

Markscheme

November 2015

Computer science

Higher level

Paper 2

This markscheme is the property of the International Baccalaureate and must **not** be reproduced or distributed to any other person without the authorization of the IB Assessment Centre.

Subject details: Computer science HL paper 2 markscheme

Mark allocation

Candidates are required to answer **all** questions in **one** Option. Total 65 marks.

General

A markscheme often has more specific points worthy of a mark than the total allows. This is intentional. Do not award more than the maximum marks allowed for that part of a question.

When deciding upon alternative answers by candidates to those given in the markscheme, consider the following points:

- Each statement worth one point has a separate line and the end is signified by means of a semi-colon (;).
- An alternative answer or wording is indicated in the markscheme by a "/"; either wording can be accepted.
- Words in (...) in the markscheme are not necessary to gain the mark.
- If the candidate's answer has the same meaning or can be clearly interpreted as being the same as that in the markscheme then award the mark.
- Mark positively. Give candidates credit for what they have achieved and for what they have got correct, rather than penalizing them for what they have not achieved or what they have got wrong.
- Remember that many candidates are writing in a second language; be forgiving of minor linguistic slips. In this subject effective communication is more important than grammatical accuracy.
- Occasionally, a part of a question may require a calculation whose answer is required for subsequent parts. If an error is made in the first part then it should be penalized. However, if the incorrect answer is used correctly in subsequent parts then follow through marks should be awarded. Indicate this with "FT".

General guidance

Issue	Guidance
Answering more than	 In the case of an "identify" question read all answers and mark positively up to the maximum marks. Disregard incorrect answers.
the quantity of responses prescribed in the questions	 In the case of a "describe" question, which asks for a certain number of facts eg "describe two kinds", mark the first two correct answers. This could include two descriptions, one description and one identification, or two identifications.
	 In the case of an "explain" question, which asks for a specified number of explanations eg "explain two reasons", mark the first two correct answers. This could include two full explanations, one explanation, one partial explanation etc.

[5]

Option A — Databases

1. (a) (i) Award up to [1 max]. Greater financial control: Sales analysis; Better customer services/information; [1] (ii) Award up to [1 max]. Cost of hardware / hardware maintenance; Cost of software / software maintenance: Training cost; [1] (b) (i) 5; [1] (ii) Award up to [2 max]. Primary key should be unique for each record; Many people can have the same name; So the customer name does not uniquely identify the record (for example Ann Low); [2] A candidate key that has not been chosen/field used for indexing; [1] (iii) SELECT Item Ordered FROM Customer Table WHERE Date ==19/04/15 AND Quantity_Ordered > 5 Award marks as follows, up to [4 max]. Award [1] for selecting correct field from the table. Award [1] for correct logical operation. Award [1] for correct date comparison. Award [1] for correct quantity comparison. Accept logically equivalent answers. [4] Reorganize database to be relational; So that records are accessed/retrieved faster; Because relational reduces repeated data; Defragment the disk: In order to reduce search time; Normalise the database; To reduce redundancy;

2. (a) Award up to [3 max].

Acts as an interface between the user's view and the way in which data is actually held; Allows users to store/restore/update information;

Protects database from accidental/deliberate corruption:

Protects from data theft:

Provides the ability to recover database in the event of system failure;

[3]

(b) Award up to [2 max].

Example answer:

The purpose of normalization is to separate different entities (columns);

To make the most efficient use of space;

And to allow for all possible relationships between data items;

[2]

[1]

[1]

(c) Example answers:

- (i) A table/relation is in 1NF if it does not contain repeating columns/attributes;
- (ii) A table/relation is in 2NF if it is in 1NF and no column/attribute that is not part of a primary key is dependent on only a portion of the primary key;
- (iii) A table/relation is in 3NF if every column/attribute is dependent on the key (and nothing but the key) / contains no transitive dependencies; [1]

(d) Award up to [7 max].

In the first normal form a primary key identifying the transaction is required

<u>Rental No</u>, HirerID, HName, HPhone, BikeID, TimeOut, TimeIn, BikeMake, BikeModel

Award [2] for identifying the new attribute as a primary key.

(Alternatively, award [2] for identifying either HirerID and TimeOut or BikeID and TimeOut as a composite key.)

Award [1] if BikeID and HirerID are used as composite key.

As there is only one primary key the 2nd Normal form is the same,

Rental No, HirerID, HName, HPhone, BikeID, TimeOut, TimeIn, BikeMake, BikeModel

Award [1] for the same table as in 1NF. Allow FT for the use of composite keys for 1NF.

In the 3rd Normal form

HIRERTABLE

(<u>HirerID</u>, HName, HPhone)

BIKETABLE

(BikeID, BikeMake, BikeModel)

RENTALTABLE

(Rental_No, TimeOut, TimeIn, BikeID*, HirerID*)

Award [4] for a completely correct answer (including the indication of the foreign keys). Otherwise award [1] for each correct table and [1] for the use of foreign keys.

[7]

3. (a) Award up to **[4 max]**.

Logical design will consist of definitions of tables and links between them; Physical design will consist of implementing these tables with data types needed;

Logical design refers to what database is;

As opposed to how it operates;

Logical design is a detailed description of the database model from the users' perspective;

Whilst physical design specifies exactly how the data will be arranged and stored;

Logical design involves user information needs, analyzing data requirements and logical grouping;

Whilst physical design involves storing this data so it can be retrieved and updated as efficiently as possible;

Logical design should be completed first;

And then the next step is physical design;

[4]

(b) Award up to **[3 max]**.

Who owns or is responsible for the data;

What data is available:

Where the data is located;

Data descriptions;

How the data is used;

Access rights / who is allowed to access data;

Etc.

[3]

(c) Award [1] for each responsibility and [1] for the explanation, up to [4 max].

Database design;

The DBA participates in the logical and physical phase;

Database implementation/operation;

This involves guiding the use of the database – this includes adding data, deleting data, controlling access to data;

Coordination with users;

Receiving and reviewing user requests for additional support;

Backup and recovery:

DBA is responsible for periodic backups and establishing processes for recovery from failures/archiving;

System security;

This involves assignment of user passwords and controlling users' access to the database:

(d) Award marks as follows up to [4 max].

Award [1] for identifying an advantage.

Award [1] for identifying a disadvantage.

Award [2 max] for a discussion that may focus on usability, performance or privacy. For example:

Advantage

All jobs are listed centrally and updated;

All jobs are accessible from one place for the applicants;

Accessible at any time by applicants;

Etc.

Disadvantage

Users' details could get hacked as stored online;

Employers may see details of current employees and know they are thinking of leaving; *Etc.*

4. (a) Award up to [2 max].

It is a repository of data collected over time;

That is created by integrating data from one or more disparate databases;

To be used for reporting and data analysis;

To find patterns;

[2]

(b) Award up to [4 max].

Data from different databases (schools', towns', districts' databases) is used to create the data warehouse/will be in different formats:

Data should be extracted:

Transformed into uniform format:

Extracted and transformed data should be loaded into the data warehouse; And finally from the data warehouse into a set of data marts that are accessible by

decision makers/school administrators, teachers, government;

[4]

(c) (i) Award up to [2 max].

Able to access student performance on external assessment;

Compare performance with predicted grades;

Can identify problems and opportunities for improvements to ensure students have the best possible education;

[2]

(ii) Award up to [2 max].

Can measure the progress of school against various criteria;

Can assess school performance over time;

Compare schools;

[2]

(d) Award up to **[4 max]**.

Answers that refer to data integrity rather than security should not get credit.

Access controls:

Effective and efficient access controls/restrictions so that the end users can access only the data or programs for which they have legitimate privileges;

Employers could defend against security vulnerabilities of the data warehouse; Addressing employee hiring, training (security awareness), periodic background checks, and transfers;

Data Encryption;

Nullifies the potential value of data interception;

Ensures the confidentiality;

(e) (i) Classification recognizes patterns that describe the groups to which an item belongs;

It examines existing data and helps to discover the characteristics of students who failed to complete the examination or the students who achieved excellent results:

Forecasting uses patterns within the data to estimate the future values;

[3]

(ii) Award marks as follows, up to [3 max].

Award [1] for stating what predictive modelling is.

Award [2 max] for a discussion.

Example:

Predictive modelling is a process by which a model is created or chosen to try to predict the probability of an outcome;

It can help to determine the most effective ways to keep/achieve the educational standards as high as possible;

And/or spend as much money as necessary to retain that standard;

The problem is that using historical data to predict the future assumes certain conditions are constant which is almost always wrong when the system involves people;

[3]

Option B — Modelling and simulation

5. (a) Award [1] for clear access to 24 readings.

Award [1] for "per day".

Award [1] for "per year".

Do not give the first mark if all months are on the same worksheet.

For example:

One worksheet/table for each year;

Each worksheet with a day per column/row;

Hourly readings in row/column;

[3]

(b) Award marks as follows, up to [5 max].

Award [1] for keeping an overall count of days >50.

Award [1] for calculating averages each day.

Award [1] for correct data sent to press.

Award [1] for calculating yearly average.

Award [1] for correct data (including suitable summary of days) sent to EU.

Accept an Excel solution with correct formulae which is illustrated and/or described.

For example:

A count set to zero (could be one count for each sheet – then totalled) at the beginning of the year

At the end of each day

average for day calculated

if >50 add one to count

send to press

At end of year

calculate overall average

present averages for each day (table or bar chart)

send to EU together with the overall count

Accept logically equivalent answers.

[5]

(c) (i) Award up to **[2 max]**.

Averages for winter and summer etc;

Year to year could be looked at;

Different time of the day;

Etc.

[2]

(ii) Introduce times/dates where central heating should (not) be turned on;

Adjust working/school hours, to change frequency of cars on the road;

Accept reasonable suggestions that link to a pattern given in (c)(i).

[2]

(d) Award up to [5 max].

Knowing the amount of reduction;

Creating a simulation of future weeks and months;

Based on the pattern so far;

Apply the reduction progressively:

Until satisfactory level achieved;

This would need repeating regularly as figures may not continue to follow the same pattern:

Award marks for a reasonable strategy that uses current data to make an estimate.

[5]

6. (a) 2, 5 or 11; [1] (b) A goes to 3, then B to 11, 14 or 15. So move to 11 or 15 to win; A goes to 6, then B can go to 15 to win; A goes to 7, then B can go to 11 or 15 to win; [3] (c) (i) Award up to [2 max]. Table / 2-D array; With i index being row and j index being column; Represent A/B with two numbers; One being its row position, one being its column position; Distinguish between A and B using another variable; Accept alternative solutions. [2] (ii) Award marks for answers consistent with the answer in (c)(i), showing A and B to be in correct position. For example: 2-D Array, A would be in 1, 2 and B in 3, 2 (accept 2, 1 and 2, 3); Three variable methods: 0 represents A and 1 represents B, so A is 012 and B is 132; Etc. [1] This will depend on the structure chosen. Award marks as follows: (iii) Adjacent squares correct in terms of chosen structure with "off board" checked; If opponent in possible square then check "jump over" square is on board; [2] (iv) For each possible move from above calculate the possible moves for the opponent: For each possible opponent move, identify all possible moves for next go; Until player or opponent gets to HOME; Choose move on shortest winning path; [3] (d) Award [2] for explaining time taken and [2] for the effect on the game, up to [4 max]. The number of possible paths would increase rapidly: For each extra row and column; This would mean more processing; And slow down the game; [4] **7.** (a) Award marks for an answer which outlines the relationship between the maps and current inputs. For example:

The camera and lasers can interpret the immediate surroundings;

Car needs a destination (path to follow);

High definition maps will combine map and input (relate present position to destination); [2]

(b) Award marks that give reference to:

Detail of environment;

Movement relative to time;

Avoidance of objects;

For example:

The 3D image shows the environment at any instant in time;

With the possibility to distinguish movement;

In relation to previous instant in time;

Hence objects can be avoided using the projected path;

[3]

(c) Visualization must be in real time;

As car is moving in real time;

Requires massive memory to manipulate multiple images in colour/greater depth;

[3]

(d) (In theory) they should be safer;

More economical/environmentally friendly than humans – human error eliminated – rules set for safety and economy;

Essential to include fail-safe measures in case of malfunction;

Computers cannot react to unknown situations;

Who is responsible/liable in case of an accident?;

Etc.

8. (a) Generally semantics is the problem. Award [2] for any aspect with elaboration:

Although "understanding" is possible the same word can have different meanings (eg box, green);

Natural language has syntax but the rules are not always clear and a word can sometimes be a verb or a noun: (eg runs, gains);

The same phrase in the same language in different parts of the world can have different meanings;

(b) Award up to [4 max].

A random set of solutions would be generated on the sample documents;

And tested against the human labelling;

Best fit solutions retained;

New generation created by mutating/crossing;

Algorithm repeated;

Until a good fit obtained;

[4]

[4]

(c) Supervised and unsupervised learning are two approaches. Award [2] for each explained/compared:

Supervised learning will involve the machine using the labelled samples; To distinguish between success or not and continue until correct;

Unsupervised learning involves the machine distinguishing clusters of information; Which have similarities with no fitness function to test against;

Credit answers which discuss the difference between cognitive and heuristic learning. [4]

(d) Award up to [3 max].

Speech recognition would be an advantage to the user who could search easily from mobile device;

in any situation;

an extra step in the machine learning;

going from spoken word or phrase to the written word;

[3]

(e) Award [2] for advantages, [2] for disadvantages and [1] for a conclusion. For example:

Advantages:

Important developments and views get published quickly; Such as in dangerous situations (war, earthquake *etc*); *Etc.*

Disadvantages:

A lot of the information is very subjective or politically undesirable; Sources easily traced can lead people into dangerous situations...;

Overall international boundaries could be opened leading to a safer world; But it depends on who uses the information...;

[5]

[4]

Option C — Web science

9. The first DVD tag is not closed; [1] (b) Allows users to add to HTMLs tag set; By using tags (tag names) of their own choice; [2] (c) Award up to [2 max]. Saves time (for the web designer) as there is no need to add/change formatting to each HTML document: Because the same CSS file can be applied to many HTML documents; Because any change in the CSS file can be applied across many documents; [2] (d) The TCP protocol causes data from the application (browser); To be divided up into packets; The IP protocol takes these packets and adds a header to them; This includes routing information; Etc. [4] (e) (i) JavaScript; Could be used to validate the input data; eg checking for the @ sign in the email address; [3] (ii) Less processing is required at the server-side; So it can handle more requests at the same time / use a less powerful server; [2] (f) Award up to [4 max]. A scripting language is present on the (user's) web page; This sends the request from the user to the web server; Where using SQL (to link to the database); The user's photograph is retrieved; And returned to the specific section of the web page / And only the photo is reloaded;

10. (a) Award up to **[5 max]**.

Lossy compression results in data being lost from a file;

Positive effect on the user's experience as the time of transmission is reduced/delivery is possible;

Because of smaller file size;

And the data removed may be irrelevant/not discernible;

eg Excess detail in a photo (which would not be discernible to the human eye);

Negative effect on the user's experience if the data removed is significant;

This outweighs the reduced time of delivery;

[5]

(b) Crawling: Exploration of every link/page in the web;

Returning a copy of that page;

Indexing: Each web page is analysed for key words;

After which the page is added to an index;

Searching: Queries are entered;

And the index is searched for matches:

[6]

(c) The ranking produced by the PageRank algorithm is based on examining (the number and importance of) back-links;

A new site will initially have few/none of these;

So will receive a low ranking;

[3]

(d) Award up to [3 max].

A (search engine's) web-crawler works continuously;

So new pages will eventually be found:

As long as they have links from other pages;

Web-site administrators can notify search engines of any new pages;

[3]

11. (a) Award up to **[2 max]**, **[1]** for the concept and **[1]** for a good description, for any two of the following:

Virtualization that allows the sharing/pooling of resources;

Use of distributed/grid systems that allows fast communication between the different components:

Self-service interface which allows customers to select required services;

Scalability which allows rapid increase in services on-demand;

[4]

(b) It might use its own private cloud for holding all sensitive data;

And for normal operations:

But use the public cloud when increased computing facilities / network capacity is required;

As over the Christmas period;

Or in periods of development when temporary extra storage is required;

Or for its mailing operations;

As public clouds have dedicated services for this;

[6]

12. (a) Web pages are designed by different people who may use different keywords; In tags describing their pages;

This makes a systematic search of the web difficult;

The use of ontologies would standardize the use of keywords in related areas;

Allowing for successful computer-based searches of the web;

[5]

(b) Using the combined intelligence of a group of people; In a beneficial way for the company;

[2]

(c) Award up to [4 max].

For example:

The use of ratings by users;

Will draw attention to the product;

Decide whether they should stop selling products;

Recording how many times items are bought/clicked on;

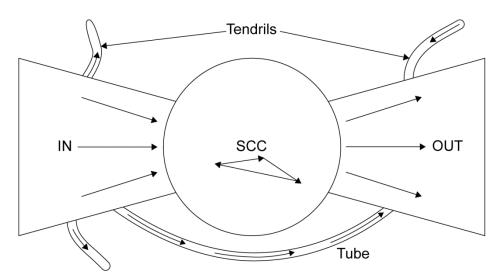
If something is not being bought/browsed, retailer can promote it;

Items can have prices adjusted (be put on sale);

Tells users what similar customers have bought ("customers who bought *x* also bought *y*");

[4]

13. (a)



Award [1] for a bowtie shaped diagram.

Award [1] for correct labelling the IN/OUT/SCC sections in diagram.

Award [1] for describing each of the IN, OUT and SCC up to [3 max].

- IN contains pages that have no in-links;
- OUT contains pages that have no out-links;
- SCC contains pages where any one page can link to any other page in that section;

[5]

(b) Award [1] for an example.

eg in-links / no. of links per page etc;

Power laws are of the form $f(x) = a.x^{(-c)}$ / means the frequency of an event grows inversely proportional to the power of the size of the event / negative exponential / the fraction of sites which are visited a lot of times;

For this example the frequency of the pages decreases as the number of in-links that it has increases;

Power law graph with correct labels (either log or not);

Award marks for any of these points that are communicated diagrammatically.

Option D — Object-oriented programming

14. (a) Award [1] for a definition, such as:

A (special) method (with the same name as the class) that instantiates an object of that class;

```
Award [1] for the example:
```

```
public Patient(String i, String n, int p)
```

[2]

(b) Award [1] for a field and its data type.

For example:

```
String phoneNumber;
```

[2]

(c) Award [1] for any indication of aggregation.

Award an additional [1] for a full description.

Example answer:

The WaitingRoom object stores up to 10 instances of the Patient object;

The WaitingRoom has a Patient;

Accept a correctly labelled diagram.

[2]

(d) Award marks as follows up to [3 max].

Award [2] for correct loop, award [1] for loop with one condition.

Award [1] for correct test.

Award [1] for assigning max.

```
Example answer:
```

```
int i=1;
while ((i<10) && (patients[i]!=null))
{
   if (patients[i].priority > patients[max].priority)
   {
      max = i;
   }
   i = i+1;
}
```

[3]

(e) Award marks as follows up to [6 max].

Award [1] for including the correct signature.

Award [1] for correct output of patient details.

Award [1] for any loop.

Award [1] for correct loop.

Award [1] for correct assignment to shift a patient instance.

Award [1] for assigning null.

```
Example answer:
```

```
public void remove(int n)
{
   System.out.println(patients[n].toString());
   for (int i=n; i<9; i=i+1)
   {
      patients[i] = patients[i+1];
   }
   patients[9] = null;
}</pre>
```

[6]

15. (a) Award [1] for a suitable definition, for example:

Encapsulation means having private variables;

Variables not directly accessible from outside the class;

Methods and variables are all included in the class definition:

Award [1] for relating an advantage to the Patient class, such as:

So that patient data cannot be accessed/modified accidentally;

So that patient data is more secure;

[2]

(b) Award [1] for identifying an ethical issue and [1] for an elaboration.

Award [1] for relating this issue to the Treatment class and [1] for an elaboration.

Example answer:

Sensitive information should be separated from identity data;

In order to avoid information misuse;

The consultation object does not include the patient's name (only an ID);

So that a sensitive diagnosis (like AIDS) is not directly linked to a name;

[4]

(c) Award marks as follows up to [3 max].

Award [1] for including String name with correct getter and setter methods.

Award [1] for including String phone with correct getter and setter methods.

Award [1] for including Boolean present with correct getter and setter methods;

Example answer:

```
Doctor

String name
String phone
Boolean present

setName(String name)
setPhone(String phone)
setPresent(Boolean present)
String getName()
String getPhone()
Boolean getPresent()
String toString() //optional
```

[3]

(d) Award [1] for identifying an appropriate need for extended character sets and an additional [1] for an elaboration of this need. Award [1] for relating to the Doctor object.

Example answer:

Different languages use different characters, which are not included in the basic 8-bit ASCII character set;

Extended character sets (like Unicode) include all possible characters from all languages;

For example, the doctor's name could not be spelled correctly in ASCII;

[3]

```
16. Award marks as follows up to [6 max].
```

Award [1] for correctly initialized method.

Award [1] for creating a new Treatment object.

Award [1] for loop through treatment file.

Award [1] for setting current equal to next in file.

Award [2] for correct if statement that checks names.

Award [1] for correct output statement.

```
void showMedicationByDoctor(String name)
{
    Treatment current;
    while (treatmentFile.hasNext())
    {
        current = treatmentFile.getNext();
        if (current.getDoctor().equals(name))
        {
            output(current.getMedication());
        }
    }
}
```

[6]

17. (a) Award marks as follows up to **[6 max]**.

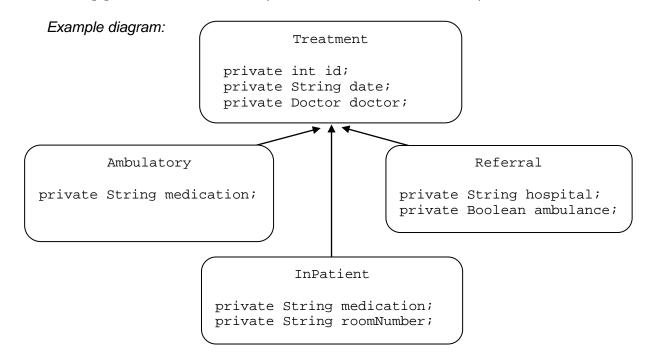
Award [1] for a common super-class.

Award [1] for all common fields: id field (integer) date field (String) doctor field (String).

Award [1] for three sub-classes and award [1] for arrows with correct direction. Award [1] for the medication fields in Ambulatory and InPatient, but not in Referral.

Award [1] for the additional fields (Hospital and Ambulance) in Referral.

Award [1] for the additional fields (RoomNumber and ReleaseDate) in InPatient.



Do not penalize if data types are not specified.

(b) Award [1] for identifying each advantage and an additional [1] for an elaboration of this advantage, up to [6 max].

Example answer:

Faster development;

Because different programming teams can work on different modules;

Easier to debug;

Because the smaller modules will have fewer mistakes than one big program;

Easier to update (in the future);

Because it is easier to update a module than the full program;

Re-usability;

Modules can be stored in libraries and reused for different programs;

[6]

[6]

18. (a) (i) A stack is a LIFO structure; [1]

(ii) Award [1] for stating an application and [1] for an elaboration. Example answer:

The system stack stores return data of interrupted processes; The last interrupted process is the first to resume;

Pre-condition: there is at least one patient instance in the list;

[2]

(b)

[1]

(c) Award [1] for each row filled correctly.

i	current.name	firstup.name	result
0	-	null	0
0	Abdul Hashim	Abdul Hashim	0
1	Iris Gotenberg	Abdul Hashim	0
2	Anh Nguyen	Anh Nguyen	2

[3]

[1]

[3]

- findNextPatientIndex returns the index of the first patient with the highest priority in the list/find the next patient with the highest priority;
- Award marks as follows up to [3 max].

Award [1] for introducing and initializing a Boolean done.

Award [1] for adding ! done to the loop condition.

Award [1] for testing for early exit (current.priority==3).

Example answers:

Add a boolean done and set it to false before the loop;

Add &&(!done) to the loop condition;

if current.priority equals 3 then make done true;

```
boolean done = false;
```

```
while ((i < PatientList.size()) && (!done))</pre>
  current = PatientList.get(i);
  if (current.getPriority() > firstup.getPriority())
    firstup = current;
    result = i;
  if (current.getPriority()==3) { done = true; }
  i=i+1;
```

while ((i < PatientList.size()) && (current.getPriority()!=3))</pre>

19. (a) A programming technique where a method calls on itself;

[1]

(b) Award up to [4 max].

Award [1] for correct result, 840.

Award [1] for applying base case to result(3,0) = 1.

Award [1] for showing at least 2 recursive calls of result(), even if the final calculation is incorrect.

Award [1] for presentation, showing the alternation of + and *.

```
result(3,4)
(3+4)* result (3,3)
(3+4)* (3+3)* result (3,2)
(3+4)* (3+3)*(3+2)* result (3,1)
(3+4)* (3+3)*(3+2)*(3+1)* result (3,0)
(3+4)* (3+3)*(3+2)*(3+1)*1 = 840
```

[4]

(c) Award up to [4 max]

Award [1] for having a return call for each branch

Award [1] for both cases returning a call to result().

Award [1] for checking the condition on x, y being non-negative.

Award [1] for correct recursive call on result4().

```
public double result4(double x, int y, double z, int v)
{
   double sum = x + y + z;
   if (sum > 12*v)
   {
      return result(v,y);
   }
   else
   {
      if ((y<x/2)&& (v-3>=0 && y-3 >=0))
      {
           return result4(x,y-3,z,v-3);
      }
      else
      {
           return result(0,0);
      }
   }
}
```