Designing the proof of Vizing's Theorem

by Josyula R. Rao & Edsger W. Dijkstra

We shall design the proof of the following theorem, due to V.G. Vizing.

Theorem For a finite undirected graph without autoloops and without multiple edges, at any vertex of which fewer than N edges meet, N colours suffice for an edge colouring such that edges incident on the same vertex are of different colour.

* * *

The graph being acceptably coloured is expressed by

(0) $(\underline{A}V{::}\ acc.V)$

where predicate acc on vertices is given by

acc.V ≡ (no two edges incident on vertex V have the same colour) .

We shall prove Vizing's Theorem constructively by showing how the entire graph can be acceptably coloured. To this end it suffices to design a procedure that, given an acceptably coloured graph with one uncoloured edge, constructs an acceptable colouring in which this last edge is coloured as well. Repeated application of this procedure, starting with a graph consisting of vertices only (whose colouring is trivially acceptable) and adding one uncoloured edge at a time, colours the whole graph acceptably. We shall design the procedure for colouring the next uncoloured edge.

In the rest of this note, we confine ourselves to acceptably coloured graphs. In order not to interrupt the subsequent development of the procedure too much, we first introduce a concept, the need of which will emerge, viz. the <u>alternating path</u>. An alternating path is a maximal path of at most two given colours. Because the colouring of the graph is acceptable, those two colours alternate along such a path; hence the name. One can show the following:

(i) an alternating path is either a cycle (i.e. without end points) or a simple path with two end points. (In order to avoid case analysis, we allow the end points to coincide, in which case the alternating path has length 0.)

(ii) a pair of colours and a vertex determine a unique alternating path of those colours and through that vertex. (We allow the two given colours to be equal, in which case the length of the alternating path is at most 1 edge.)

(iii) swapping the colours of the edges of an alternating path leaves the colouring acceptable (because alternating paths are maximal, i.e. have no end points where they can be extended); its significance is that it provides a way for changing the sets of colours at the end points while leaving the colouring acceptable.

Before the procedure starts, we define for each vertex V a colour $c.V$, in terms of which we define predicate f given by

$$f.V \equiv (\text{no edge incident on } V \text{ has colour } c.V).$$

The fact that the number of colours is higher than the number of edges meeting at any vertex is exploited by choosing c in such a way that initially $(\underline{A}V :: f.V)$ holds.

In the following algorithm for the procedure, X, Y, and Z are variables of type vertex. The procedure's functional specification in terms of pre- and postcondition is given in full; intermediate assertions are only named and will be determined later. Explanation and missing definitions will be given afterwards, when we design the algorithm. (We give this text now for future reference. At this stage, the reader should not try to understand this program fragment.)

{ ($\underline{A}$V:: acc.V) $\wedge$ ($\underline{A}$V:: f.V) $\wedge$
(XY is the only uncoloured edge)}
{P0 : invariant}
<u>do</u> the c.Y-path ends in Y $\rightarrow$ {P1}
determine Z so that edge
XZ has colour c.Y {P2}
; give edge XY colour c.Y and
uncolour edge XZ {P3}
; Y:= Z {P0}
<u>od</u> {P4}
; swap the colours along the c.Y-path {P5}
; give edge XY colour c.Y
{ ($\underline{A}$V:: acc.V) $\wedge$ (all edges are coloured)} .

The term

(0) ($\underline{A}$V:: acc.V)

is a conjunct of all intermediate assertions; the term

(1) (XY is the only uncoloured edge)

is a conjunct of all, except P3 . Note that, in view of the absence of autoloops, (1) implies that X and Y are different vertices.

We shall now show how the above algorithm can be designed. The leading principle is to keep things as simple as possible and not to introduce complications unless forced to do so. In particular: the simplest way, if possible, of asserting a term that has been asserted before is to maintain it in between.

We now start the design of the algorithm, beginning with its last statement. Because (0) – i.e. $(\underline{A}V:: acc.V)$ – occurs in both pre- and post-condition, we decide – see previous paragraph – to maintain (0) all through the algorithm. Consequently, the final act of the algorithm is to colour the uncoloured edge with an acceptable colour. Let, as initially, the uncoloured edge be XY; colouring it while maintaining (0) means that has to be coloured with a colour that is incident on neither X nor Y. Initially, since $f.X \wedge f.Y$ holds, $c.X$ is not incident on X and $c.Y$ is not incident on Y. Somewhat asymmetrically we decide to give edge XY the colour $c.Y$. This choice of last statement yields – note that in view of the definition of f

$$f.Y \equiv (c.Y \text{ not incident on } Y) \text{ –}$$

for its precondition $P5$:

$P5$: $(0) \wedge (1) \wedge (c.Y \text{ not incident on } X) \wedge f.Y$.

We now consider $P5$ as the postcondition to be established. Because $(0) \wedge (1) \wedge f.Y$ is implied by the precondition, we concentrate our attention on establishing P5's conjunct

$$(c.Y \text{ not incident on } X)$$

while maintaining its three other conjuncts. If an edge incident on X has colour c.Y, its colour has to be replaced by a colour not incident on X. Since initially f.X holds, (i.e. colour c.X not incident on X), we propose, in view of (iii), to swap the colours c.X and c.Y along the alternating path through X that has those colours. This ensures that c.Y is no longer incident on X.

Since we need the concept a number of times, we define, for graphs satisfying $(\underline{A}V :: acc.V) \wedge f.X$ and for any colour p, the <u>p-path</u> to be the alternating path through X with colours p and c.X. Due to f.X, i.e. c.X not incident on X,

- the p-path with p = c.X is empty
- a non-empty p-path starts at X with an edge of colour p and ends at a vertex different from X.

In the above terminology, we propose to establish (c.Y not incident on X) by swapping the colours along the c.Y-path; this maintains $(0) \wedge (1)$, but maintains f.Y <u>only</u> provided the c.Y-path does <u>not</u> end in Y. (Because of f.Y, a c.Y-path ending at Y does so with an edge of colour c.X ; changing that colour into c.Y would falsify f.Y .) Collecting our requirements, we get as precondition P4 for the colour swap along the c.Y-path

P4: (0) $\wedge$ (1) $\wedge$ f.X $\wedge$ f.Y $\wedge$
(the c.Y-path does not end in Y) .

We now consider P4 as the postcondition to be established. Observing that the first four conjuncts of P4 are implied by the precondition, we apply the same heuristic principle as before and decide to look for a statement that establishes the last conjunct of P4 while maintaining the others. The following methodological interlude explains why we look for a statement of the form of a repetition.

<u>Methodological Interlude</u> Suppose we look for a program S that satisfies for some P, Q

$$\{P\}\ S\ \{P \wedge Q\} \quad .$$

In those initial states in which Q holds, S may act as a skip. Sometimes this requires no explicit measures. (We saw an example of this in the statement we derived earlier: swapping the colours along an alternating path automatically reduces to a skip if the path is empty.) Otherwise we need an explicit guard in order to distinguish between the cases Q and $\neg Q$. Unless an S of the form

$$\underline{\mathrm{if}}\ Q \rightarrow \mathrm{skip}\ [\!]\ \neg Q \rightarrow D\ \underline{\mathrm{fi}}$$

obviously does the job, it is wiser to begin looking for an S of the form

$$\underline{\mathrm{do}}\ \neg Q \rightarrow D\ \underline{\mathrm{od}} \quad ,$$

since this allows us to separate the concerns for partial correctness and termination. If, while studying termination, we learn that D is executed at most once because it establishes Q, we are still free to replace the repetition by the alternative construct. (End of Methodological Interlude.)

The repeatable statement we are looking for should have the potential of falsifying the guard

$$(\text{the c.Y-path ends in } Y) \quad .$$

Because of an example in which it is impossible to falsify this guard by only changing c and not changing the colouring of the edges, and because, for simplicity's sake, we would not like to change both, we decide to try to keep c constant and let the repeatable statement affect the edge colouring only.

The next constraint we adopt is that the repeatable statement is confined to the combination of colouring the uncoloured edge and uncolouring a coloured one. This is inspired by the observation that this is the simplest transformation whose repeated application can transform a colouring with one uncoloured edge into any other such colouring. It is a constraint —and therefore not adopted without optimism— because, in addition, we intend to maintain at each step

$$(0) \wedge (1) \wedge f.X \wedge f.Y \quad .$$

As asymmetrically as before, we decide to give the uncoloured edge the colour $c.Y$; on account of the initial validity of $f.Y$, this colouring act maintains $acc.Y$. In order to maintain $acc.X$ as well, the first edge of the $c.Y$-path is uncoloured. Thus we have designed the repeatable statement given above. (The variable Z has primarily been introduced to ease the discussion.) Our gamble is that we can prove invariance and termination. Let us concentrate on the invariance first.

As we shall see shortly, all is available for the invariance of $(0) \wedge (1) \wedge f.X$; the invariance of $f.Y$, however, poses a problem. The invariance of $f.Y$ requires the conjunct $f.Z$

in P3 , and the simplest way of justifying it there is by requiring $f.Z$ as conjunct in P2 . (Besides this being the simplest way, our termination argument, as we shall see later, relies on the fact that no false f-value is truthified.) In order to justify $f.Z$ in P2 the invariant has to be strengthened with a new conjunct; $(\underline{A}V :: f.V)$ would justify $f.Z$ but is too strong to be maintained. We could also justify $f.Z$ from

for any colour p such that the p-path is not empty:

$$f.S \quad ,$$

where S is the second vertex on the p-path .

This, however, is also too strong: the repeatable statement transforms the p-path $XZ...Y$ into the p-path $XY...Z$ with XY of colour $c.Y$, so $\neg f.S$ for $p = c.Y$. The following slightly weaker conjunct does the job:

(2) for any colour p such that the p-path is not empty:

$$f.S \vee c.L \neq p \quad ,$$

where S is the second vertex of the p-path and L its last vertex .

Now the time has come to list and justify assertions P0 through P3 .

P0: $(0) \wedge (1) \wedge (2) \wedge f.X \wedge f.Y \wedge X \neq Y$

Assertion P0 is implied by the precondition of the procedure: its first two conjuncts occur in the precondition, the next three follow from $(\underline{A}V:: f.V)$, and $X \neq Y$ follows from (1) and the absence of autoloops. We return later to the re-establishment of P0 at the end of the repeatable statement.

P1: P0 $\wedge$ (the c.Y-path ends in Y) $\wedge$
(the c.Y-path starts at X with an edge of colour c.Y)

The first two conjuncts follow from the topology of the program; the last conjunct is a property of nonempty p-paths and the c.Y-path is nonempty because it connects the different vertices X and Y .

P2: P1 $\wedge$ (XZ has colour c.Y) $\wedge$ $X \neq Z$ $\wedge$
$f.Z \wedge c.Z \neq c.Y \wedge Y \neq Z$.

The first conjunct P1 , which does not refer to Z , is maintained and implies the existence of a unique vertex Z such that XZ has colour c.Y ; XZ is an edge and not an autoloop, hence $X \neq Z$; the conjunct f.Z follows from (2) with the instantiation $p, S, L := c.Y, Z, Y$; from f.Z and the fact that XZ has colour c.Y , we conclude $c.Z \neq c.Y$, from which, with Leibniz's Principle, $Y \neq Z$ follows.

P3: (XZ is the only uncoloured edge) $\wedge$
(XY has colour c.Y) $\wedge$ (0) $\wedge$
$f.X \wedge \neg f.Y \wedge f.Z \wedge$ (2) .

Since X,Y,Z are three distinct vertices, XY and XZ are distinct edges and the (un)colouring statement is well-defined; with (1) it establishes the first two conjuncts. For (0) and the next three conjuncts we need only consider vertices X,Y,Z since they are the only vertices with incident edges changing colour; acc.X and f.X are maintained because the bag of colours incident on X remains the same; to the bag of colours incident on Y the colour c.Y is added and, because of f.Y in P2, this maintains acc.Y and falsifies f.Y; acc.Z and f.Z are maintained because the bag of colours incident on Z decreases. For the invariance of (2), we distinguish two cases.

In the case $p \neq c.Y$, we observe that the (un)colouring statement leaves the p-path unchanged; in particular, if the p-path is not empty, its vertices S and L remain the same. The constancy of L ensures the constancy of $c.L \neq p$; for the constancy of f.S we observe that Y is the only vertex whose f-value changes, while $S \neq Y$ because XS has colour p and XY was uncoloured.

(Here we have used the absence of multiple edges.)

In the case $p = c.Y$ we observe that the p-path $XZ...Y$ is replaced by the p-path $XY...Z$, i.e. we have to demonstrate (2) for the instantiation $p, S, L := c.Y, Y, Z$, i.e.

$$f.Y \vee c.Z \neq c.Y \quad ,$$

which follows from P2. Hence P3 is justified.

The axiom of assignment tells us that P3 is for $Y := Z$ a precondition strong enough for the re-establishment of P0.

For the termination argument we observe that, the function c not being changed, f-values can only be changed by changing edge colours, i.e. by the (un)colouring statement. Comparison of P2 and P3 shows that it decreases $(\underline{N}V :: f.V)$ by 1 (by falsifying $f.Y$); hence the repetition terminates.

And this concludes our design of the proof of Vizing's Theorem.

Acknowledgements We thank R.E. Tarjan for having drawn our attention to the ugliness of the existing proofs; for constructive criticism we thank the ATAC — in particular K.L. Calvert and J. Misra — and the ETAC — in particular W.H.J.

Feijen — ; for their active assistance we thank A.J.M. van Gasteren and David Gries.

In retrospect

The above is our nth explanation of the second algorithm we designed to prove Vizing's Theorem. (This algorithm is much better than our first one, which was probably still too much influenced by the published proofs.)

We are very pleased by the modesty of what we needed: the function c , the predicates acc and f , the notion of the p-path , and three variables X, Y, Z of type vertex (of which X is a constant and Z really only a local variable of the repeatable statement).

We also have reservations, for the algorithm is probably more subtle than our heuristic explanation suggests. The conclusion that in the repeatable statement the uncoloured edge should be allowed to migrate over the graph is not surprising; the decision to restrict that migration by keeping X constant (and confining Y to range over the neighbours of X) is much less obvious. (Note that at the beginning of the repeatable statement, the situation is symmetric in X and Y: they are connected by a path of even length along which the edge colours c.Y and c.X alternate.) The only justification we can think of for the optimistic decision to keep X constant is that thus, independently of the size of the graph, the number of possible values of XY is bounded by N. The underlying hope was obviously to prevent the shape of the graph from generating a case analysis.

Moreover it is only fair to say that we worked very hard on our definitions. The a priori restriction to acceptably coloured graphs simplified the introduction of the alternating path; the restriction to graphs with f.X as well simplified the introduction of the p-path; the formulation of (2) in terms of universal quantification over the dummy p of type colour simplified its proof of invariance. (It is straightforward predicate calculus to eliminate p from (2): "For each L such that the nonempty c.L-path ends at L, etc."; in the case p = c.Y we have to compare XZ...Y with XY...Z: the end points differ, the colour that alternates with c.X remains the same. Hence the colour p, and not the end point L, should be the dummy in (2).)

We are pleased that we needed only one case analysis and no pictures; we are glad that we could avoid notations such as $\mu_{\beta\gamma_j}[a,z]$ —used by Claude Berge—, but we would like to master a more calculational style of reasoning about graphs.

Austin, 22 September 1990

Josyula R. Rao
prof. dr. Edsger W. Dijkstra
Department of Computer Sciences
The University of Texas at Austin
Austin, TX 78712-1188
USA