence time may vary, and the basic idea of periodic main shocks must be modified.

在这个例子里，前两句共同阐明了积累张力的速度（Rate Of Strain Accumulation)。然而，第一句里的旧信息并没有放在第二句的开始。读者读到第三句的时候通常就不明白这段到底要讲什么了。更清晰的描述应该如下：

Large earthquakes along a given fault segment do not occur at random intervals because it takes time to accumulate the strain energy for the rupture. The rates of strain accumulation at the boundaries of tectonic plates are approximately *uniform*. Therefore, *nearly constant time* intervals (at first approximation) would be expected between large ruptures of the same fault segment. [However?], *the recurrence time* may vary; the basic idea of periodic main shocks may need to be modified if subsequent main shocks have different amounts of slip across the fault.

新段现在着重阐明了地震的发生频率。下划线标明了以前描述过的旧信息。很明显，新旧信息的连接是理解这段的关键。从旧信息到新信息的流动是使读者轻松阅读的最佳方式。**写文章的目的不是去测试读者的阅读能力，而是考验作者的表达能力。不能怪人没看懂，只能怪自己没写清楚**。常常听到这样的抱怨：那审稿人连这都不懂! 审稿人也可以说：连这个也写不清楚。

读者对表格和图示的期望

一些没有耐心的读者会直接通过图表来判断一篇文章是否值得一读。怎样能使读者不需读正文就能理解图表是至关重要的。

对于表来说，由于我们是从左向右阅读的，我们熟悉的信息应该出现在左边而新的信息出现在右边。例如，下面列出的表1和表2是仅仅调换了两列。比较一下那个表格更易理解。

表 1：

Temp (°C)	Time
25	0
27	3
29	6
32	12
32	15

表 2:

Time	Temp (°C)
0	25
3	27
6	29
12	32
15	32

显然因为我们更熟悉时间作为独立变量，表2就比表1容易读些。做表的另一条规则是把最好的留在最后。也就是最能使人感兴趣的结果应该放在最右边一列或在最后一行，因为这些地方是读者结束阅读并能留下印象的地方。下面的例子比较了各种方法的精度。 最后一行展示了现在得到的结果。

Table 3:

Benchmark	SALIGN	Lindahl	PROSPECTOR 3	LiveBench 8
Method	Alignment	MaxSub	MaxSub	MaxSub
SPARKS	53.1%	325.9	529.0	38.3
SPARKS2	54.9%	341.0	591.0	40.7
This work	**56.6%**	**349.2**	**601.9**	**42.2**

对于图，我们至少应该对所有的标签（数字，座标，说明）使用大的黑体Helvetica字体。只画出重要的区域。尽量不用彩色就能使曲线达到最大的区分（彩色的图很贵)。

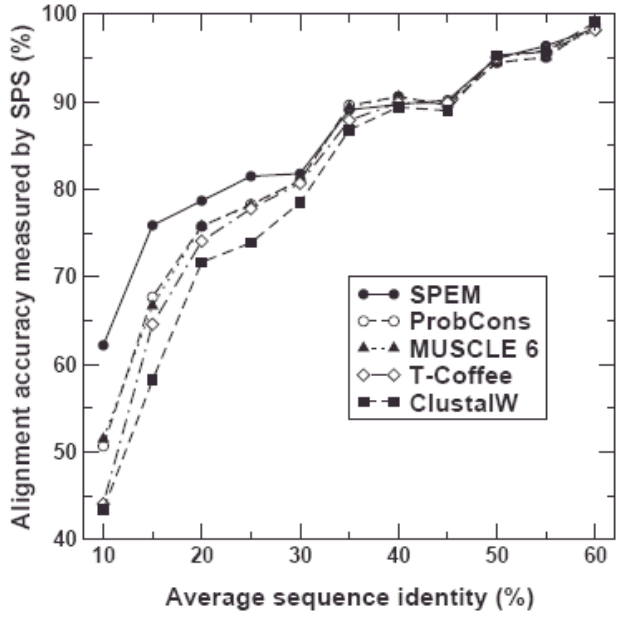

Fig. 3. Alignment accuracies (measured by SPS) as a function of average sequence identity given by methods SPEM, ProbCons, MUSCLE 6.0, T-Coffee and ClustalW, shown as labeled. Each point is represented by the lower bound of sequence identity at each bin.

这是我喜欢的一个图。它说明了一些画图的原理。对你的工作用实线而对别人的工作用点画线。间隔使用实心和空心符号来使曲线之间的不同更加明确。详细说明X和Y座标，标题不用缩写。

审稿人要什么？

文章在发表前必须经过审稿人的评审。他们一般是相关领域的专家甚至是你的竞争者。他们会尽力寻找你文章中的毛病。有时，由于不同的观点和竞争的需要，审稿人或许会试图阻止你的文章发表。因此，文章必须写得理由充足。**在被别人挑剔之前，自己必须首先鸡蛋里挑骨头，预先回答审稿人的可能质疑**。

怎样满足审稿人？

1). 只提出“一”个中心命题。论文里的观点太多，不但不好写，问题也容易多，读者也不易记住你要说什么。

2). 在这个中心命题的基础上，用一个迷人(但决不能夸张)的标题来吸引审稿人的兴趣。审稿人只审批感兴趣的论文。如果你不能引起审稿人的兴趣，那最好不要发表那篇文章。编辑们有时候会很郁闷，因为找不到有兴趣的审稿人。无偿审稿也只有科学界才有。

3). 合理解释每一个参数，合理说明每一个步骤。审稿人没时间考虑细节。程序和参数的合理化显示出你知道你在做什么，而不是凑数据。**即使你是在凑数据，也要把凑数据的过程合理化**。

4). 问问你自己是否提供了足够重复你工作的所有细节。审稿人(或读者)越容易再现你的工作，他就越可能接受你的文章。当然，审稿人并不会真正去重做你的工作，但你必须通过你的描述使他相信可以重做。

5). 必须有说服力！尽量做彻底而不是半成品的工作！用多方面测试来证明你的中心命题。要使文章象律师证明无罪官司，预先回答一切可能提出的疑问。

6). 引用所有重要的研究工作，特别是经典力作。写作的时候要再做全面文献检索。

为了达到这些目标，写科学论文的时候必须遵照一定的框架结构。

文章的结构

典型的科学论文包括标题，摘要，引言，方法/实验步骤，结果，讨论，感谢，和参考文献。这样的结构是用来帮助读者快速找到他们感兴趣的信息。把信息放错地方会使读者糊涂。常犯的错误是混淆事实(结果)和解释(讨论)。讨论是对结果的解释及说明它的意义，而不是重复结果的描述。

一篇论文是从摘要，引言开始，这里建议从方法和结果部分开始写，因为你对方法和结果最熟悉，此外只有更好地理解方法和结果，才能确定中心命题。而标题，引言和讨论的写作都需要中心命题。我们应该从最熟悉的事情开始，就像读者从他们最熟悉的地方开始理解一样。

方法/实验步骤

如果文章是关于新的方法，技术，或算法，要非常详细地写它的新颖之处。要用有逻辑的、合理的方式来描述它。这会帮助读者抓住新方法的要领。如果这个方法使用参数，则要把每一个参数(或参数的取值)合理化，或者是以前用过的，或者可以从物理或数学推导出来，或者通过了广泛的测试及优化。如果无法保证它的合理性，那就必

须描述改变它会造成的影响(实际的结果应该在结果部分或讨论部分，方法部分仅包含影响的描述)。如果没有测试它们的合理性，你应该解释为什么(做的代价太贵了？太费时间了？或者需要延期到将来做)。参数改变造成的影响可以衡量方法是否Robust。Robust的方法应该是在参数改变很大的时候，结果也不会太大变化。

对于新方法的发展，你同样需要设计不同的方法来测试。让人信服就需要做尽可能多的测试。你所能找到或设计的测试越多，你的工作就越会被其他人所接受和使用。

当完成了方法部分以后，问一问自己以下的问题：1)新的术语是不是都定义了？2)如果你是第一次读这部分，你能否得到重复整个工作的所有信息？记住，不要隐藏任何窍门或使用的捷径。人们如果不能重复你的结果的话就不会相信你的论文。永远不要弄虚作假！别人不是傻子。一山更比一山高。聪明的大有人在。如果你伪造数据，心存侥幸不会被人发现。如果真的没人发现的话，那就是没有任何人想重复或使用你的结果，那只能说明你的结果根本不值得发表，毫无意义。若要人不知，除非己莫为，这是千真万确的真理。

结果部分

当你开始写结果部分时，先考虑一下结果的意义。也就是说，你理解你的结果吗？这些结果是不是告诉了你更深刻的东西？你能从很多不同角度来理解结果吗？你能设计证明或者反驳你的一些解释的新测试吗？

如果你发现了新现象，你必须证明你的结果不是你方法制造出来的(讨论部分的一个好内容)。它可以在不同的条件下重复吗？如果你发展了一个新方法，你必须证明这个方法的重要性。它是否改进了现有的方法？你的结果部分必须用不同的角度或多重测试来支持新发现或验证新方法的重要性。

一旦你对结果有更好的理解，你需要决定卖点，也就是说这篇文章最有意义的一个观点是什么？确定这篇文章的中心命题之后要组织所有的段落来证明、支持它，用数据(有必要的话再加数据)来证明它。同时也要排除其他可能性。放弃与中心命题无关的数据，即使这些数据是很辛苦得来的。

标题

当你有了中心命题之后，就该决定文章的标题了。标题可以为你的方法，你的结果或结果的隐含意义做广告。文章的标题一般只有一句。应该把最重要，最吸引人的信息放进标题。比如，标题 “Steric restrictions in protein folding: an alpha-helix cannot be followed by a contiguous beta-strand” 主要突出了结果。另一方面，标题“Interpreting the folding kinetics of helical proteins” 突出了结果的含义。用标题 “Native proteins are surface-molten solids: Application of the Lindemann criterion for the solid versus liquid state” 的话，同时突出了方法和结果的含义。注意标题 “Native proteins are surface-molten solids” 是结果的解释，而不是结果本身。用既广泛又具体的标题，这样才能吸引更多的读者。

引言部分

中心命题和标题都决定了以后，就该写引言了。第一件该做的事就是围绕中心命题来收集所有相关文献。搜索并研究所有最近和相关的文章(通过对中心命题关键字的搜索或用引用索引)。确认你有所有最新的论文。引用所有重要的文章。如果你不引用别人的文献，别人也不会引用你的！如果你想谁引用你的工作，你要先引用他的。你引用的文章越多，他们越可能阅读并引用你的文章。因为人们更加关注引用他们的论文。仔细读你所引用的文章，避免引用错误。在引用上，不要偷懒。

引言的第一句最难写，因为它决定了你整个引言的走向。我的办法是把第一句和文章的标题连起来。在第一段以最基本和常见的术语来定义标题里用的一些术语。从这个术语，引入研究的领域和它的重要性。第二段应该对这个研究领域作一个鉴定性的论述。如果中心命题是关于解决一个问题的方法。这一段就应该指出这个当前研究中现存未解决的问题。描述解决这个问题的难度或挑战。 第三段引入你提出的办法和它大致会带来什么效果。你可以大略地描述你的结果和它的含义。这里有个例子。

Assessing secondary structure *assignments* of protein structures by using pairwise sequence-alignment benchmarks

The secondary structure of a protein refers to the local conformation of its polypeptide backbone. Knowing secondary structures of proteins is essential for their structure classification[1,2], understanding folding dynamics and mechanisms[3-5], and discovering conserved structural/functional motifs[6,7]. Secondary structure information is also useful for sequence and multiple sequence alignment[8,9], structure alignment[10,11], and sequence to structure alignment (or threading)[12-15]. As a result, predicting secondary structures from protein sequences continues to be an active field of research[16-18] fifty six years after Pauling and Corey[19-20] first predicted that the most common regular patterns of protein backbones are the α-helix and the β-sheet. Prediction and application of protein secondary structures rely on prior *assignment* of the secondary-structure elements from a given protein structure by human or *computational methods*.

Many *computational methods* have been developed to automate the assignment of secondary structures. Examples are DSSP,STRIDE, DEFINE, P-SEA, KAKSI,P-CURVE, XTLSSTR, SECSTR, SEGNO, and VoTAP. *These methods* are based on either the hydrogen-bond pattern, geometric features, expert knowledge or their combinations. However, *they* often **disagree** on their assignments. For example, **disagreement** among DSSP, P-CURVE, and DEFINE can be as large as 25%. More beta sheet is assigned by XTLSSTR and more pi-helix by SECSTR than by DSSP. The **discrepancy** among different methods is caused by non-ideal configurations of helices and sheets. As a result, defining the boundaries between helix, sheet, and coil is problematical and a significant source of **discrepancies** between different methods.

Inconsistent assignment of secondary structures by different methods highlights the need for a criterion or a benchmark of "standard" assignments that could be used to assess and compare assignment methods. One possibility is to use the secondary structures assigned by the authors who solved the protein structures. STRIDE, in fact, has been optimized to achieve the highest agreement with the authors' annotations. However, it is not clear what is the criterion used for manual or automatic assignment of secondary structures by different authors. Another possibility is to treat the consensus prediction by several methods as the gold standard. However, there is no obvious reason why each method should weight equally in assigning secondary structures and which method should be used in consensus. Other used criteria include helix-capping propensity, the deviation from ideal helical and sheet configurations, and structural accuracy produced by sequence-to-structure alignment guided by secondary structure assignment.

In this paper, we propose to use sequence-alignment benchmarks for assessing secondary structure assignments. These benchmarks are produced by 3D-structure alignment of structurally homologous proteins. Instead of assessing the accuracy of secondary-structure assignment directly, which is not yet feasible, we compare the two assignments of secondary structures in structurally aligned positions. We assume that the best method should assign the same secondary-structure element to the highest fraction of structurally aligned positions. Certainly, structurally aligned positions do not always have the same secondary structures. Moreover, different structure-alignment methods do not always produce the same result. Nevertheless, this criterion provides a mean to locate a secondary-structure assignment method that is most consistent with tertiary structure alignment. We suggest that this approach provides an objective evaluation of secondary structure assignment methods.

在这个例子里，标题推荐了一个评估指派蛋白质二级结构的方法。第一段以二级结构的定义开始（与标题相连)。整段描述了二级结构的重要性。最后一句过渡到指派二级结构的计算方法（下一段的主题)。注意“计算方法”放在句

子的最后是为了强调而且和第二段的开始连接在起来。第二段则聚焦在计算方法中存在的问题。旧信息“计算方法”逐渐的变到了“他们的不一致”。第三段的第一句把主题从“不一致”（旧信息）转变成了“评估的办法”（新信息)。然后，介绍了这个领域已有的工作。第四段引入新方法并讨论了新方法的优点。第五段（这里没有给出）将会简要地讨论结果。每一个引言应该 包括研究领域的介绍和意义，做这工作的具体原因，结果和隐含的意义。一般而言，读者读完引言，对论文的来龙去脉就应该清清楚楚了。

讨论部分

现在到了你写论文的最后一部分。很多人认为讨论部分最难写。他们常常不知道该写什么。学生常常不能把结果从他们的解释，含意和结论中分离出来。此外，他们不善于思考可能存在的其他解释。好的讨论通常以得到的结果和解释的评论开始。其它可用于讨论的内容有：参数改变对结果的影响，与其他研究相比还有待解决的问题，将来或正在进行的工作（防止别人从事你显而易见的，立刻就能实现的后续工作)。这里有一段文章中的讨论部分。

> One question about the complex homopolymer phase diagram presented here is whether it is caused by the discontinuous feature of the square-well potential. We cannot give a direct answer because the DMD simulation is required to obtain well-converged results for the thermodynamics. However, the critical phenomena predicted for a fluid composed of particles interacting with a square-well potential are as realistic as those predicted for a fluid composed of particles interacting with a LJ potential. Also an analogous complex phase diagram is found in simulations of LJ clusters. The present results for square-well homopolymers may well be found in more realistic homopolymer models and even in real polymers.

这一段探究了可供选择的解释。

摘要部分

整篇文章写完了。你需要写文章的摘要了。典型的摘要包括课题领域的重要性（回到标题)，要研究的问题，你方法的独特性，结果的意义和影响。这里有个例子。

> How to make an objective assignment of secondary structures based on a protein structure is an unsolved problem. Defining the boundaries between helix, sheet, and coil structures is arbitrary, and commonly accepted standard assignments do not exist. Here, we propose a criterion that assesses secondary-structure assignment based on the similarity of the secondary structures assigned to structurally aligned residues in sequence-alignment benchmarks. This criterion is used to rank six secondary-structure assignment methods: STRIDE, DSSP, SECSTR, KAKSI, P-SEA, and SEGNO with three established sequence-alignment benchmarks (PREFAB, SABmark and SALIGN). STRIDE and KAKSI achieve comparable success rates in assigning the same secondary structure elements to structurally aligned residues in the three benchmarks. Their success rates are between 1-4% higher than those of the other four methods. The consensus of STRIDE, KAKSI, SECSTR, and P-SEA, called SKSP, improves assignments over the best single method in each benchmark by an additional 1%. These results support the usefulness of the sequence alignment benchmarks as the benchmarks for secondary structure assignment.

前两句陈述了问题。第三句提出了解决办法。这些句子后面跟着结果。整个摘要以总结收尾。

总结

1. 认真对待写作。尽你最大努力花时间写作。它是科学研究的重要一环。文章没写好，没人看，没人用，等于没发表。

2. 除非这个研究是全面彻底的，而且你试了所有可以支持你结论的方法，否则不要去发表。

3. 重新思考，并合理解释为什么做这项工作，做了什么，什么是最重要的发现？为什么用这个方法？为什么用这些参数？什么是以前做过的（更新文献搜索)？不同在什么地方？

4. 要从批判的角度来看你的工作。只有这样，才能找到弱点，进一步发展。我的许多论文是在反复讨论中大幅度修改，许多计算经常要重做。只有理顺和理解结果，文章才会更有意义。

5. 要能回答所有合理的质疑。如果你自己有疑问，一定要搞清楚，否则别人又怎会相信。

6. 不要隐藏任何事实，不做假，不要低估其他科学家的智慧。让你的研究可重复。把所有的材料和数据上网。

7. 从头（标题）到尾（结论或讨论）要从旧信息过渡到新信息。永远不要在句子的开头引入新信息。切忌在术语被定义之前使用它们。

8. 在段首要有阐明整段主题的句子，在段尾要有连到下段的过渡句。从标题到结论都要连贯。句句相扣，段段相连，让一篇论文是一个整体而不是杂乱无章地把句子堆积在一起。这样才能使读者享受阅读你的文章。

9. 写，重写，再重写。没有人能第一次就写好。不花时间，不下功夫，写不好。我的文章一般要修改十次以上。

感谢

此文中的一些例子出自 “The Science of Scientific Writing” by G. D. Gopen and J. A. Swan, Scientific American, 78, 550-558, 1990. 我在杜克大学Gopen教授1995年年度短训班受益非浅。我要特别感谢我的导师Martin Karplus(哈佛大学)，George Stell (纽约州州立大学－石溪校区), Harold L. Friedman (纽约州州立大学－石溪校区)和 Carol Hall (北卡罗来纳州立大学)的鼓励和指导。没有他们，我不会有那么多机会练习英文写作。最后，我要感谢我的学生和博士后。他们对科学的贡献使我可以继续写论文，基金申请，或评论。此文中的一部分例子来自与他们合作的文章。此文初稿是用英文写的。由于我的中文打字速度太慢，特别感谢徐贝思帮我翻译成中文初稿。如果有不妥的地方是我的问题，请多指教。

二零零七年六月一日于印地安那