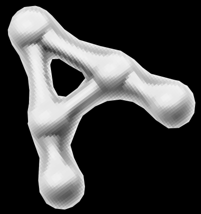
CHAPTER 2

Shortest path problem (Dijkstra's algorithm)



Learning objectives

After studying this chapter, you should be able to:

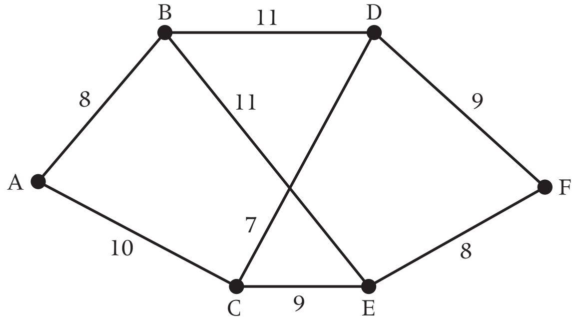
- apply Dijkstra's algorithm to a network
- trace back through a network to find a route corresponding to a shortest path
- apply Dijkstra's algorithm to a network with multiple start points
- understand the situations where Dijkstra's algorithm fails.

2.1 Introduction

When you plan a journey, there are different factors you might consider:

- do you want to go the shortest distance?
- do you want to take the minimum time?
- do you want to minimise the cost?

Autoroute is a computerised route planner. You enter the start and finish points of a journey and the program works out the different routes, depending on the criterion that is to be minimised.
Another innovation used today is satellite navigation systems. These keep a motorist in constant contact with a central computer. The central computer monitors traffic flow and updates the motorist on the optimum route from his or her current position. These systems are invaluable for haulage and coach companies.
How do these systems work?
The basic principle is to find the shortest route from one part of the diagram to another, in this case from A to F.



There are four routes from A to F:

> ABDF, length 28; ABEF, length 27; ACDF, length 27;
> ACEF, length 26

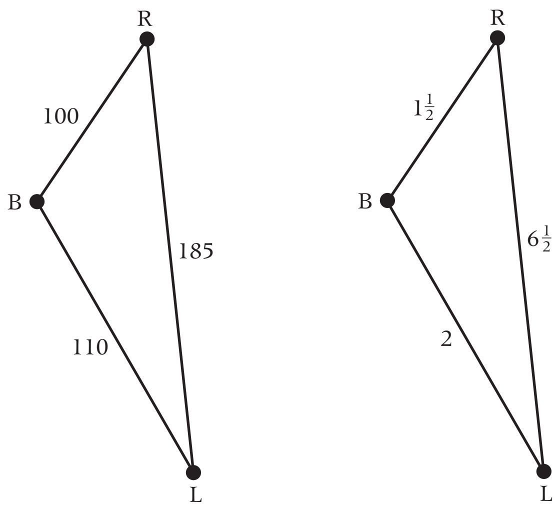
The shortest route is ACEF.

In this example there are only four possibilities to consider, but if the network were more complex then this method, called a complete enumeration, would become impractical. This chapter uses an algorithm to find the shortest path.

2.2 Triangles in networks

Consider a real-life situation in which we wish to travel from Royton to London.

The diagrams below show roads connecting Royton, Birmingham and London.



The numbers on the first diagram show the distances, in miles, between the towns.

There are two possible routes: Royton direct to London, which is 185 miles, or Royton to Birmingham and then Birmingham to London, a total distance of 210 miles.

The right-hand diagram shows the times of the same journeys. The time taken on the direct route from Royton to London is $6\frac{1}{2}$ hours. The time from Royton to Birmingham is $1\frac{1}{2}$ hours and Birmingham to London 2 hours. Hence we have a triangle with sides of length $1\frac{1}{2}$, 2 and $6\frac{1}{2}$! The right-hand diagram is an impossible triangle.

This is acceptable because the diagram shows a network representing a real-life situation and not a scale drawing.

> Real-life problems may not obey the triangle inequality.

2.3 Dijkstra's algorithm

In 1959, Edsger Dijkstra invented an algorithm for finding the shortest path through a network. The following is a simple set of instructions that enables students to follow the algorithm:

Step 1 Label the start vertex as 0.

Step 2 Box this number (permanent label).

Step 3 Label each vertex that is connected to the start vertex with its distance (temporary label).

Step 4 Box the smallest number.

Step 5 From this vertex, consider the distance to each connected vertex.

Step 6 If a distance is less than a distance already at this vertex, cross out this distance and write in the new distance. If there was no distance at the vertex, write down the new distance.

Step 7 Repeat from step 4 until the destination vertex is boxed.

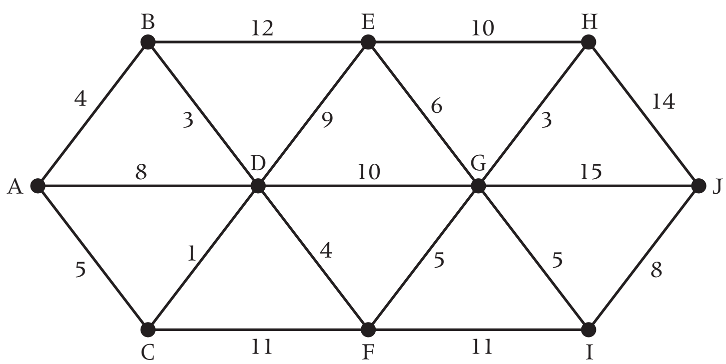
Sometimes edges of networks have arrows that show you which directions you must go from one vertex to the next. These are called **directed networks**. To apply Dijkstra's algorithm to a directed network, at each stage only consider edges that lead **from** the vertex.

Note: When a vertex is boxed you do not reconsider it. You need to show all temporary labels together with their crossings out.

If a vertex is boxed then you do not write down a new temporary value. This would be a complete enumeration.

Worked example 2.1

Find the shortest distance from A to J on the network below.

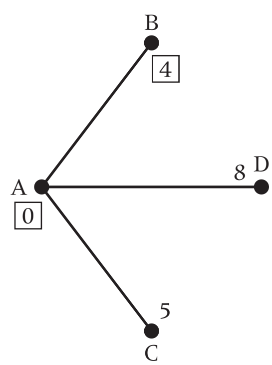


Solution

Step 1 Label A as 0.

Step 2 Box this number.

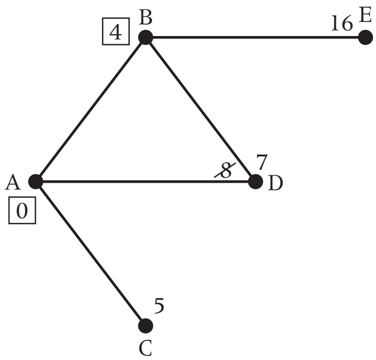
Step 3 Label values of 4 at B, 8 at D and 5 at C.



Step 4 Box the 4 at B.

Step 5 From B, the connected vertices are D and E. The distances at these vertices are 7 at D (4 + 3) and 16 at E (4 + 12).

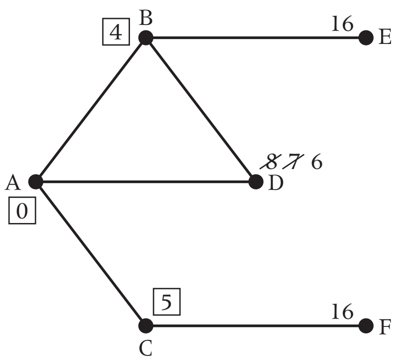
Step 6 As the distance at D is 7, lower than the 8 currently at D, cross out the 8.



Step 4 Box the smallest number, which is the 5 at C.

Step 5 From C, the connected vertices are D and F. The distances at these vertices are 6 at D (5 + 1) and 16 at F (5 + 11).

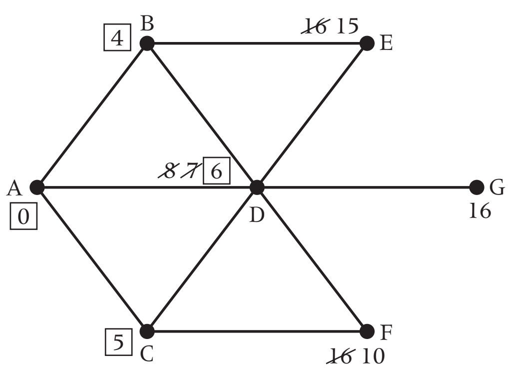
Step 6 As the distance at D is 6, lower than the 7 currently at D, cross out the 7.



Step 4 Box the smallest number, which is the 6 at D.

Step 5 From D, the connected vertices are E, F and G. The distances at these vertices are 15 at E (6 + 9), 16 at G (6 + 10) and 10 at F (6 + 4).

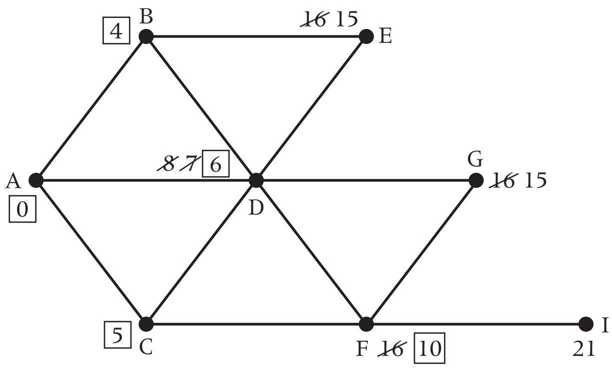
Step 6 As the distance at E is 15, lower than the 16 currently at E, cross out the 16. As the distance at F is 10, lower than the 16 currently at F, cross out the 16.



Step 4 Box the smallest number, which is the 10 at F.

Step 5 From F, the connected vertices are G and I. The distances at these vertices are 15 at G (10 + 5) and 21 at I (11 + 10).

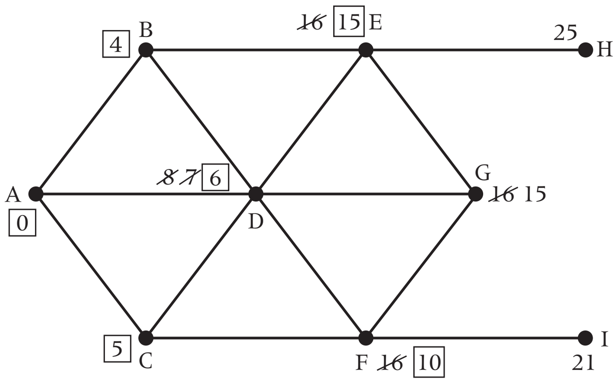
Step 6 As the distance at G is 15, lower than the 16 currently at G, cross out the 16.



Step 4 Box the smallest number, which the 15 at either E or G (it doesn't matter which you chose).

Step 5 From E, the connected vertices are H and G. The distances at these vertices are 21 at G (15 + 6) and 25 at H (15 + 10). Do not write down the value of 21 at G as this is greater than the number already there.

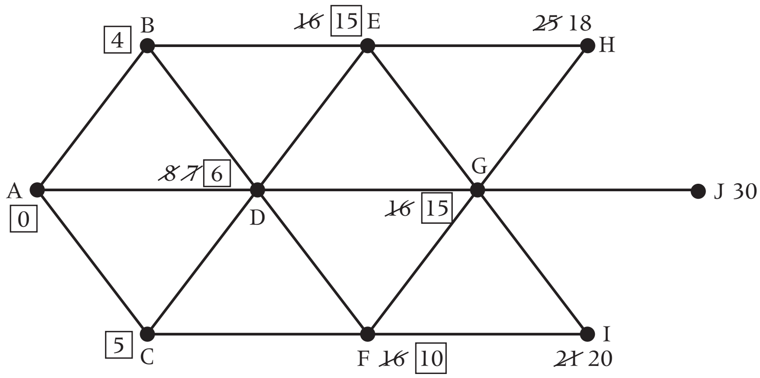
Step 6 There are no improvements, so there is no crossing out.



Step 4 Box the smallest number, which is the 15 at G.

Step 5 From G, the connected vertices are H, I and J. The distances at these vertices are 18 at H (15 + 3), 20 at I (15 + 5) and 30 at J (15 + 15).

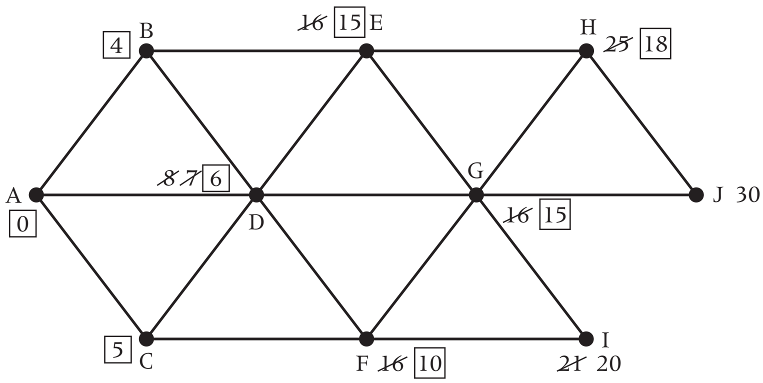
Step 6 As the distance at H is 18, lower than the 25 currently at H, cross out the 25. As the distance at I is 20, lower than the 21 currently at I, cross out the 21.



Step 4 Box the smallest number, which is the 18 at H.

Step 5 From H, the connected vertex is J. The distance at this vertex is 32 (18 + 14). Do not write down the value of 32 at J as this is greater than the 30 already there.

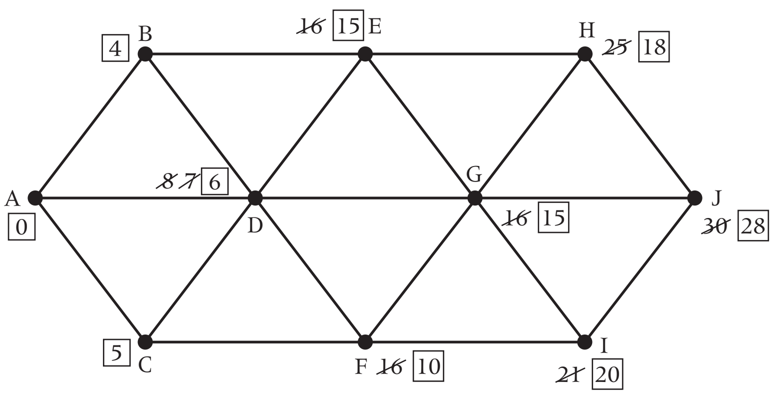
Step 6 There are no improvements, so there is no crossing out.



Step 4 Box the smallest number, which is the 20 at I.

Step 5 From I, the connected vertex is J. The distance at this vertex is 28 (20 + 8).

Step 6 As the distance at J is 28, lower than the 30 currently at J, cross out the 30.

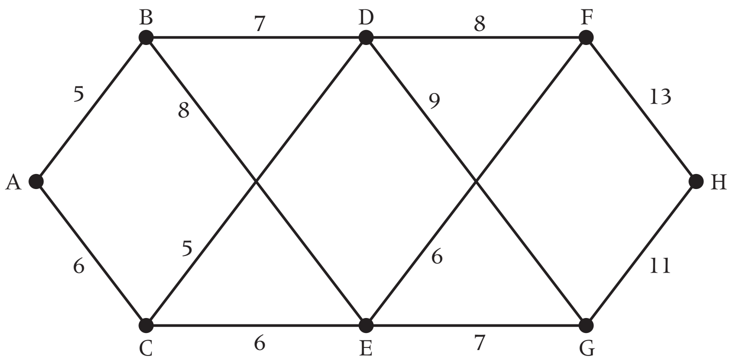


Step 7 The final vertex, in this case J, is not boxed. The boxed number at J is the shortest distance. The route corresponding to this distance of 28 is ACDFGIJ, but this is not immediately obvious from the network.

How do we retrace the route that corresponds to this shortest network? This problem will be dealt with later in the chapter.

Worked example 2.2

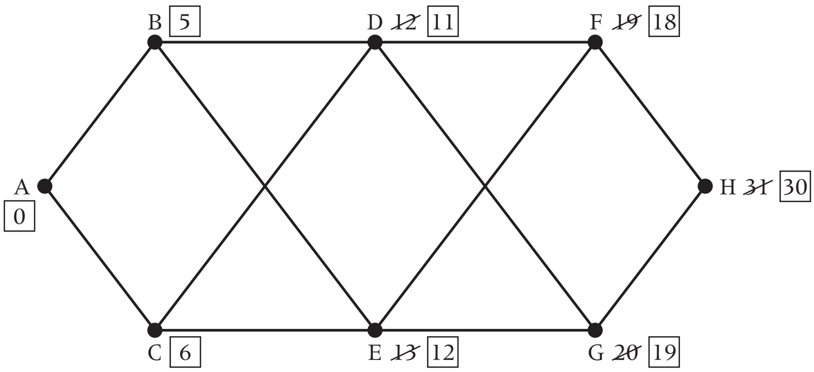
Find the shortest distance from A to H on the network below.



Solution

The fully labelled diagram below shows the values, both temporary and permanent, at each vertex.

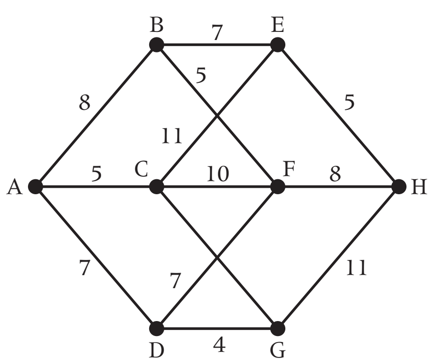
Any answer to an exam question must have exactly the same amount of detail as shown here.



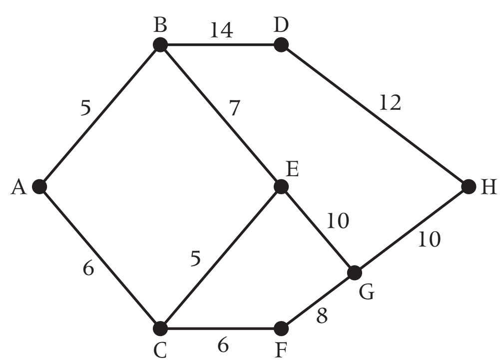
EXERCISE 2A

1 Use Dijkstra's algorithm on the networks below to find the shortest distance from A to H.

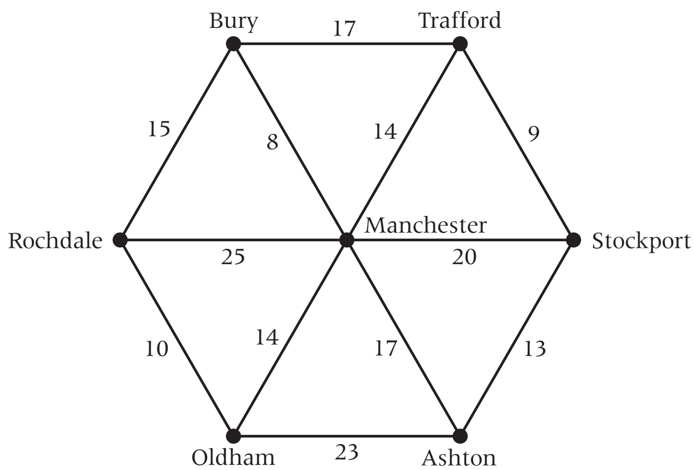
(a)



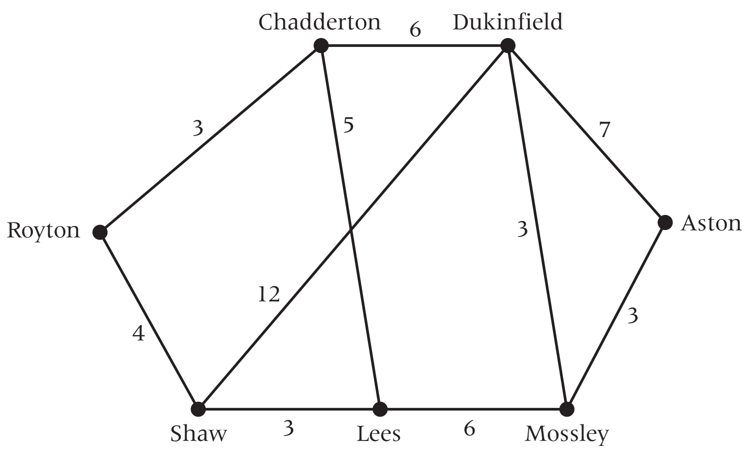
(b)



2 The diagram below shows roads connecting towns near to Rochdale. The numbers on each arc represent the time, in minutes, required to travel along each road. Peter is delivering books from his base at Rochdale to Stockport. Use Dijkstra's algorithm to find the minimum time for Peter's journey.



3 The diagram below shows roads connecting villages near to Royton. The numbers on each arc represent the distance, in miles, along each road. Leon lives in Royton and works in Ashton. Use Dijkstra's algorithm to find the minimum distance for Leon's journey to work.



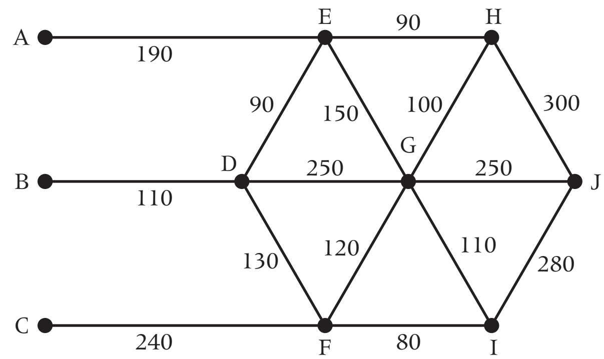
2.4 Multiple start points

How do we cope with a situation in which instead of having a single starting point and a single end point there are multiple start points and a single end point? For example, in the Monte Carlo rally cars start from a number of different countries but all end up at the same finishing point. Dijkstra's algorithm gives a method of finding the shortest distance, as in the previous example from A to J, but it is an identical problem to find the shortest distance from J to A.

> If there are multiple start points, then we apply Dijkstra's algorithm from the end point until we have reached each of the starting points. In this way we can find the shortest route.

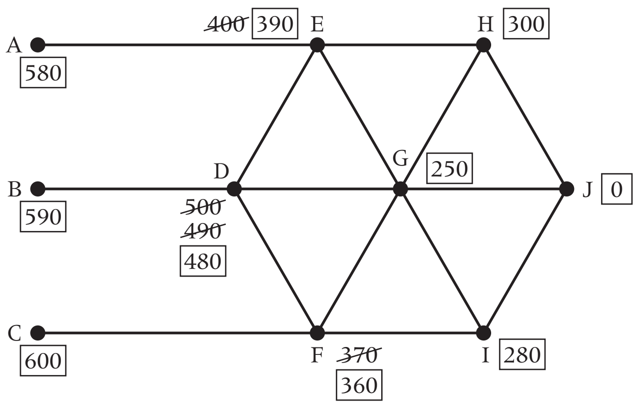
Worked example 2.3

Three boys walk to school, J, from their homes at A, B and C. The diagram shows the network of roads near their homes and school.
The numbers on each arc represent the distance, in metres, along each road. Use Dijkstra's algorithm from J to find which boy lives nearest to the school.



Solution

If we were to apply Dijkstra's algorithm in the normal way, the workings would be difficult to follow. There would be temporary and permanent labels from each of the three starting points. Working backwards from J, and applying the standard algorithm, the worked solution below is obtained.

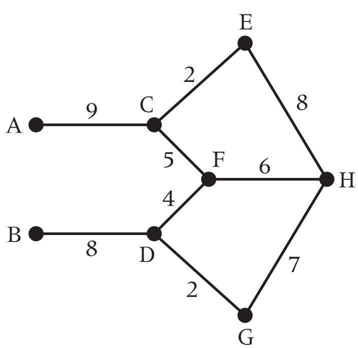


The boy living in house A lives closest to the school.

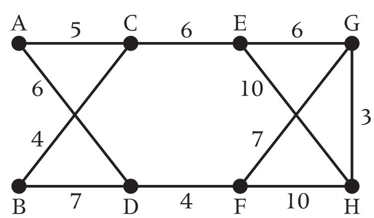
EXERCISE 2B

1 Use Dijkstra's algorithm, starting from H, on the networks below to find which of the vertices A or B is nearer to H.

(a)



(b)



2.5 Limitations

Whilst Dijkstra's algorithm gives an exact method of finding the shortest distance connecting two points, it doesn't work if any of the edges have a negative value. How can you have an edge of negative length? If we are talking about distances then obviously something like this is impossible, but consider the following situation.

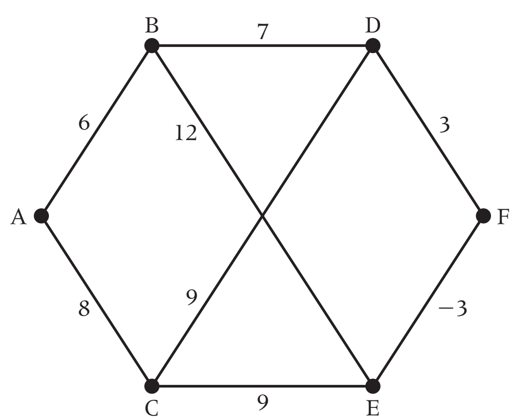
A catalogue delivery firm calculates the cost of delivering along a variety of routes. Each of the distances is represented by an actual cost. However, if whilst on a delivery a delivery van also makes a collection, then the value of making this collection en route may save more money than it would normally cost by driving along this road. Hence you may have an edge that has a negative value.

> If we use Dijkstra's algorithm on a network containing an edge that has a negative value it does not work.

If we accept that negative lengths are possible, then it will become immediately obvious that Dijkstra's algorithm won't work because using the algorithm we don't revisit anywhere that has already been boxed, yet if we come to a vertex then from there a negative length may produce a value that is shorter than one that is already boxed. In such situations, where Dijkstra's algorithm fails, the normal approach is to use dynamic programming (see D2).

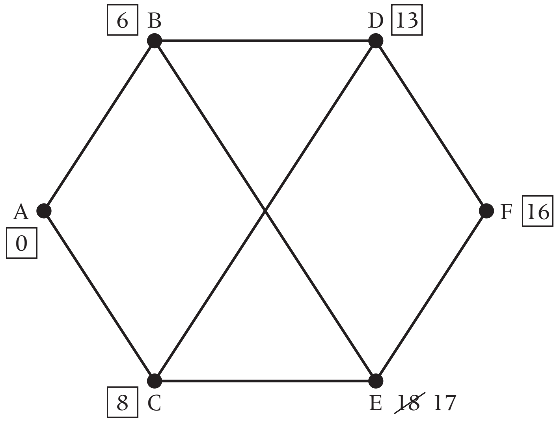
Worked example 2.4

The following network shows the cost, in pounds, of travelling along a series of roads. Use Dijkstra's algorithm to find the minimum cost of travelling from A to F.



Solution

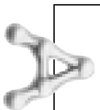
Working through the network in the standard way, the worked diagram below is obtained.



At F, the figure of 16 is boxed and hence the question is finished. However, from E to F a value of −3 reduces the minimum cost of getting to F to 14!

2.6 Finding the route

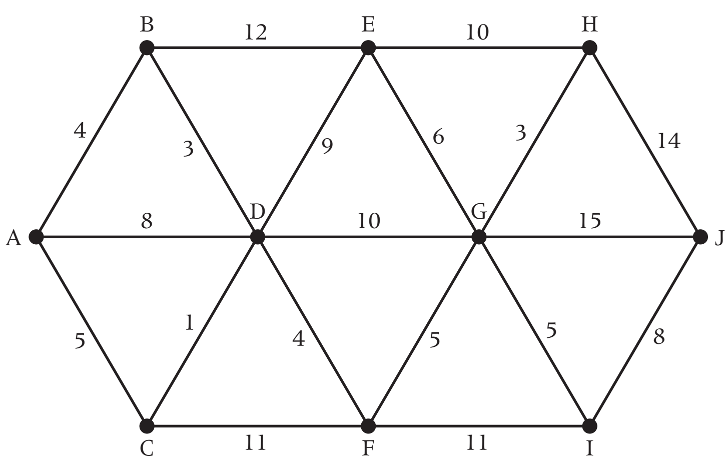
We can use Dijkstra's algorithm to find the length of the shortest route between two vertices in a network. However, if we want to find the corresponding route, we need to record more information as we apply the algorithm.



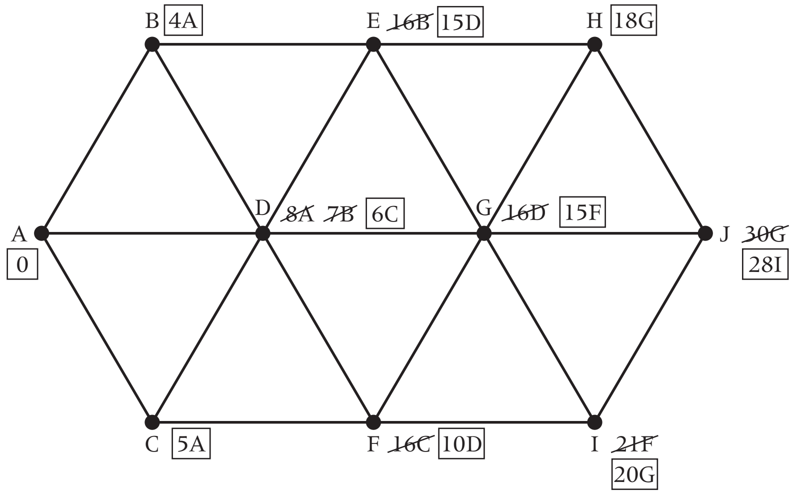
> Instead of listing temporary values, we put a letter after each value, which indicates the preceding vertex on the route. We find the route by backtracking through the network from the finishing point.

Worked example 2.5

The network below was used in Worked example 1.



Solution



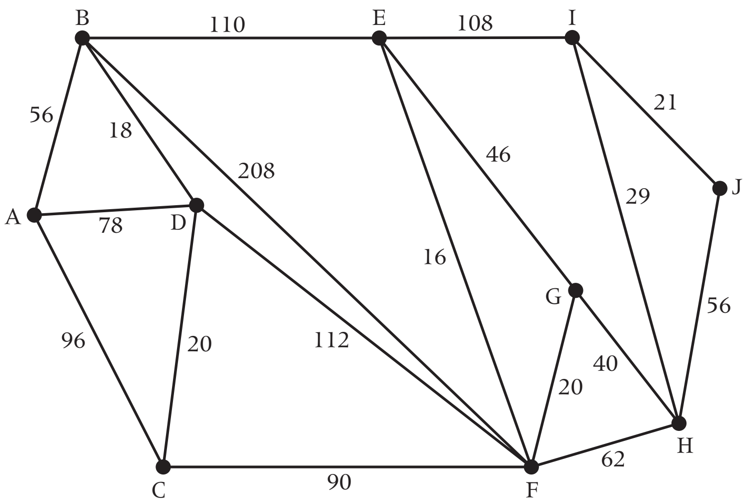
Box A has a value of zero.
At B, the distance from A is 4 so we now write 4A.
At D, the distance from A is 8 so we write 8A and at C we write 5A.
We then box the value of 4 at B, so 4A is now boxed.
From B, the value at E becomes 16B. At D we get 7B.
Box the smallest number, which is the value of 5A at C, so 5A is boxed at C.
From there, there is a value of 16C at F and a value of 6C at D.
We box the smallest value, which is 6C at D.
At D there is 8A crossed out, 7B crossed out and 6C, which has been boxed. This tells us that to get to D the smallest distance is 6 and we came from vertex C.
Working from the finishing point J we have a boxed value of 28I so we now look at vertex I. Here the boxed value is 20G so we now look at vertex G, and so on until we return to A. Hence the shortest path is ACDFGIJ, with length 28.

EXERCISE 2C

Repeat Exercises 2A and 2B to find the routes that correspond to the minimum distances.

MIXED EXERCISE

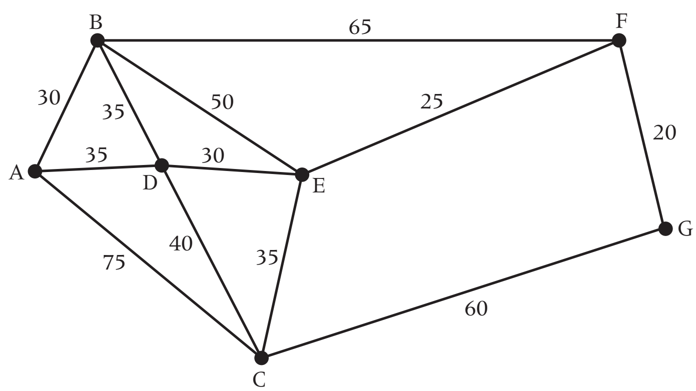
1



(a) Use Dijkstra's algorithm on the above diagram to find the minimum time to travel from A to J, and state the route.

(b) A new road is to be constructed connecting B to G. Find the time needed for travelling this section of road if the overall minimum journey time to travel from A to J is reduced by 10 minutes. State the new route. [A]

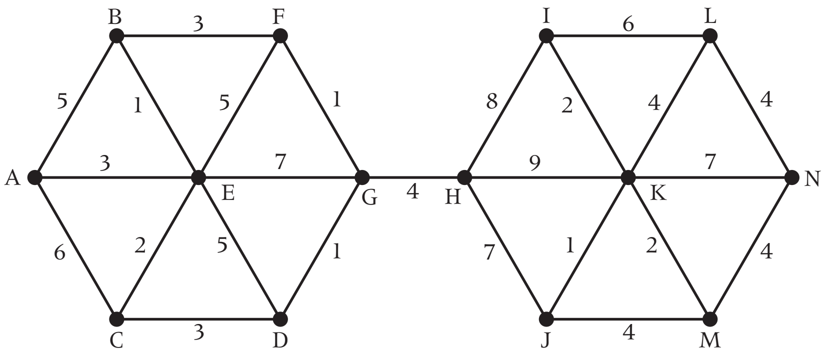
2 The following network shows the time, in minutes, of train journeys between seven stations.



(a) Given that there is no time delay in passing through a station, use Dijkstra's algorithm to find the shortest time to travel from A to G.

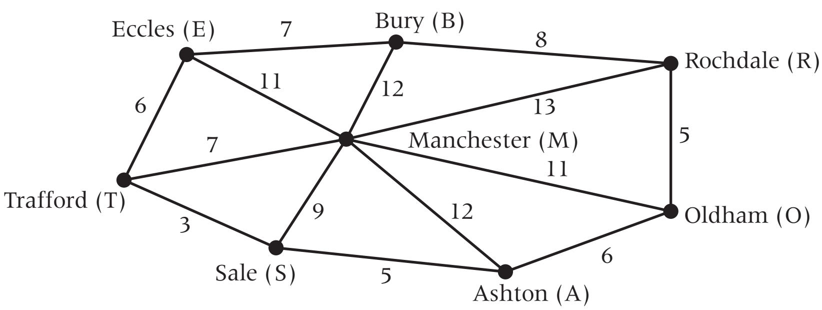
(b) Find the shortest time to travel from A to G if in reality each time the train passes through a station, excluding A and G, an extra 10 minutes is added to the journey time. [A]

3 The following network shows two islands, each with 7 small towns. One road bridge connects the two islands. Values shown represent distances by road, in miles.



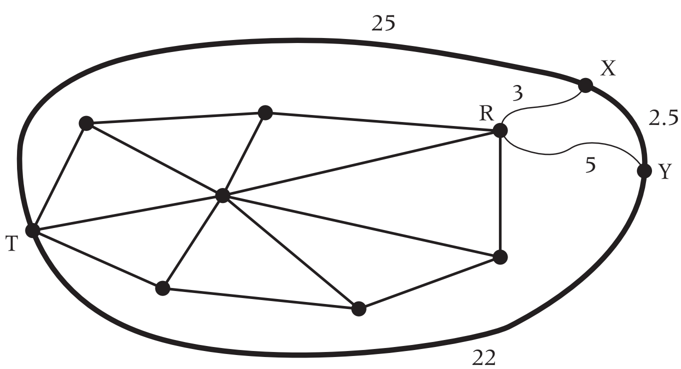
Use Dijkstra's algorithm to find the shortest distance between A and N, stating the route. [A]

4 The following diagram shows main roads connecting places near to Manchester, where the values shown represent the distances in miles. Mark lives in Rochdale and works in Trafford.



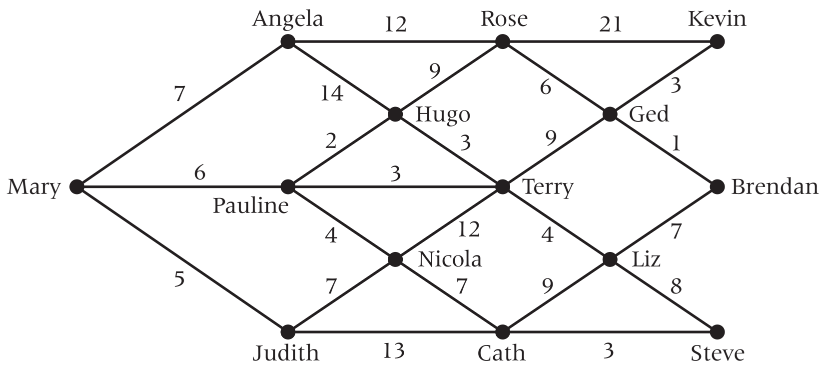
(a) Use Dijkstra's algorithm to find the shortest distance from Rochdale to Trafford. Write down the corresponding route.

(b) A new orbital motorway is built around Manchester as shown in the diagram below. The values shown represent the distances in miles.



Mark has access to the new motorway at points X and Y. Due to traffic conditions, he can drive at 20 mph on all main roads and 40 mph on the motorway.
Find the minimum time for Mark to travel from Rochdale to Trafford and state the route he should take. [A]

5 Every day, Mary thinks of a rumour to spread on her way to school. The rumour is then spread from one person to another. The following network shows the route through which the rumour spreads. The number on each arc represents the time, in minutes, for the rumour to spread from one person to another.



(a) Use Dijkstra's algorithm to find the time taken for the rumour to reach each person.

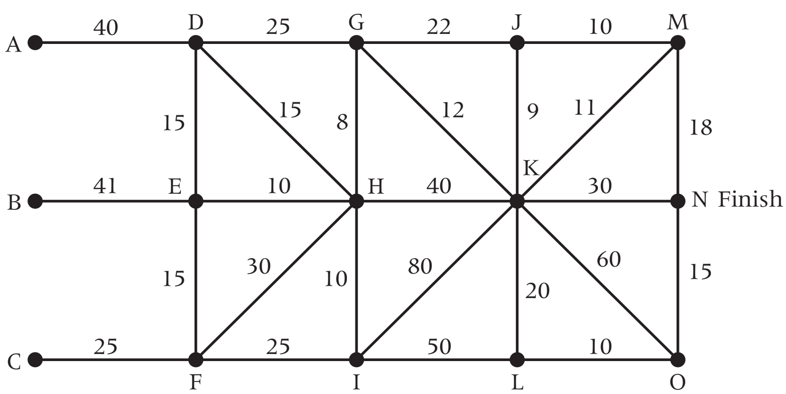
(b) List the route through which Brendan first hears the rumour.

(c) On a particular day Pauline is not at school. Find, by inspection, the extra time that elapses before Brendan first hears the rumour for that day. [A]

6 Three boys, John, Lee and Safraz, are to take part in a running race. They are each starting from a different point but they all must finish at the same point N.

John starts from the point A, Lee from the point B and Safraz from the point C.

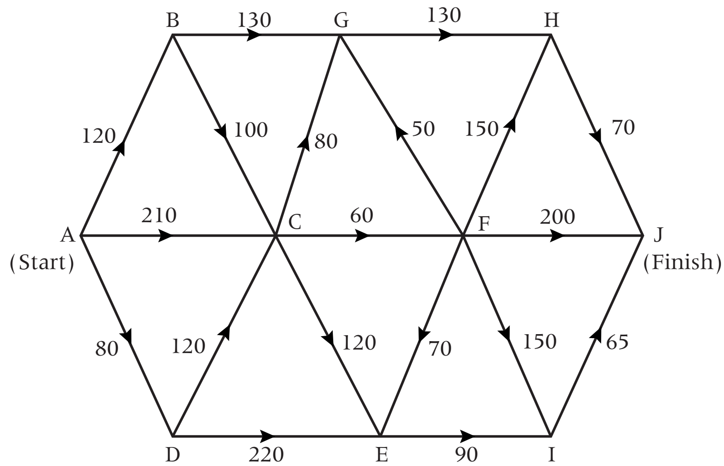
The following diagram shows the network of streets that they may run along. The numbers on the arcs represent the time, in seconds, taken to run along each street.



(a) Working backwards from N, or otherwise, use Dijkstra's algorithm to find the time taken for each of the three boys to complete the course. Show all your working at each vertex.

(b) Write down the route that each boy should take. [A]

7 A school is organising a short road race in which pupils have to start at A and use their own choice of route to reach J as quickly as possible. The diagram below shows the network of roads available and, for each road, the minimum completion time, in seconds, for a girl in the school.

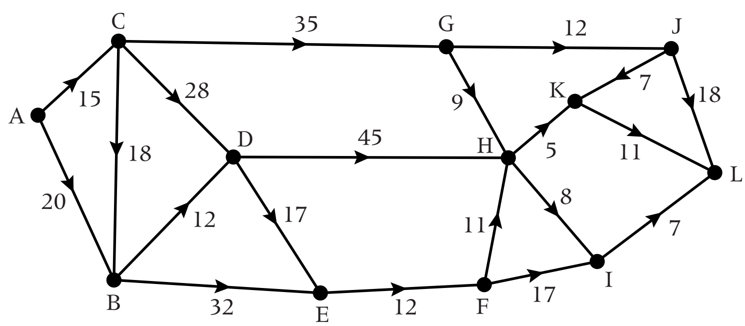


Mia, the best girl runner in the school, is planning her strategy. Given that she is able to run each section in the minimum time, use Dijkstra's algorithm to find the route she should take and her time for the race. [A]

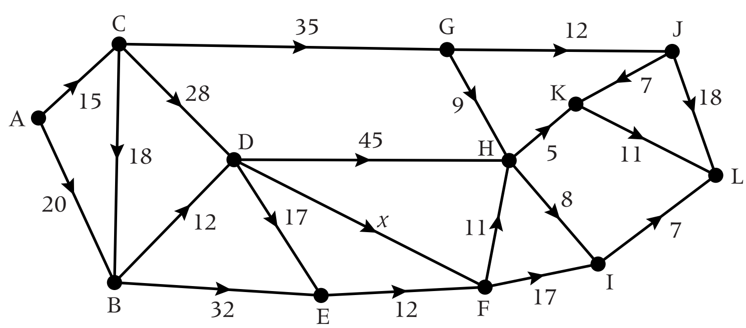
Hint for Question 7: See margin note on page 25.

8 An insurance salesman has to drive from his home at A to his head office at L. The time, in minutes, for each section of the journey is shown in the diagram below.

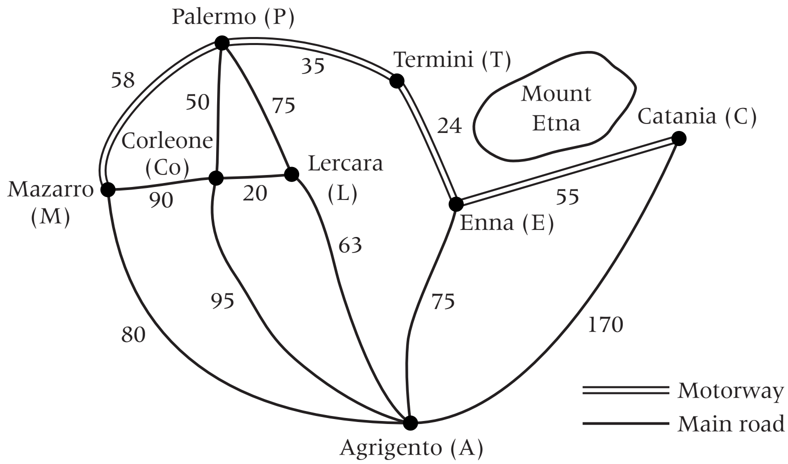
(a) Use Dijkstra's algorithm to find the minimum time for the total journey and state the route the salesman should take.



(b) A new road is constructed joining D to F, as shown below. The journey time for this section of road is x minutes. Find an expression, in terms of x, for the minimum time for the journey from A to L using the new road. [A]



9 The following diagram shows motorways and main roads connecting towns in Sicily. The numbers represent the times taken, in minutes, to drive along each road. There are two airports on the island, one at Catania (C) and one at Palermo (P).



Stella plans to fly to Catania and then drive to Agrigento.

(a) Find, by inspection, the minimum time for Stella to drive from Catania to Agrigento.

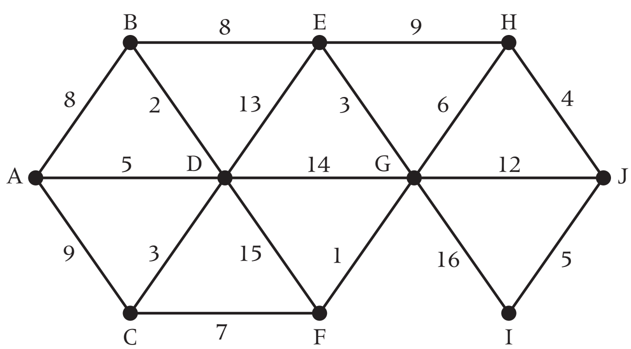
(b) Due to the volcano at Mount Etna erupting, Stella's flight is diverted to Palermo.

(i) Use Dijkstra's algorithm to find the minimum time to drive from Palermo to Agrigento.

(ii) State the route that she should take.

(c) Stella drives at 50 km/h on main roads and 100 km/h on motorways. Given that she keeps her driving time to a minimum, find the extra distance that she would have had to drive if she had landed at Catania airport rather than at Palermo airport. [A]

10 The following diagram shows the lengths, in miles, of roads connecting 10 towns.

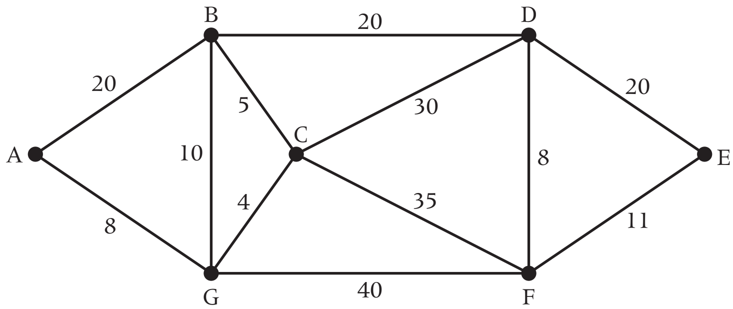


(a) Use Kruskal's algorithm, showing the order in which you select the edges, to find the minimum spanning tree for the network. Draw your minimum spanning tree and state its length.

(b) **(i)** Use Dijkstra's algorithm to find the shortest distance from A to J. State the route corresponding to this minimum distance.

(ii) A new road is built connecting F to I. The length of this road is x miles, where x is an integer. A shorter route from A to J than that found in **(b)(i)** is now available.
Find the maximum value of x. [A]

11 A railway company is considering opening some new lines between seven towns A–G. The possible lines and the cost of setting them up (in millions of pounds) are shown in the following network.



(a) Use Dijkstra's algorithm to find the minimum cost of opening lines from A to E. Show all your workings at each vertex.

(b) From your workings in **(a)** write down the minimum cost of opening lines:

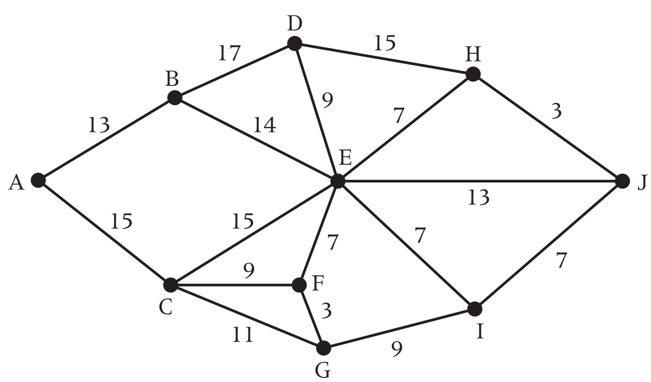
(i) from A to B

(ii) from A to F.

(c) Use Kruskal's algorithm to find the minimum cost of opening lines so that it is possible to travel between any two of the towns by rail, and state the lines which should be opened in order to achieve this minimum cost.

(d) The rail company wants to open some of the lines so that it is possible to travel by rail starting at one town, finishing at another and passing through each of the other five towns exactly once. Which lines should the rail company open in order to do this as cheaply as possible? [A]

12 The diagram represents the roads joining 10 villages, labelled A to J. The numbers give distances in kilometres.

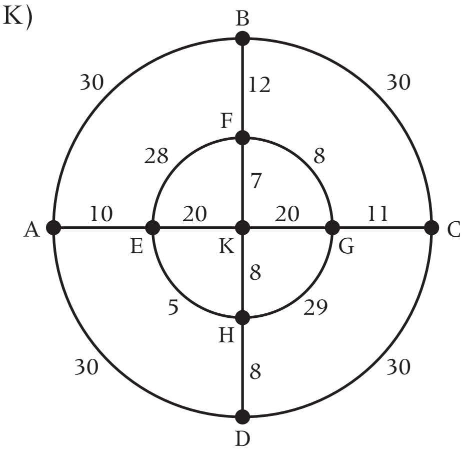


(a) Use Dijkstra's algorithm to find a shortest route from A to J. Explain the method carefully, and show all your working. Give a shortest route and its length.

(b) A driver usually completes this journey driving at an average speed of 60 km/h. The local radio reports a serious accident at village E, and warns drivers of a delay of 10 minutes.

Describe how to modify your approach to **(a)** to find the quickest route, explaining how to take account of this information. What is the quickest route, and how long will it take? [A]

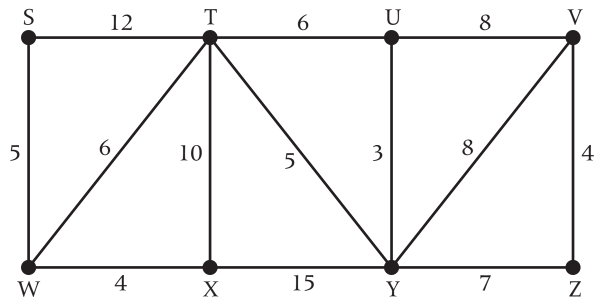
13 The network shows the roads around the town of Kester (K) and the times, in minutes, needed to travel by car along those roads.



(a) A motorist wishes to travel from A to C along these roads in the minimum possible time. Use Dijkstra's algorithm to find the route the motorist should use and the time that the journey will take. Show all your workings clearly.

(b) The four sections of ring road AB, BC, CD and DA each require the same amount of time, and next year there will be improvements to the ring road in order to reduce this time from 30 minutes to m minutes. This will enable the motorist to reduce the minimum time for a journey from A to C by 2 minutes. Find the value of m and state his new route. [A]

14 The network shows the distances in kilometres of various routes between points S, T, U, V, W, X, Y and Z.



Use Dijkstra's algorithm to find the shortest path from S to Z. Show your working. [A]

Key point summary

1 Real-life problems may not obey the triangle inequality. *p24*

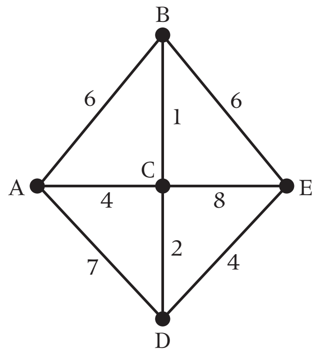
2 **Dijkstra's algorithm** enables the shortest path between two points to be found. *p25*

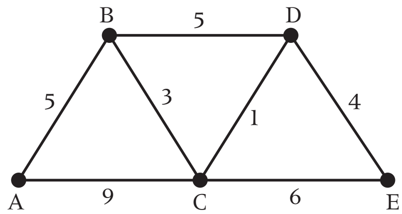
3 Dijkstra's algorithm is equally valid when used backwards through a network. *p30*

4 Dijkstra's algorithm fails if there are negative edges in a network. *p32*

5 Tracing a route through a network can be easily found if careful labelling of Dijkstra's algorithm is used. *p33*

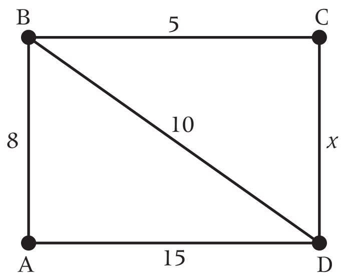
Test yourself	What to review
1 Use Dijkstra's algorithm to find the shortest distance from A to E in the networks below.	*Section 2.3*

(a) 

(b) 

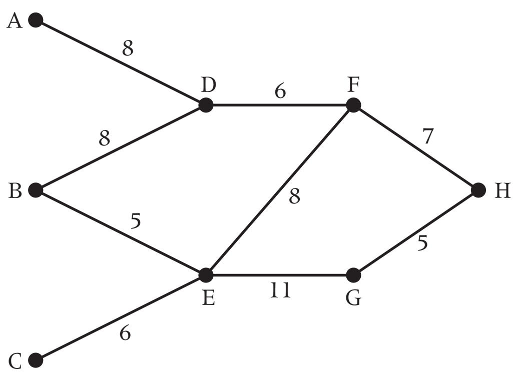
2 Find the route corresponding to the shortest distance for question 1. — *Section 2.6*

3 The network below has four vertices and five edges. — *Sections 2.3, 2.6*



Given that the second shortest route from A to D is ABCD, find the range of values of x.

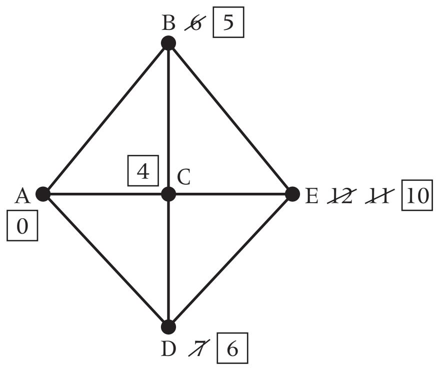
4 The network below has eight vertices and nine edges. — *Section 2.4*



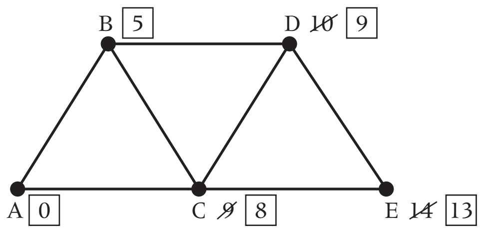
Find which of the vertices A, B or C is nearest to vertex H.

Test yourself ANSWERS

1 **(a)**



(b)



2 **(a)** ACDE

(b) ABCDE

3 $2 < x < 5$

4