Writing up solutions

There are two stages to solving a mathematical problem. The first is to understand the problem and arrive at a solution. The second is to write the solution up. Both are important. Sometimes the second is more difficult than the first. (This often happens with combinatorial problems.) You need to be clear enough that your strategy and argument can be followed by a reader. There are ways of writing up solutions effectively; this little guide will indicate some of them.

You should realize that a solution is an act of communication between two individuals. You must assume that there is a **reader** who is not a mind reader. You have worked through your problem and have in your mind some concept of how the problem was approached and a solution found. These mental processes will not always be evident to your reader, and so you must try to bring the reader along into your way of thinking. Sometimes this can be achieved with very few words to provide a cue. In particular, on a contest, the marker must understand where you are coming from. Often markers feel that students have the essence of a good idea, but are unable to identify it. This frustration translates into a lack of marks that might otherwise be had.

Think of a solution as a flow of logic that carries you step by step from what you are given to the desired conclusion.

Use complete English or mathematical sentences in your solution.

Paragraph your solution so that each section deals with a single main idea. It should be possible for the reader to glance over your solution and get a sense of its architecture, the main steps and your methods of dealing with them.

If the solution seems to be complex in terms of explaining it, first make a plan, indicating in bulleted form what you intend to do. Occasionally, there is a significant result as part of your solution which could be separated out as a Lemma, with its own separate proof. However, this should be used sparingly. Small steps can be justified within the body of your solution, and standard results you need can just be appealed to.

Sometimes it is useful to number or asterisk a step, preferably in parenthese to the right, that you may wish to refer to later. Do not overdo this; your primary goal is to be helpful to the reader.

Avoid the use of disembodied symbolism that forces the reader to try to fill in gaps.

It is often necessary to edit or rewrite your solution to make it clear and to present the ideas in the correct logical order. This should certainly be done when time is not an issue (for example, in doing the correspondence program). Getting in the habit of doing this will, even on contests, help you arrive quickly at more polished solutions.

Here are a few specific issues:

1. **Notation.** Any notation that is not defined in the statement of the problem or given by convention must be explained in the solution. Use the same notation in the solution as in the statement of the problem. Every symbol in the solution should have exactly one meaning. Avoid using different symbols to represent the same quantity.

Avoid fussy notation. Subscripted notation is useful when you have to deal with terms in a sequence or a variable is determined by a parameter. Avoid using words as subscripts. However, novices often overuse subscripts and use subscripted notation when the use of single letters is clearer. For example, rather than letting x_1 and x_2 be two arbitrary elements of a set, consider using x and y.

2. **Appealing to formulae.** It is generally better to refer to a formula by name rather than quoting it in algebraic terms, especially if the notation in the formula conflicts with the notation in the problem. Often, you can set up a calculation in such a way that it is clear what you are appealing to; this is true for example when applying the quadratic formula for the solution of a quadratic equation, Pythagoras's formula or the Cauchy-Schwarz inequality.

3. **Appealing to results.** Avoid idiosyncratic designations of reasons; not everyone will know what the "Rockcrusher theorem" is. Don't try to show off your erudition by referring to obscure results; if the marker cannot verify it or look it up, then you may be out of luck. Generally speaking, results appealed to should be more elementary than the problem. (Do you really need Dirichlet's result about primes in certain arithmetic progression to do that contest problem about primes?) Stick to standard names and designations; the use of ASA, SSS and SAS is acceptable for justifying congruency of triangles. When referring to similar or congruent triangles, it is helpful to the reader if you list the vertices in the order of the similarity or congruence.

When writing international competitions, note that a result may be known by different names in different parts of the world. Consider in this situation giving a brief description of the result in parentheses or as a footnote.

"Two-column proofs" where each assertion is numbered and a justification reference is put in a second column are often very hard to read, because they are generally tedious and it is hard to get the thread of the solution and a clear sense of the logic. This is a situation where it is better to use paragraphs and connect related statements.

4. **Mechanical work.** *General issues.* It is a matter of judgment how much mechanical work should be shown. You want the reader to be able to follow your manipulations without being mired in a morass of detail. The general rule is that the reader should understand how you are proceeding and be able to easily verify the final result. While you may go through contortions to arrive at an equation, you should not reproduce these steps in your solution if the equation can be verified in a straightforward way. It is frequently helpful to say one or two words to describe the activity ("squaring both sides and transferring terms to the left side, we obtain").

Make sure that there is a logical flow from left to right. If $A = C$ because A can be manipulated to B and B manipulated to C, present this as $A = B = C$ rather than $A = C = B$.

Equations and inequalities. In dealing with equations or inequalities, there are essentially three ways in which one may lay out the solution. Suppose, for example, that one wishes to show $A \leq B$. Here are the options:

(i) Present it as a chain of equivalent inequalities:

$$A \leq B \Leftrightarrow C \leq D \Leftrightarrow \cdots \Leftrightarrow Y \leq Z \ ,$$

where the last is clearly correct and where A can be manipulated to C, B to D, *etc.*. This should in general be avoided, particularly when each member of an inequality is a reworking of the corresponding member of the previous inequality. This often entails a lot of tedious repetition and runs the risk of a logical error as you proceed from what needs to be shown to something that is known to be true. Not every logical step may be reversible.

However, there are situations in which this approach is desirable in setting up an inequality in a tractable way. For example, the given inequality may contain surds and you might want to square both sides to get an equivalent inequality to analyze. Or it may involve fractions and you might wish to clear the denominators. Or you may want to take the logarithms or exponents of the two sides. In such cases, the process $A \leq B \Leftrightarrow C \leq D$ at the outset may be the prelude to a clear and convenient presentation. However, check that the reasoning is reversible (each inequality implies the other) and that, where a multiplication or division has occurred, whether the sense of the inequality remains the same. There may be other restrictions under which an equivalence of two pairs of equations is valid that you should keep in mind, for example when squaring or taking a logarithm. For equations, you should check to make sure that extraneous solutions are identified.

(ii) Work from one side of the inequality to the other. In this case, one manipulates one side of the inequality until an inequality can be invoked to give a term leading to the other side:

$$A = C = E = \cdots = Y \le Z = \cdots = F = D = B \ .$$

This works when it is just a matter of replacing an expression by an equivalent expression and has the advantage of providing a logical and mechanical flow to carry the reader along.

(iii) Take the difference. Consider looking at $B - A$ and try to show that this is not less than 0 (or equal to 0 in the case of an equation). This is often the best approach. It allows you to factor or rearrange terms (for example, to form a sum of squares) and shortens the solution considerably in some cases.

Be careful when dealing with equations and inequalities if you multiply or divide by anything. Consider the possibility that a divisor may vanish, and perhaps peel this off as a separate case. With an inequality, you must note that, when the divisor is negative, the sense of the inequality is reversed. Check carefully the sense of the inequality if you take reciprocals.

5. **Work smarter rather than harder.** Try to avoid situations where you have to consider a great many cases or get involved in heavy computations that are hard to keep accurate. The first is often a risk in a combinatorial problem or in solving inequalities or equations, the second in a geometry problem where an analytic geometry solution is attempted. Look for other approaches first. An analytic geometry argument should be regarded as a last resort. However, there are situations where it can be quite efficient, especially if you make a wise selection of coordinates. Try to avoid using unneeded variables. A geometry result does not generally depend on the scale of the diagram, so you might use convenient numbers for the coordinates. Here is an example of a problem that illustrates the effective use of analytic geometry:

Problem. ABC is a triangle and M the midpoint of side BC. Squares $ACDE$ and $BAFG$ are erected on the outer side of sides AC and BA (with the vertices E and F adjacent to A). Prove that the lines AM and EF are perpendicular.

Comments on the solution. Taking $A \sim (a_1, a_2)$, $B \sim (b_1, b_2)$, *etc.*, leads to a truly horrible solution. Noting that scale is not an issue, try taking $M \sim (0,0)$, $B \sim (-1,0)$ and $C \sim (1,0)$. Since A can be anywhere, let $A \sim (a,b)$. Then $E \sim (a+b, b+1-a)$ and $F \sim (a-b, b+1+a)$ and the slopes can be readily checked.

However, there is an even easier analytic geometry argument even through the number of variables is increased. Put the origin at A and let $B \sim (a,b)$ and $C \sim (u,v)$. This has the advantage that the points E and F are quickly located as is the slope of AM: $E \sim (-v,u)$, $F \sim (b,-a)$, $M \sim (\frac{1}{2}(a+u), \frac{1}{2}(b+v))$, and the slopes of AM and EF are respectively $(b+v)/(a+u)$ and $(-a-u)/(b+v)$.

A slight modification of this gives a third argument. Taking the origin at A as before, but letting $M \sim (0,-1)$, we can take $B \sim (-s, -1-t)$ and $C \sim (s, -1+t)$. Then $E \sim (1-t, s)$ and $F \sim (-1-t, s)$, from which it is clear that EF is horizontal and thus perpendicular to AM.

Very often, the use of analytic geometry can be "hidden" in an argument using complex numbers. One advantage of complex numbers is that we get more than just a vectorial approach; multplication by a complex number z corresponds to a rotation through the argument of z followed by a dilation of factor $|z|$. In particular, multiplication by i corresponds to a counterclockwise rotation through $90°$.

Corresponding to the first argument above, we let M, B and C corresponds to the numbers 0, -1 and 1 in the argand plane, and let A be positioned at z. Then E is positioned at $z+i(1-z)$ and F at $z+i(1+z)$. The complex number corresponding to the vector $\overrightarrow{FE}$ is $2zi$, so that EF is found by rotating MA through an angle of $90°$ counterclockwise about M, doubling its size and translating into position.

Corresponding to the second argument, we let A be located at 0, B at z and C at w. Then E is located at iw and F is located at $-iz$, and we are again led to the solution.

The third argument looks like this using complex numbers. Take A at 0 and M at $-i$, with B and C at $-i-z$ and $-i+z$, respectively. Then E is at $-i(i+z)=1-iz$ and F is at $-i(-i+z)=-1-iz$. Then the vector $\overrightarrow{EF}$ corresponds to the real number -2, and so is perpendicular to and twice the length of AM. ♠

There have been some notorious instances of students making heavy weather of a problem with an elementary solution. V. Pandelieva recounts a case of a student on a 4-hour Bulgarian national contest who solved a Diophantine equation by systematically working out 96 different cases. In the end, he got all the solutions and therefore qualified for a full mark. However, the marker was likely less than happy to read through all the material. In the 1978 Putnam competition, the first (socalled "starter" problem) was to show that, if 20 numbers were selected from among the terms of the finite arithmetic progression $\{1, 4, 7, \cdots, 100\}$, then two of them would add up to 104. This was intended to be a simple pigeonhole problem, but many students did not see it that way and produced long and involved solutions with many cases that threatened to bog down the marking completely. For the first problem on a contest, it is reasonable to expect a short and reasonably straightforward solution; think about it.

It is highly likely on a contest that the setters have in mind a reasonable insightful solution to a problem, and getting into heavy work should be a last resort. Apart from anything else, it eats up a lot of time that might be more productively used.

6. **Geometric constructions.** There are four steps to solving a geometric construction problem: (a) giving an analysis; (b) describing the construction; (c) proving that the construction delivers what is needed; (d) checking feasibility, *i.e.*, conditions under which the construction can be carried out. A complete solution for marking requires that (b) and (c) be submitted, and you should make sure that these two steps are clearly delineated. Often, the problem is such that you should deal with feasibility as well, so it is wise to be on the safe side and do this, if convenient.

An analysis often consists of assuming that the construction has been implemented and then working back to elements upon which it depends and which can be easily found from the given conditions. This is a natural way to start a construction problem, but carries the risk of working backwards and missing the import of some logical step. Some students present their analysis as the construction, which can create difficulties for the marker who has to track back to ensure that the construction can indeed be carried out from the given situation. Others appeal to the analysis as the proof of the construction, which again carries the risks of working backwards. This should be avoided. However, including an analysis, clearly marked as such, with your solution is not a bad idea, as it gives the marker some idea of where you are coming from.

7. **Range of parameters.** There are problems, often inequalities, where something has to be established over a range of parameters or the maximum or minimum value of some parameter needs to be found for something to work. Be careful to check that there is an example that correspondends to extreme values of a parameter.

8. **Logic.** Normally, the logic of your solution should go forward from what is given or known to what has to be established. Avoid reasoning backward from the solutions in your write-up. (This is often the way you arrive at the solution.) Avoid an extended use of "if we can show this, then this happens" which leads to a convoluted winding back and forth of the logic.

Many students are addicted to contradiction arguments; often such an argument can be avoided by a direct argument. Check to see whether this is possible for your solution. For example, you may have to show that a certain condition implies that $x = y$. Rather than start with "suppose $x \neq y$" and proceeding to $x = y$ as a contradiction, check whether your argument goes through directly from the condition and leads to the conclusion $x = y$. However, there are situations where a contradiction argument cannot be conveniently avoided, and attempts to do so will muddy the waters. This is a judgment that will come with experience.

Logical relationships between statements must always be indicated. These include the use of such words and symbols as:
"therefore", "implies", "it follows that", "whence", "hence". $\Longrightarrow$,
"since", "because", "from"
"is implied by", $\Longleftarrow$
"is equivalent to", $\Longleftrightarrow$

To prove that a general statement is false, it is necessary only to come up with one counterexample that satisfies the hypotheses but not the conclusion. To prove that a statement is possible, you need to find just one example that works. To prove that a general statement is true, try to think of a broad reason that might justify it so that you reduce the number of cases that have to be considered as much as possible. Checking many options is risky, because it is easy to leave some cases out. To prove that a statement is impossible, try to think of a general reason. You need to show that it *never* holds, and this cannot be done simply by verifying its impossibility in a number of special cases that do not cover all situations.

9. **Answer the question!** At the start, read the question carefully and make sure that you understand what is required. When you have solved the problem, reread the question and make sure that you have a concluding statement in your argument that answers the question. On a contest, if an answer in the form of a number or an example is asked for, it is a good strategy to put the answer at the front of your solution, and then follow it by the justification. This will alert the marker that you have indeed solved the problem and indicate that it is worthwhile to wade through your work. If, alas, the answer is not right, it will still arouse the curiosity of the marker to figure out where you came unstuck and at least allow for the awarding of part marks.

10. **Missteps.** It is not uncommon to head off in a wrong direction and have to revise your solution. When this occurs, you must indicate clearly to the reader what you do not wish to have considered. Box the rejected material and gently put a large X through it. However, do not obliterate it. If by any chance your purported correct solution has something wrong with it, the examiner may want to look at the rejected material to see if anything can be salvaged. It occasionally happens that the student was on the right track to begin with and got spooked on the way.

11. **Improving the odds.** You want to make sure that the reader of your solution does not miss anything and that you get all the credit that is due to you. **Always** start each new solution on a new page and **never** answer two or more problems on the same page or mix the solution of one problem with another. Number your solutions. Problems do not have to be solved in order. On contests, there is frequently a different marker for each problem and the scripts may have to be sent over long distances. Space your solutions reasonably, and leave a space between paragraphs. Do not challenge the possibly failing eyesight of the marker too severely. If you use a pencil, make sure that the lead is neither too

hard (so it hardly makes a mark) nor too soft (so that it smudges). If you use a pen, pick one of sufficient quality that the ink does not smudge or go through the paper. If your handwriting is unclear or loopy, consider writing on every other line. Get in the habit of writing symbols so that they cannot be confused with others. For example, printing and crossing the letter "z" distinguishes it from a "2" or "3". Printing an "x" as two tangent semicircles and crossing it distinguishes it from "times" ($\times$). Make sure that your "q" does not look like "9" (a serif on the bottom of the "q" will achieve this), and that your "u" can be distinguished from "v". Writing rather than printing a lower-case "l" (ell) will distinguish it from "1". Europeans often cross a "7" to distinguish it from "1". Don't let your solutions be a morass of symbolism, so that the reader feels as though there is some kind of code that has to be deciphered. Try to put everything that has to do with the same problem on contiguous pages; if this cannot be done, indicate to the marker where other work on the same problem can be found. Make sure that you have filled out everything needed on the cover page of the solutions - especially your name or examination number. If you are writing in examination books, indicate on the cover of the first book how many books you have used and number the books. If you have loose pages, identify each with your initials or examination number, as well as with the number of the problem.

Olymon and the Olympiads. Qualification for the International Mathematical Olympiad is based primarily on performance in national and international competitions: the Asian-Pacific Mathematical Olympiad, the Canadian Mathematical Olympiad, the USA Mathematical Olympiad. When a decision on the IMO team is difficult, other factors may come into play, in particular how students fared at the various training camps, in particular the winter training camp preceding the IMO. The purpose of the Mathematics Olympiads Correspondence Program (Olymon) is to give students the opportunity to write up solutions to problems and have them graded. The program is also useful in helping the Canadian Mathematical Society identify students who might be invited to regional training camps or encouraged to take part in the Open competition in November. It may be used in conjunction with other information to help narrow down the selection for the winter training camp or the IMO team. It is a good idea to submit Olymon solutions during the early fall, so that students can be identified and encouraged to write the Canadian Open Competition in November. A good grade in the Open is the key that opens the door to later opportunities.

No claim is made that the Olymon problems are original: after all, if a nice problem were created, it would be submitted for consideration on a contest – good new problems are hard to come by. Thus, a zealous student could undoubtedly find Olymon problems on the net. However, this defeats the purpose of the Olymon, and does no one any good. To obtain benefit, students must try the problems on their own and acknowledge any outside help received. Integrity demands this. There are few, if any, qualities more important than integrity in the choice of the IMO team. A complete collection of Olymon problems can be found on the website <*www.math.utoronto.ca/barbeau/*>. The individual Olymon sets and their solutions are also found on the website <*www.cms.math.ca*>. You should try the problems and check your solutions against those provided. These solutions, many due to students, can be used as models to help you improve your presentation.

Ed Barbeau