N17/4/COMSC/SP2/ENG/TZ0/XX/M

Markscheme

November 2017

Computer science

Standard level

Paper 2

19 pages

Option A — Databases

1. (a) *Award marks as follow up to* ***[2 max]****.*
Award ***[1]*** *for: a unit of work/logical action;*
performed on a database;
independent of other transactions;
changes state of the database;

Example answer 1
A database transaction is a logical unit that is independently executed;
For data retrieval or updates;

Example answer 2
A database transaction is a unit of work;
That is either executed in full or not executed at all;

Example answer 3
A database transaction is a way of representing a state change;
And has four properties, known as ACID;

Example answer 4
A database transaction usually means a sequence of steps, treated as a unit;
For the purposes of satisfying a client's request;

Example answer 5
A database transaction is a process carried out on a database;
Which may change its state, for example: moving money between bank accounts; **[2]**

(b) *Award up to* ***[2 max]****.*

Example answer 1
Durability is important because transaction data changes must be available;
Even in the event of database failure;

Example answer 2
Durability means that if the system says the transaction has been committed;
The client does not need to worry about it because transactions that have been committed will survive permanently;

Example answer 3
Durability in databases is an important property because it ensures transactions are saved permanently;
And do not accidentally disappear or get erased; **[2]**

(c) *Award up to* ***[2 max]****.*
One to one;
One to many;
Many to one;
Many to many; **[2]**

(d) *Award* ***[1]*** *for stating what a data dictionary is and up to* ***[3 max]*** *for describing its contents.*

It is a set of tables/a database that provides information/meta-data about the database;

A data dictionary contains
The definitions of records/entities in the database;
How much space has been allocated for and is currently used;
Default values for columns/attributes;
Integrity constraint information;
Names of (database) users, their privileges and roles;
Auditing information such as who has accessed or updated data records;
etc (any other general database information);

They do not contain any actual data from the database, it contains only information for managing it (without a data dictionary a DBMS cannot access data from the database);
Data dictionaries are usually hidden from users to prevent them from (accidentally) deleting/changing/destroying its contents; **[4]**

(e) *Award up to* ***[4 max]***.
To direct and perform all activities related to a database system;
Installing and configuring software;
Creating new databases;
Designing the database schema and creating any necessary database objects;
Ensure database security is implemented to safeguard the data;
Back up and recover the database;
Work closely with application developers and system administrators to ensure all database needs are being met;
Apply patches or upgrades to the database as needed;
Training employees who use the database; **[4]**

2. (a) *Award **[1]** for stating what data is.*
*Award **[1]** for stating what information is.*
*Award **[1]** for stating the difference using an example.*

Example answer 1
Data is a raw fact, for example, “Joe Skrin” is data/a string;
Whilst information is data which has meaning;
In this table “Joe Skrin” is the name of an art student/who had an examination/ who passed an examination;

Example answer 2
For example, number 99 is data;
In this database, total 99 is information;
Because it has meaning, it represents a student’s mark/it shows that the student failed; **[3]**

(b) (i) An entity is some unit of data that can be classified and has stated relationships to other data units;
A real-world object with attributes that is represented as data in a database; **[1]**

(ii) In relation to given example, the entity is a single student about whom data can be stored as a record in this relation; **[1]**

(c) String/varchar/char/alphanumeric; **[1]**

(d) *Award up to **[4 max]**. Award **[1]** for stating the purpose of data verification and **[1]** for further explanation, up to **[2 max]**. Award **[1]** for stating the purpose of data verification and **[1]** for further explanation, up to **[2 max]**.*

Data verification is a process that ensures the accuracy of data;
(Accurate data is important because strategies devised based on incorrect data lead to inconsistent decision making;)
Data verification is conducted using proofreading/double entry checks/new data cleansing software and technologies have been developed to automate the data verification process;

Data validation ensures the data is logical and reasonable;
Data validation is a computer-generated process using codes to validate a range of data;

Data verification and data validation (applied in combination) provide quality assurance/make sure that processes and strategies are not driven in the wrong direction; **[4]**

(e) *Award up to **[2 max]**.*

***Example answer 1** (assuming the value of the Total is calculated)*
The calculated value for the total mark must be an integer;
In the range from 0 to 300;
(If it is not, an error message could be output;)

***Example answer 2** (assuming the value of the Total is entered)*
The entered value for the total mark must be an integer;
In the range from 0 to 300;
The sum of marks on the three exams can be calculated and should be equal to the entered value for Total;
(If it is not, an error message could be output;) **[2]**

Example answer 3
Total is a validated field;
Therefore, if all the individual test fields are validated, the Total value will automatically be within the specified range;

(f) (i) *Award **[1]** for selecting the correct field from the table and **[1]** for a correct comparison, up to **[2 max]**.*

```
SELECT Student_Name FROM Class_Table
WHERE Exam_Two == 100
```

[2]

***Note:** Accept logically equivalent answers.*

(ii) *Award marks as follows up to **[4 max]**.*
*Award **[1]** for selecting the correct field from the table.*
*Award **[1]** for each of the correct comparisons, up to **[4 max]**.*
*Award **[1]** for using logical operators correctly.*

Example answer 1

```
SELECT Student_ID FROM Class_Table
WHERE Exam_One >= 30 and Exam_Two >= 30 and
  Exam_Three >= 30 and Total >= 150
```

Example answer 2

```
SELECT Student_ID FROM Class_Table
WHERE not (Exam_One < 30 or Exam_Two < 30 or
  Exam_Three < 30 or Total < 150)
```

[4]

***Note:** Accept logically equivalent answers.*

3. (a) (i) *Award **[1 max]**.*
Redundant data means data that is held in two different places within a database; **[1]**

(ii) *Award **[1]** for identifying an issue caused by data redundancy and **[1]** for a brief explanation up to **[2 max]**.*

It could give the system unwanted/unexpected results;
due to the use of inaccurate data;

It may lead to additional storage requirements;
As data is used more times than necessary; **[2]**

(b) *Award up to **[3 max]**.*
Each attribute has a single value/is atomic;
All values for a given attribute are of the same data type;
Each attribute is unique;
This is a unique key;
There are no repeating fields;
There are no two identical tuples in this relation;
Order of attributes/tuples is not significant for the relation;
Key (Date + Code) is unique for each tuple; **[3]**

(c) *Award **[1]** for identifying why a compound key is used for the `SCHOOL_VOLUNTEERS` relation and **[1]** for a brief explanation up to **[2 max]**.*

The alternative is to use an autonumber field;
But this would use additional storage space;

Are used because it is not possible to designate a primary key from a single field;
Neither the code nor the date field on their own uniquely identify a record;
Is based on two primary keys in other tables; **[2]**

(d) *Award up to **[5 max]**.*

Example answer 1
2NF = 3NF if there are no transitive relationships/if any non-key attributes are more dependent on another non-key attribute than the key field;
`Schools_Table` could be the above if a school had more than 1 address;
Then the `Address` would depend upon the `School_Name`;
And the `Schools_Table` code would be split as follows:
(`Code`, `School_Name`)
(`School_Name`, `Address`);
If the school had only 1 address then 2NF = 3NF;
The `Volunteers_Table` has no transitive dependencies;
There is no redundant data;

Example answer 2
A relation is in 3NF if it is in 2NF and it contains no transitive dependencies;

Assuming that schools name is not unchangeable;
OR there are two schools with different names and same addresses;
OR two schools with same name and different addresses;
Then the school name cannot be treated as a key;

From the 2NF (two created relations above) the functional dependencies are not evident in the relation **School_Table** (School_Name, Code, Address):

The relation given above (in 2NF) is also in 3NF;
School_Table (School_Name, Code, Address)
Volunteers_Table (Code, Date, Num_Volunteers) **[5]**

Option B — Modelling and simulation

4. (a) *Award up to **[4 max]**.*
A computer model is a representation of a system;
Made up of variables and formulae/mathematical representation;
Whereas a computer simulation is a process that uses the model;
In order to see the outcome(s) when different values are used for the variables (in the model); **[4]**

(b) Either all of the variables are not known/difficult to define;
Or the relationships between them cannot be expressed accurately/mathematically; **[2]**

(c) (i) ***Example***
Experiments on living animals in a high school science class;
Such as effects of altering diet;
It is not ethically acceptable to harm animals (for such purposes);
***Note:** Accept reasonable examples, provided they are sufficiently explained.* **[3]**

(ii) *Award up to **[3 max]**.*
Computer models/simulations allow the designers to:
Make alterations and quickly see the outcomes;
Repeat tests several times over;
Model dangerous situations safely;
Learn from “what if?” scenarios;
Saves costs if several different models have to be built; **[3]**

5. (a) *Award up to **[6 max]**.*
Different sheets will be dedicated to different categories, *eg* salaries, running expenses, sales, different products, different months *etc.*;
***Note:** Award **[2 max]** for identifying at least two categories.*

Each sheet will contain a list of items for that category including associated values;
Formulas will be included as necessary;
Each sheet/category will include a formula that totals the values in that category;
Intermediate values will be calculated, *eg* tax to be paid;
The final profit will be determined from the previous totals; **[6]**

(b) Figures from previous years can be entered into the model;
With the results checked against the previously calculated results; **[2]**

(c) *Award up to **[4 max]**.*
Different (“what-if”) scenarios can be run;
In which the values of different variables are changed;
For example, the number of items sold / the increase in the level of salaries;
***Note:** Award **[2 max]** for identifying at least two items.*

Selling prices/other (acceptable) parameters can be adjusted to achieve the desired profit; **[4]**

6. (a) *Award up to **[4 max]**.*
Modern computer systems have become increasingly more powerful;
Which allows more complicated systems to be simulated;
In a short/acceptable period;
The understanding of (the science of) weather has steadily improved/more historic weather patterns can be accessed;
Allowing more accurate modelling to take place;
More data can now be retrieved;
Through satellites, ground stations *etc*; **[4]**

(b) Only the major/known variables can be input into the model;
Approximations have to be made to represent complex processes;
These or other minor/unknown variables will not have a significant effect in the short term;
But will have (unknown) effects in the long-term/small errors have a cumulative effect over the long-term/butterfly effect; **[4]**

(c) *Award **[1]** for each valid point discussed, and a further **[1]** up to **[5 max]** for a good expansion of this point.*

Investigate past weather patterns in order to see if past forecasts were correct or not and make adjustments to your model appropriately;
Look at specific events in the past (*eg* appearance of El Nino) to see how they affected future weather patterns in different areas, and then apply this to new occurrences of these events;
However, historical data does not take into account new factors, such as how carbon emissions are contributing to global warming; **[5]**

(d) (i) The representation of (abstract) data;
In a way that is understandable by humans; **[2]**

(ii) Standard simulations produce numerical values (which are not easily understood by the general public);
Modern simulations include
Graphics that are related to the data output;
eg rain drops/a smiling sun;
CGI/animation can produce motion;
eg showing a belt of rain crossing the country;
Which are more easily understood by the general public; **[6]**

Option C — Web science

7. (a) *Award up to **[2 max]**.*
Mark-up language;
Can contain scripting commands/code;
Can imbed objects;
Uses tags to structure page;
Provides constructs to build hyperlinks;
Can be integrated with CSS and JavaScript;
***Note:** Accept other reasonable characteristics.* **[2]**

(b) *Award marks as follows up to **[6 max]**.*
*Award **[1]** for identifying an advantage, **[1]** for identifying a disadvantage, and a further **[4 max]** for expanding on these advantages/disadvantages.*

Advantages
It saves time (in web development);
Because modification of style needs to only be made to the CSS;
It saves space;
Because the code (of the style) is not repeated in each HTML file;
Quicker load of many pages / Quicker download of many pages from the same site;
HTML files are smaller;
It gives a uniform appearance to the entire website;
Which makes it more attractive/appealing/easy to navigate;
It can be integrated with outside sources (*eg* RSS feeds);
Which makes the website more interactive/dynamic for the user;
Quicker to make changes to layout/formatting/page positioning;
Therefore, it helps separating the jobs/tasks of illustrators (designer) from those that produce content;
Increases accessibility of authoring to non-experts;
Supported by most browsers;
Increasing the number of users;

Disadvantages:
Downloading pages from different sites can be slow (general browsing);
Because each page may have its own CSS;
(However, downloading many pages from the same site is instead faster;)
CSS syntax is different from HTML and can be slightly ambiguous;
Not a user-friendly language (but available as IDE) that the developer needs to learn;
(*eg* the same cascade style name may have slightly different effects when used by different browsers;)
An HTML page saved without CSS, and seen offline, will not display nicely;
Because it would require access to external files, including CSS (including images, video, sound);
Anybody with read/write permission to the CSS can easily override just the CSS;
Showing unwanted information to an entire web-site (risk of hacking in just one point/internal threat with high risk);
Some browsers do not support exactly the same CSS (some styles/fonts for example);
The view can be slightly different using different browsers/risk of slight incompatibility;
(For large, corporate sites) Integration of HTML + CSS with other content management systems (CMS) may be tricky;
Because CMS also use their own CSS, and there can be collisions, and it requires technical competence; **[6]**

(c) *Award up to* ***[3 max]****.*
(When a request of download is received at the Application level (eg HTTP with a browser), appropriate data packets are generated and sent to the transport layer/TCP;)

TCP uses the header of the data packets to order them and to **check** their contents;
And **sends an acknowledgement** signal to the transmitter upon correct reception, before preparing a new data packet with a further header for the IP/Internet level;
If the transmitter does not receive the acknowledgement from the TCP, the **transmitter re-sends** the packet;
Therefore, all correct packets will be available to the TCP at some point to be made available to the IP and specifically to the IP destination/receiver (in the header);
IP ensures that the package reaches the correct address; **[3]**

(d) *Award marks as follows up to* ***[4 max]****.*
Award ***[1 max] each*** *for outlining the* **two** *techniques (lossy and lossless).*
Award ***[up to 2 max]*** *for relating both to actual applications (****[1]*** *if related to file types instead).*

Images and audio can be compressed using lossy techniques;
This will create a smaller file/faster transmission time;
But the data loss will not seriously affect quality;

However, CSS and HTML files might lose possible scripting;
As you need to recover the original data;
Lossless compression is more suitable; **[4]**

8. (a) *Award up to* ***[4]****.*
Server-side scripting does not require the installation of extra resources on a computer, for example the latest plug-in to run the script;
The log-in process will generally be slower as a round trip has to be made;
However, for regular visitors of the site, it speeds up the login, as the server script might be linked to data storage and recognize their access for further reference;
The server script is independent from the version of the browsers that is used, the user just sees the HTML interface (this is the typical problem with client-side scripting);
A security breach on the server could lead to all log-ins being compromised;
However, server-side is more secure as client-side is more easily hacked into;
There can be occasional downgrade in performance if too many users are running simultaneously their scripts on the server;
However, for the nature of the files, this should not generate bottlenecks in the network (small files, not intensive streaming); **[4]**

(b) *Award up to **[4 max]**.*
Your email address and associated password are private;
Both are needed to register with the initial website, but at most only the email address is stored there;
They are forwarded to the third-party website for identification;
The third-party identifies/confirms whether or not the person is who they say they are, based on their email/social network services that the third party provides;
This yes–no answer/confirmation provides the authentication for the open-source developers; **[4]**

(c) (i) *Award **[1]** for stating how the URL is generated, and **[1]** for giving an example, up to **[2 max]**.*

It is generated by taking the path of the website extended with the path for the code;
eg www.OpenSourceDev.org/newcode/json/fakecode007; **[2]**

(ii) The script accesses the file of new URLs (in a file);
And adds the URLs to an email template;
Then accesses the file of email addresses of subscribers (and sends email); **[3]**

(d) *Award marks as follows, up to **[3 max]**. (**Note:** there are **[5]** marking points)*
*Award **[1]** for the value passing from the form to php via `POST`.*
*Award **[1]** for the operation of the code to generate the short URL with final `$v`.*
*Award **[1]** for access to update database.*
*Award **[1]** for a new correct triple in database.*
*Award **[1]** for remote access operated by server through the database to retrieve the URL.*
***Note:** Competent use of terminology is needed to award full marks. Generic answers that show general good understanding of the process, but lack use of competent language should not be awarded full marks.*

Example
The value of the URL that is inputted (and submitted) via the form is stored in the variable URL which Interacts with the php code, because the method `POST` is specified in the form, therefore starting a communication process;

The process generates a different value for the name for the URL that is used to update a database of URLs together with the IP address;

The IP address is retrieved by the server through a remote call requested from the insertion in database;

The new value for the URL that is generated, given the one in input, is `thenhpid.td` and is stored in `$v`;

(detail of intermediate operations:

```
$x = thendorgbinphpid
$y = thenhpid
$z = td
$v = thenhpid.td)
```

The new record inserted in the database url_db will be the following triple
www.the2nd.org/bin.php?id=70, thenhpid.td, 172.16.254.1; **[3]**

9. (a) *Award* ***[3 max]*** *for* ***three*** *differences stated.*

Dynamic allows user interaction;
Allows parts of the content to be changed without uploading the complete page;
Can connect with server-side databases;
Gives different views for different users;
Includes the use of scripts / server side scripting, examples of;
Dynamic makes use of templates;

Static webpages are changed only by administrator;
It gives the same view to all users;
It displays exactly the information that is stored in the html file (and it may still include multimedia);
Simpler to develop/maintain, because it does not require major design/knowledge of web applications;
Source code edited directly in HTML generally shorter/less cumbersome than the one developed with web applications; **[3]**

(b) *Award* ***[1]*** *for identifying a service and* ***[1]*** *for stating a benefit, for* ***two*** *services up to* ***[4 max]****.*

Book/pre-pay online for tickets;
More convenient for visitors (avoids closures/queues *etc*);

Add a virtual tour of the museum online;
Increases the visibility of the museum / It becomes better known;

Add an audio-guide for smartphone to be downloaded from the portal;
So that users are already equipped before they visit;

(Targeted) pop up adverts/links;
Can increase revenue to the museum; **[4]**

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

More possibility for the users to interact will make the website more popular;
More services online provides more ways for the museum to be linked from/to other external organizations/institutions (*eg* bank/tourist office);
The two combined would generate more traffic and more authoritative in-links;
This increases the possibility for the site to be found in a variety of ways when searched;
And this increases the website’s ranking; **[4]**

(d) *Award marks as follows up to **[3 max]**.*
*Award **[1]** for an example/idea of a meta-tag and its purpose.*
*Award **[1]** for understanding the link between meta-tags and indexing via the search operation by crawlers.*
*Award **[1]** for understanding the link between indexing and ranking via optimization/better classification by SEO.*

Use of correct keywords allows the search engine to search and find those terms;

For Example
Title tags / keywords associated to the title of the page;
A description with tags may induce more visitors to visit the page;
Other keywords that appropriately refer to the content of the web;
(Even irrelevant keywords or hidden content however keyword stuffing might actually be punished by more recent algorithms for ranking);
Meta attributes for robots (such as follow/nofollow, index/noindex) that give indication about whether or not to index the page;
Meta language descriptors (for example to recognize in some language rather than English) to appropriately index content based on the dictionary of that language;
This is related to indexing, because web crawlers visit pages and inspect their content based on the terms that they have to search;
The ranking algorithms give value to the pages once that they are indexed, based on parameters that depend also on the ease of retrieving the pages/number of authoritative in-links/number of hits (or visitors);
The Search Engine can be optimized so that the best use of meta-tags can be made with the objective of indexing and ultimately ranking; **[3]**

Option D — Object-orientated programming

10. (a) *Award **[1]** for identifying the code and **[1]** for identifying the new position.*
`private int fiCount = 0` and `private int diCount = 0`;
In the variable/attribute section of the class and not in the constructor); **[2]**

(b) ***EITHER***
The values of these (static) variables;
Are the same for all objects;
They belong to the class (not to the objects);
And are only created/declared once;
Note: *Do not accept "cannot be changed".*

OR
The static methods;
Are class methods (not object methods);
Are independent of the objects in the class; **[2]**

(c) *Award **[2 × 2]** for each clear example. Only award **[2]** for an example if the benefit is clearly shown;*

Having different data types allow different operations to be carried out depending upon the type;
eg
Calculations in the integer "quantity" variable;
which are not possible for a String;

Double instead of integer allows the use of decimals;
which mirrors real-life scenarios/allows more precise calculations;

Using an array (of objects);
Allows individual items to be accessed/processed more easily;

Memory usage can be reduced;
eg "`fiCount`" takes up less space as an integer than it would as a double (assuming it's just a number); **[4]**

(d) The variable `fi` is declared as an array of `FoodItem` objects / of type `FoodItem`;
With (a maximum of) 100 values;
It cannot be directly accessed/it is encapsulated within the Payment class; **[3]**

(e) *Award **[1]** for each correct line of code.*

```
public void addFoodItem(FoodItem f)
{
  fi[fiCount] = f;
  fiCount++;
}
```

[3]

11. (a) *Award **[1]** for each correct value.*
2;
0.2;
d103; **[3]**

(b) (i) `System.out.println(tables[40].getDiCount());` **[1]**
Allow variations of the `get` method name.

(ii) `System.out.println(tables[2].getFi(2).getItemCode());` **[1]**
Allow variations of the `get` method name.

***Note:** Ignore minor syntax errors in both parts (b)(i) and (b)(ii).*

(c) *Award marks as follows up to **[6 max]**.*
*Award **[1]** for correct initialization.*
*Award **[2]** for loop that checks entries, but has early exit.*
*Award **[1]** for a loop that checks all 200 entries with no early exit.*
*Award **[1]** for correct assignment.*
*Award **[1]** for correct comparison.*
*Award **[1]** for correct return value.*
*Award **[1]** for both the assignment and comparison **IF** `get` methods are not used.*

```
public static double findPrice(Item[] pl, String c)
{
  int x = 0;
  double price = 0.0;
  boolean found = false;
  while(!found)
  {
    if ((pl[x].getCode()) == c)
    {
      price = pl[x].getPrice();
      found = true;
    }
    x++;
  }
  return price;
}
```

[6]

(d) *Award marks as follows up to* ***[7 max]****.*
Award ***[1]*** *for correct method header.*
Award ***[1]*** *for correct initialization.*
Award ***[1]*** *for correct loop.*
Award ***[1]*** *for finding item code.*
Award ***[1]*** *for use of* `findPrice` *method.*
Award ***[2]*** *for the calculation if completely correct, award only* ***[1]*** *if taxes are wrong and/or no use of accessor.*
Award ***[1]*** *for consideration of both food and drink objects (whether correct or not);*

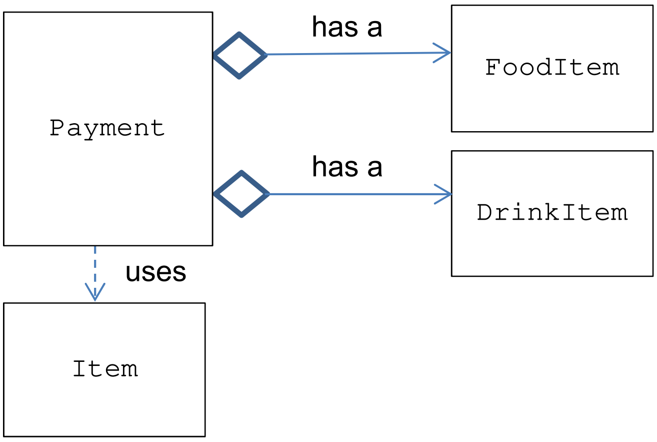
```
public double calculateBill(Item[] pl)
{
  double total = 0.0;

  for(int x = 0; x < fiCount; x++)
  {
    String c = fi[x].getItemCode();
    double price = findPrice(pl,c);
    total = total + fi[x].getQuantity()*price*(1 + foodTax);
  }

  for(int y = 0; y < diCount; y++)
  {
    String c = di[y].getItemCode();
    double price = findPrice(pl,c);
    total = total + di[y].getQuantity()*price*(1 + drinkTax);
  }
  return total;
}
```

[7]

12. (a) *Award marks as follows, up to* ***[3 max]****.*
Award ***[1]*** *for 4 classes connected with arrows/lines.*
Award ***[1]*** *for "has a" label or correct arrow from* `Payment to FoodItem` ***and*** *"has a" label or correct arrow from* `Payment` *to* `DrinkItem` *(with no additional arrows).*
Award ***[1]*** *for "uses" label or correct arrow going from Payment to Item (with no additional arrows).*



[3]

(b) *Award* ***[1]*** *for identifying a feature,* ***[1]*** *for an example in the given context, and* ***[1]*** *for an elaboration, for two features up to* ***[6 max]****.*
Encapsulation places all attributes and methods that relate to a particular object/entity together;
For example, `Payment` class includes attributes such as the food and drink arrays and methods such as `calculateBill()`;
This provides a clearer view/understanding of each section of the problem;
Which can lead to more efficient programming (faster, less errors *etc*);

Encapsulation protects the values of the data stored within the object;
From (accidental) changes made by other classes;
For example, quantity in the `FoodItem` class cannot be altered through another variable called quantity in another class;
This allows programmers to select any variable names they wish/no restriction on choice of variable names; **[6]**

(c) *Award up to* ***[4 max]****.*
A generic `Payment` class can be created;
This would contain data/variables/methods required by all units of the company;
Each of the different operations could then inherit this class;
Adding new variables/methods that relate only to them;
Overriding the superclass methods as necessary;

Note: *Allow a similar answer that deals with the different items (*`FoodItem`*,* `DrinkItem` *etc).* **[4]**