这是一篇有些实验性和推测性的文章。这周我参加了我是其中一位组织者的IPAM关于机器辅助证明的研讨会。我们有一系列有趣而多样的演讲，既有计算机科学家介绍最新的工具，可以正式验证证明或自动化证明写作或证明发现的各个方面，也有数学家描述他们使用这些工具来解决他们的研究问题的经验。可以在IPAM的youtube频道上找到这些演讲的视频；我也在我的Mathstodon账户上发布了关于活动期间的演讲的内容。我当然不是最客观的人来评判，但从我收到的反馈来看，似乎会议成功地实现了它的目标，即将对这个话题感兴趣的不同社区聚集在一起。

由于会议的原因，我开始思考现在可能开发出什么计算机工具，可以广泛地用于数学家，特别是那些没有编写代码或安装软件的细节方面的先验专业知识的人。一个我想到的想法是一个潜在的工具，可以将，比如说，一个arXiv预印本作为输入，返回一种图表，详细描述了论文中主要定理和引理的逻辑流程。这目前是由作者手工完成的，在一些但不是所有的论文中（并且通常也可以从正式验证的证明中自动生成，例如在IPAM研讨会附带的图形中，或者这个图表，它是从Massot的蓝图软件中生成的，该软件从手动输入的一组定理和依赖关系作为正式化证明的前奏[感谢Thomas Bloom提供这个例子]）。例如，这是我和我的合作者Rachel Greenfeld为一篇最近的论文绘制的图表：

文本

中度可信度描述已自动生成

这个特殊的图表包含了一些关于布局的主观设计选择，这些选择导致了一些结果被指定为足够重要，需要一个专门的框（而不是被视为一个从一个框到另一个框的工具），以及如何描述每一个结果（以及如何给它们上色）。这仍然是一个非常人力密集的任务（我和我的合作者经过了这个特殊图表的几次迭代，有很多反复的讨论，直到我们都满意）。但我可以看到创建一个自动工具的可能性，它可以提供一个这样的图表的初始“近似”，然后人类用户可以根据自己的喜好进行修改（也许使用一些方便的GUI界面，例如LaTeX中绘制交换图的Quiver在线工具的一些变体）。

作为一个粗略的第一次尝试自动生成这样的图表，一个可能的方法是开发一个工具来抓取一个LaTeX文件，找出文本中所有定理环境的实例（即，所有正式标识的引理、推论等），并且对于每一个这样的定理，找出一个看起来与该定理相关的证明环境实例（要做到这一点，可能需要一些机器学习，不过也许可以希望证明环境实例与定理环境实例的接近程度在很多情况下就足够了）。然后识别出证明环境中对其他定理的引用，开始构建蕴含关系的树，然后可以用类似上面的图表来表示。这样的方法可能会漏掉很多蕴含关系；例如，因为许多引理可能不是用正式的证明环境来证明的，而是用一些更自由流动的文本讨论，或者也许是一行简单的证明，比如“通过结合引理3.4和命题3.6，我们得出结论”。另外，一些对论文中其他结果的引用可能不是通过直接引用进行的，而是通过一些更间接的证明，比如“利用前面的引理，我们得到”或者“通过重复第3节中的论证，我们有”。不过，即使是这样一个粗略的图表也可能是有帮助的，既可以作为作者制作一个改进图表的起点，也可以作为一个试图理解一个冗长论文的学生得到一些逻辑结构的初步想法。

更高级的功能可能是试图利用论文的更多文本来给每个结果分配一些重要性的度量（然后相应地对图表进行加权，突出显示更重要的结果），试图给每个结果一个自然语言的描述，并且以某种方式捕捉那些没有被整齐地封装在一个定理环境实例中的关键语句，但我想象这样的任务应该推迟到一些更粗糙的概念验证原型可以展示之后。

无论如何，我很想听听关于这个想法（或者它的一些修改）是否（a）用当前的技术实际可行（或者更好的是，已经以某种形式存在），以及（b）对研究数学家有兴趣的意见。