UNIVERSITY OF CAMBRIDGE COMPUTER LABORATORY

Computer Science Tripos

Marking Scheme
and
Classing Convention

A Guide to Computer Science Examining in Cambridge from the setting of questions to the preparation of class lists

February 2012

Preface

This document is concerned with Part IA, Part IB, Part II and Part III Computer Science Tripos examinations, administered by the Computer Laboratory. Paper 1 of Part IA is also taken by candidates offering subject Computer Science of Part IA of the Natural Sciences Tripos and by those taking the Introduction to Computer Science option in Part I of the Politics, Psychology and Sociology Tripos.

The first edition was originally prepared for, and approved by, the Computer Science Syndicate in 1997.

Subsequent editions have principally been concerned with the unification of Part IA of the Computer Science Tripos with Part IA of the Natural Sciences Tripos. The classing procedures for both these examinations now make use of a common order-of-merit table.

Changes from October 2010 mainly relate to candidates who cannot be classed by the Examiners. In addition, the cohort-tracking process in Part IB and Part II is amplified.

Changes from October 2011 and February 2012 consist of a change of viva requirements for Part II and a new section on Part III.

Contents

1 — Overview of the Examinations

Parts IA, IB and II of the Computer Science Tripos (CST) include written papers and assessed practical work. The following table shows the components of the different Parts and indicates the maximum mark available for each:

Examination	*Papers*	*Assessed Practical*	*Maximum Mark*
Part IA	1, 2 (+ 2 others)	20 Assessed Exercises	375
Part IB	3, 4, 5, 6	7 Assessed Exercises	400
Part II	7, 8, 9	Project/Dissertation	400
Part III	6 *modules*	Project/Dissertation	900

Each written paper and the Part II Project/Dissertation is marked out of 100 (though Papers 1 and 2 are later scaled to be out of 80). Part III modules are marked out of 100 in the same way as MPhil modules. Details of the way credit is awarded for Assessed Exercises are given in the sections relating to Notes on Part IA and Part IB below.

Note that Paper 1 is an option for candidates taking Part IA of the Natural Sciences Tripos (NST) or Part I of the Politics, Psychology and Sociology Tripos (PPST).

The table omits numerous important details and further explanation is necessary...

Notes on Part III

A Part III candidate offers 6 modules and is also required to submit a dissertation (worth 3 modules) of up to 12,000 words.

Notes on Part II

A Part II candidate offers only three papers but is also required to submit a dissertation of up to 12,000 words. The dissertation mark is scaled so as to be out of a maximum of 100 and the score is added to the total mark on the three written papers to give a theoretical overall maximum of 400 for the examination.

Notes on Part IB

A Part IB candidate is also required to submit a portfolio but only seven ticks are on offer. Five of these ticks are awarded for exercises undertaken by a candidate working as an individual and two further ticks are awarded for working on a group project, one being for the group report and the other for the candidate's personal report.

Every candidate is expected to gain a full set of ticks but, instead of assigning positive marks for each tick, any shortfall from the required 7 ticks incurs a penalty of 14 marks per missing tick.

Any such penalty is subtracted from a candidate's overall total for the four written papers. Such penalties are rare.

Notes on Part Ia

Five categories of candidates may be distinguished:

1. CST+NST Candidates

   Computer Science candidates who offer Papers 1 and 2 of the CST and subject Mathematics and a bench subject of the NST.

2. CST+Mathematics Candidates

   Computer Science candidates who offer Papers 1 and 2 of the CST and Papers 1 and 2 of the Mathematical Tripos.

3. CST+PPST Candidates

   Computer Science candidates who offer Papers 1 and 2 of the CST and subject Mathematics of the NST and Paper 3 of the PPST.

4. NST+CST Candidates

   Natural Science candidates who offer Paper 1 of the CST as one of their four NST options, the other options being subject Mathematics and two bench subjects.

5. PPST+CST Candidates

   Politics, Psychology and Sociology candidates who offer Paper 1 of the CST as one of their options, the other options being regular PPST papers.

Since 2004, Part Ia of the Computer Science Tripos and Part Ia of the Natural Sciences Tripos have been treated almost as a unified examination. All Part Ia CST and NST candidates offer four subjects and the only difference of consequence is that separate class lists are published.

Each of CST Papers 1 and 2 is treated as though it were an NST bench subject; in each case there is a written paper and associated assessed practical work. Each written paper counts 80 raw marks and the associated practical work counts 20 raw marks. This is exactly as for the NST subject Chemistry.

As with NST subjects, the raw totals (out of $80 + 20$) for each of Papers 1 and 2 are post-processed by a procedure known as *norm-referencing* to ensure uniformity across subjects. Thus a given processed mark in, say, Chemistry is intended to indicate merit equivalent to the same mark in, say, CST Paper 1.

For the relevant candidates, the raw marks of Papers 1 and 2 of the Mathematical Tripos and of Paper 3 of the Politics, Psychology and Sociology Tripos are also norm-referenced to ensure uniformity.

In all cases, the norm-referencing results in a mark in the range 0 to 100 but the four-subject total is not out of 400. A footnote to Regulation 19 of the Natural Sciences Tripos requires the norm-referenced mark for Mathematics to be multiplied by 0.75 before totalling, so the four-subject maximum is actually 375 marks.

CST+Mathematics candidates are the only group noted above who do not offer NST Mathematics. To ensure that these candidates have the same four-subject maximum, their

norm-referenced Mathematical Tripos marks for *both* Paper 1 and Paper 2 are multiplied by 0.875 which means that the maximum total for these two papers together is 175 marks.

Every candidate is required to submit a portfolio of assessed Computer Science exercises. There are 10 exercises associated with each of CST Papers 1 and 2.

The exercises are undertaken during the year leading to the Tripos. A satisfactory exercise gains what is colloquially known as a 'tick'. A rejected exercise may be resubmitted, in principle any number of times, but it is rare to require more than one resubmission.

Shortly before the written papers are taken, the Part IA examiners are supplied with a final tick list and each tick accounts for two raw marks. Fractions of ticks are not awarded so the raw marks for each exercise will be 0 or 2.

Regulations require Part IA candidates to submit real portfolios; securing ticks is not sufficient. Submission after the final deadline specified in the Announcement by the Head of Department incurs a penalty of 2 marks per day late. Fractions of days are rounded up and penalties are subtracted from the raw marks awarded to Paper 1.

Notes on Part IA — PPST+CST Candidates

The fifth category of candidates who may offer Paper 1 relates to PPST candidates who borrow this paper which, for them, is PPST Paper 9.

The procedure is to examine any such candidates exactly as for any other candidate offering CST Paper 1. The PPST examiners are supplied with the raw marks of each such candidate together with any requested summary information. [This is exactly as happened for many years with Mathematics candidates who used to borrow CST Paper 1.]

2 — Marking Scheme for Parts IA, IB and II Written Papers

The questions set in the written papers are required to accord with the published Form and Conduct Notice (see Appendix A).

The 9 papers of the Computer Science Tripos each contain between 8 and 14 questions and candidates are always asked to attempt 5 questions with 20 marks available per question.

Every question is based principally on material presented in a particular course of lectures and is normally set and marked by the lecturer who gave the relevant course.

The Structure of the Written Papers

The structure of the papers is planned in time for copies to be made available to every candidate at the start of the academic year (see Appendix A). Every entry in the structure relates to a question and specifies the course on which the question will be based.

Each entry in Papers 3 to 9 includes the initials of the (main) course lecturer.

A typical question is shown in Appendix B. Each question is heralded by the title of the most relevant course and the breakdown of the 20 available marks is clearly indicated. Questions from previous years are available via the Computer Laboratory's web site (see

Appendix A). Each question is vetted by at least two internal examiners and, for Parts IB and II, the External Examiner.

Candidates for Part IA and Part IB are expected to attend all the courses which are arranged for them. Accordingly, Papers 1 and 2 are arranged in sections with a rubric which ensures that candidates attempt questions relating to a spread of subjects, and each of Papers 3 to 6 concentrates on a particular area of Computer Science (Programming, Applications and Professionalism, Systems, and Theory).

Candidates for Part II commonly plan (in consultation with Directors of Studies) to attend a subset of the courses which are arranged for them. The availability of the examination structure at the beginning of the year facilitates the planning.

Marking the Written Papers

The scripts constituting the attempts at a given question are normally sent for marking to whoever set the question. The scripts are sent with a covering note, a typical example of which is shown in Appendix C. This includes a reminder that the question should be marked out of 20 and it is implicitly understood that the marking scheme indicated in the question will be adhered to.

A typical candidate attempts 20 questions in an examination and these questions are likely to have been set by 15 or more different individuals. Accordingly, the scripts from a given candidate are marked by a range of markers. As a means of moderating the marking process, the External Examiner scrutinises a number of marked scripts selected at random.

3 — Marking Scheme for Part IA and Part IB Assessed Exercises

As noted, candidates for Part IA and Part IB are required to submit portfolios of assessed exercises. University Ordinances require formal Announcements by the Head of the Department outlining the work which is to be undertaken. Copies of the announcements are issued at the start of the academic year (see Appendix A).

Assessed exercises are of two kinds: individual exercises and group projects. Group projects apply only to candidates for Part IB.

Satisfactory solutions to the individual exercises are awarded ticks by those who are responsible for the associated practical classes. A submission has to pass a threshold of acceptability *and* the candidate has to satisfy an assessor at a short interview (5 to 10 minutes).

An interview enables an assessor both to check that a candidate did not just copy the submitted solution and also to make critical comments before awarding a tick. Given the importance attached to interviews, the system of anonymous marking, which is universally employed for the written papers, is not used for assessed exercises.

Only a minimalist marking scheme is involved. The guidance given to assessors is simply that most candidates are expected to gain a tick at the first submission and that all candidates are expected to gain all their ticks before some specified deadline (close to the start of the examinations).

The assessment of group projects is somewhat different. Each member of each group must attend *all* the formal meetings and write a personal report. Each such report includes summary assessments of the contributions made by the other members of the group. The proprietors of the group projects take these assessments into account when awarding ticks. No resubmission is possible for group projects but most candidates duly acquire their two ticks.

Further information about group projects is given in the Group Project Briefing Booklet which may be found via:

`http://www.cl.cam.ac.uk/teaching/GroupProjects`

4 — Marking Scheme for Part II Dissertations

Managing and assessing Part II projects is undoubtedly the most difficult aspect of the entire examining process. Very considerable thought has been directed at establishing satisfactory procedures and valuable advice has been provided by external examiners. Key features of the current practice are:

- All Part II candidates are required to attend two briefing lectures which describe exactly what is expected of them.
- The information given in the briefing lectures is amplified in a *Briefing Document* which gives full details of the assessment procedure and also includes a copy of the Guidelines for Assessors (see Appendix A).
- Every candidate has the support of a Supervisor and every project is vetted and monitored by two Teaching Officers colloquially known as 'Overseers'.
- All Part II candidates are required to write a Progress Report and to make an associated presentation.
- Every dissertation is required to follow the same format which includes five standard chapter headings.
- Dissertations carry the names of their writers. Anonymous marking would be difficult to achieve in a laboratory environment. Over a period of several months, assessors necessarily get to know many candidates and become familiar with the projects they are working on.
- Every dissertation is read by at least three readers (the three internal examiners).

The Guidelines for Assessors shows the marking scheme and the Assessment Form. This form asks for provisional marks and their total (out of 50 at this stage) and asks whether a *viva voce* examination or additional assessment by an expert should be considered.

The idea of additional expert assessment was suggested by an external examiner. If the three first-readers feel that the subject of a dissertation is outside their area of competence, the opinion of a specialist is sought. A specialist may also be brought in when there is disagreement about a particular case.

When all the dissertations have been read and marked by three readers, the raw marks are processed using the procedure described in detail in Appendix D.

In essence, the raw marks are first normalised so as to give a common mean and standard deviation for each assessor. The normalised marks for each dissertation are then averaged to give a provisional mark for the project. This is referred to as the Calculated mark in Appendix D.

An order-of-merit table is drawn up and the readers (perhaps in the presence of the External Examiner who will also have read some of the dissertations) then discuss every dissertation in turn. When the individual marks indicate broad agreement, the Calculated mark is normally accepted with little discussion and becomes the Decided mark.

When there is disagreement there may be considerable discussion, one or more experts may be consulted, and the Decided mark may be several marks different from the Calculated mark. Candidates may be invited to a *viva voce* examination, at which additional expert assessors may participate It is expected that in any year, around 5–10% of candidates will be examined by viva.

Consulted experts are not asked to provide marks (since they typically read an insufficient number of dissertations for self-calibration) but are instead invited to supply comments. These comments and the views of the External Examiner may lead to the provisional mark being adjusted. An interesting feature of triple-marking is that it is not unusual for the wayward mark of three normalised marks to be closest to the Decided mark.

5 — Classing Convention: Computer Science Tripos (Parts I and II)

There are three principal stages in classing Parts I and II of the Computer Science Tripos:

- For each candidate, determine an overall grand total mark.
- Order the candidates by their overall marks, thereby deriving an order-of-merit table.
- Partition the order-of-merit table into classes.

Further details are given in the following sections.

Unsurprisingly, the procedures used in Part IA are the most complex, principally because so many options are available. In consequence, there is far more to say about Part IA than Part IB or Part II.

Marks: Written Papers, Assessed Exercises, Dissertations and Penalties

In every written paper in the Computer Science Tripos candidates are asked to attempt five questions where each question is marked out of 20.

In Part IA the total mark for each written paper is multiplied by 0.8 to reduce the maximum possible mark to 80; there is then further scaling (norm-referencing) which is described in the next section.

There is no scaling of total marks in Part IB or in Part II and there is no scaling of marks for individual questions in any part of the Computer Science Tripos.

In Part IA each successfully completed assessed exercise is awarded two raw marks. Late submission of the portfolio incurs a penalty of two raw marks per day late.

In Part Ib any shortfall from the required total of 7 ticks incurs a penalty of 14 marks as explained earlier.

In Part II the dissertations are marked using the procedure described previously. Late submission of a dissertation incurs a penalty. If n is the integer part of the number of days late, the penalty is:

$$penalty = \frac{10 + n}{40} \times mark$$

Part Ia — Norm-Referencing CST and NST Papers

The procedure for deriving the overall grand total mark for each candidate in Part Ia of the Computer Science Tripos exactly follows that used by examiners for the Natural Sciences Tripos. A full description of this procedure is given in the documentation supplied to NST examiners but, in outline, it is interpreted as follows:

- For each of Paper 1 and Paper 2, each candidate is assigned an adjusted raw mark m thus:

  $$m = 0.8w + 2t$$

  where w is the raw mark out of 100 for the written paper and t is the number of ticks (successfully completed associated assessed exercises). Given that the maximum value of t is 10, the possible range for m is 0 to 100. Any penalty for the late submission of a portfolio is deducted from the Paper 1 adjusted raw mark.

- For each paper, all candidates who offered the paper are ranked into descending order of their m totals. Those, such as possible PPST candidates, who are neither CST nor NST candidates should be *included* but those who were absent (possibly having withdrawn from the examination as a whole) should be *excluded*.

- Naïvely partition the ranked list into three groups: the top 25%, the next 65%, and the bottom 10%. In NST terminology, these three sections are labelled the top section, the middle section and the bottom section respectively for the paper.

- The naïve partitioning probably results in candidates with the same marks being in different sections, so shift each partition up or down (normally in whichever direction requires the less movement) to ensure that candidates with the same marks are in the same section.

- Piecewise-linearly scale the marks as follows:

Informal Label	*Raw Range*			*Scaled Range*		
Top Section	max	to	a	100	to	70
Middle Section	a	to	b	70	to	50
Bottom Section	b	to	0	50	to	0

  where: max is the maximum possible mark for the paper
  a is lowest raw mark in the Top Section, and
  b is lowest raw mark in the Middle Section.

- In CST Papers 1 and 2 (but not in every NST subject) $max = 100$.

- In CST Papers 1 and 2 (but not in every NST subject) every raw mark is an integer. In particular the lowest raw mark by a top section candidate, a, is an integer and the highest possible raw mark for a middle section candidate is therefore $a - 1$. The norm-referencing algorithm will scale this mark to $50 + 20\,(a - b - 1)/(a - b)$ which is less than 70. In general, a scaled mark will not be an integer.

- A scaled mark strictly less than 40 is sometimes regarded as an informal Fail for the paper but this and the other labels have no formal significance. Candidates are not required to pass each CST paper and each NST subject or any paper borrowed from the Mathematical Tripos or from the PPST individually.

- The above steps result in norm-referenced marks out of 100 for both CST papers and all NST papers. The norm-referencing of the two CST papers is undertaken by the CST examiners and the norm-referencing of each NST paper is undertaken by the examiners for that paper.

[Note: analysis of data over several years shows that norm-referencing (as just described) leads to results which are very close to those that would stem from the more formal use of Z-values.]

Part Ia — Norm-Referencing Mathematics and PPST Papers

The Mathematical Tripos examiners and the PPST examiners do not use norm-referencing, so the CST examiners have to norm-reference the marks of CST candidates who offer Papers 1 and 2 of the Mathematical Tripos and the marks of CST candidates who offer Paper 3 of the PPST.

The CST examiners therefore have to ask the Mathematics examiners and the PPST examiners not only for the raw marks obtained by CST candidates but also for the associated values of max, a and b in the three cases.

It is well known that Mathematics uses an elaborate system of alphas and betas in lieu of norm-referencing but, to ensure harmony with the other options available to CST candidates, the CST examiners do not use these alphas and betas.

Accordingly, the CST examiners make use of the raw marks awarded for Paper 1 and Paper 2 of Part Ia of the Mathematical Tripos where, for both papers, the maximum is well defined: $max = 140$.

The first year that CST candidates were able to borrow Papers 1 and 2 of the Mathematical Tripos was 2009 when valuable advice was given by John Lister of the Faculty of Mathematics. In particular, he noted that the appropriate proportion of candidates for the top section is 30% rather than 25%.

This advice was heeded and, for the Mathematics papers, a is the mark of the 30-percentile candidate; b is again the mark of the 90-percentile candidate.

In practice, the Mathematical Tripos examiners handed over the raw data for *all* candidates who offered Papers 1 and 2 and left it to the CST examiners to determine a and b.

Attending to the CST candidates who offered Paper 3 of the PPST is relatively simple. The PPST examiners supplied the raw marks of the CST candidates and max, a and b.

In this case, $max = 100$ and a and b are the marks of the 25-percentile and 90-percentile candidates.

In summary, the CST examiners norm-reference the raw marks of five papers: the two CST papers, Papers 1 and 2 of the Mathematical Tripos, and Paper 3 of the PPST.

Part Ia — Classing

When all four norm-referenced marks are available for every CST and NST candidate, one can order the candidates by their overall totals. The result is the unified order-of-merit table which includes all CST and NST candidates.

In forming the overall totals, the marks for NST Mathematics are scaled by a factor of 0.75 and the marks for each of Papers 1 and 2 of the Mathematical Tripos are scaled by a factor of 0.875 so the overall totals are, for all candidates, out of a maximum possible 375 marks. The marks should not, of course, be rounded before totalling.

Unfortunately the realities of the examination timetable make it impractical to use the unified order-of-merit table itself for classing. Also, there are Natural Science examiners who view the incorporation of CST candidates into their table as introducing bias.

In consequence, the class boundaries required for classing both NST candidates and CST candidates are obtained from the NST-only order-of-merit table and not the unified order-of-merit table.

Analysis of the NST-only order-of-merit tables and the unified order-of-merit tables in recent years shows that including CST candidates does indeed change the boundaries slightly but using the unified table would not have changed the class of any CST candidate. This is fortuitous but fortunate.

In outline, the NST-only order-of-merit table is partitioned 25:35:32.5:7.5 to determine the three threshold marks for a Class I, Class II.1 and Class II.2. There is no divided second class in NST so the NST examiners use only two of the threshold marks. The CST examiners use all three.

The detailed procedure for establishing the class boundaries is:

- Draw up the NST-only order-of-merit table by ordering the NST candidates by their overall total marks out of 375.
- Subject to the rules on Exclusions and Inclusions given in the next section, note the overall total marks of the three candidates who are respectively at the bottom of the top 25-percentile, the bottom of the top 60-percentile and the bottom of the top 92.5-percentile.
- The bottom mark of the top 25-percentile and the bottom mark of the top 92.5-percentile are taken, by the NST examiners, as the threshold marks for Class I and Class II respectively. There is no divided second class in Part Ia NST.
- These same bottom marks are taken by the CST examiners, as the threshold marks for Class I and Class II.2 respectively and the bottom mark of the top 60-percentile is taken as the threshold mark for Class II.1.

- Note that the threshold for a II.2 is the mark of the bottom 92.5-percentile candidate and not the mark of the bottom 90-percentile candidate (the percentile used for norm-referencing). This is a result of a decision taken by the NST examiners in 2007.
- By long-established tradition 150 marks out of 375 is taken as the threshold for a III. Candidates whose total is below this threshold are deemed unclassed and for each such candidate a note is written by the examiners for consideration by the Applications Committee.

Norm-referencing has been used by Natural Sciences Tripos examiners for many years to ensure uniformity across the Natural Science subjects. Part Ia CST examiners depend on this uniformity to ensure that candidates offering different combinations of subjects are treated equally. In particular, the procedure ensures that Computer Science candidates who offer Papers 1 and 2 of the CST, and Natural Science candidates who offer only Paper 1, are treated in the same way.

The Part I PPST examiners almost certainly operate a quite different procedure. They are sent full details about the performance of those of their candidates who offer Paper 1 of the CST.

Part Ia — Exclusions and Inclusions

Given an order-of-merit table, determining the top 25% sounds straightforward until one poses the question, '25% of what?'.

For the two CST papers, the individual NST subjects, the two Mathematical Tripos papers and the PPST paper, everyone who offers the paper or subject is counted except any candidates who are absent. In particular, candidates who are neither CST nor NST candidates are included.

When considering the unified NST+CST order-of-merit table, the question of who should be included is more difficult. It is important to compare like with like so PPST candidates, certain Education candidates and occasionally other candidates (such as those on an exchange programme) are excluded.

There are several special cases and the following rules have been established:

Excluded Candidates

- Anyone who is not formally entered for either Part Ia NST or Part Ia CST.
- Anyone who withdraws from the Examination or who is absent from at least one paper. Some of these candidates may end up under the heading 'Declared to have deserved honours'.

Included Candidates

In general, every candidate who is not ruled out by the two exclusions is counted; in particular, the following are included:

- Any not-for-honours candidate.
- Anyone who is below the threshold for a Third.

- Any candidate who requests not to appear on the published class list.

These explicit inclusions are important because they do not accord with strict Baxter analysis (see page 11). Baxter himself can work only from published class lists and cannot know how a not-for-honours candidate performed and cannot infer the existence of failures. Baxter's own notes suggest that he would like to include not-for-honours candidates and failures and, although the information is not available to him, it *is* available to the examiners.

Part Ia — Anomalous Firsts

It is quite possible for a candidate to be in the top 25% of the order-of-merit table for the examination as a whole without being in the top 25% of any individual CST paper or other paper. It is thus possible, though rare, for a candidate to obtain a First overall without obtaining a single top-section mark.

Part Ia — Published Marks

When the NST marks are published, it is the norm-referenced marks that apply. For each NST bench subject, these marks are broken down into marks for the written paper and marks for the practical work. For NST candidates who offer CST Paper 1 a similar breakdown is required by the NST examiners...

An NST candidate who scores w marks for the Paper 1 written paper and obtains t ticks for the associated practical work has an adjusted raw mark $m = 0.8w + 2t$ and m is the value which is norm-referenced. Suppose:

$$n = N(m)$$

Here n is the norm-referenced value and N is the piecewise-linear norm-referencing transformation function. In general $n \neq m$ and it is n which has to be partitioned into a norm-referenced written-paper score and a practical mark. In conformance with NST practice, n is partitioned as follows:

$$\text{Written-paper score} = 0.8w\,\frac{n}{m} \qquad \text{Practical mark} = 2t\,\frac{n}{m}$$

Clearly, these marks are in the ratio $0.8w : 2t$ which makes sense. The value of $\frac{n}{m}$ is not the same for all candidates, and two candidates with the same number of ticks, t, may be shown with different published marks for the practical work.

Part Ia — The Baxter Average: Introduction

The so-called Baxter Average for a Tripos examination is the outcome of applying a procedure originally due to Norrington. The result is a numerical score that provides a merit figure for an entire set of candidates.

As noted earlier, Baxter himself can work only from published class lists and these hide information that *is* available to examiners.

Examiners for the Computer Science Tripos make use of the information that Baxter would like to have but, in general, does not have. The procedure actually used is to consider all

included candidates (not-for-honours candidates are considered too) and score 5 points for a First, 3 for a II.1, 2 for a II.2, 1 for a III and 0 for unclassed candidates.

Excluded candidates are omitted from the analysis. The average score for all included candidates is the overall Baxter Average for the examination.

Part Ia — The Baxter Average: Worked Example

Suppose there are 69 included candidates in CST Part Ia with 11 in Class I, 26 in Class II.1, 25 in Class II.2, 6 in Class III and one unclassed. The Baxter Average is therefore:

$$\frac{5 \times 11 + 3 \times 26 + 2 \times 25 + 1 \times 6 + 0 \times 1}{11 + 26 + 25 + 6 + 1} = \frac{189}{69} = 2.74$$

The figure is a quality indicator that can, *inter alia*, be used to compare the candidates in one Part Ia year with those in earlier years.

Part Ia — The Reference Examination

Part Ia can reasonably be regarded as a reference examination. The examiners for each individual NST paper and subject apply the 25:65:10 partitioning to the entire set of candidates offering that paper or subject regardless of origin. In general there will be Computer Scientists, Natural Scientists and candidates from elsewhere.

There are around 700 candidates in the unified set of NST and CST candidates. The classing procedure followed by NST examiners (where the top 25% are awarded Firsts, and so on) leads to a Baxter Average close to 3 but there is no guarantee that the Baxter Average for the CST subset will be close to this value. In a good year it may be a little higher and in a bad year it will be lower as, for example, the 2.74 figure quoted above.

This rough means of measuring CST candidates against a much larger field is the first of two reasons for regarding Part Ia as a reference examination.

The second reason is that data from Part Ia can be used by Part Ib and Part II examiners when classing their candidates. The procedure used is known as cohort tracking...

Class Boundaries in Part Ib and Part II — Cohort Tracking

In Part Ib and Part II, order-of-merit tables based on overall grand total marks are drawn up, using the same rules for inclusion and exclusion as in Part Ia. The provisional partitioning is chosen to reflect past performance...

Suppose, in a hypothetical case, that only those who gained Firsts in Part Ia went on to Part Ib. It would be unduly harsh to award only 25% of them Firsts in their second year. Equally, if only those who gained II.2s in Part Ia went on to Part Ib, it would be unduly lenient to award 25% of them Firsts in their second year.

Such extreme cases are unlikely, but there is considerable evidence that those who continue from Part Ia to Part Ib are not uniformly distributed across the Part Ia class list. Some of the stronger Part Ia candidates leave Computer Science (often for Part Ib NST) and some of the weaker candidates leave too (to take their chances in other subjects).

The cohort-tracking procedure takes these departures into account and ensures that the class distribution of those who continue to Part Ib is much as it was in Part Ia. Of course, individual candidates may move across class boundaries.

The first step in the cohort-tracking procedure in Part Ib is to identify the subset of candidates who took Part Ia CST in the previous year. This is the cohort that is to be tracked. The next step is to note how many in the cohort were in Class I, Class II.1, Class II.2, Class III or were unclassed and their Part Ia Baxter Average is computed. [Note: there are usually Part Ib candidates who did not take Part Ia CST in the previous year.]

In the provisional partitioning, the class distribution of the cohort in Part Ib should be the same as it was in Part Ia and will thereby produce the same Baxter Average.

Note that the Baxter Average for the entire Part Ib class may be higher or lower than that of the subset, depending on whether the incomers who did not do Part Ia in the previous year are stronger or weaker than those they join.

The cohort-tracking procedure in Part II is a simple extension of that used for Part Ib. In Part II the subset that is identified is that which took Part Ib the previous year *and* took Part Ia two years previously.

The Part II average for the cohort should *not* be lower than their Part Ia and Part Ib averages but nor should it be significantly higher.

Further information and worked examples of the use of cohort tracking are presented in Appendix E where the procedures used to produce the Part Ib and Part II class lists in a typical year are described.

Other Points

It is, fortunately, very rare not to be able to class a candidate and, in such cases, the guidelines provided by the Board of Examinations are followed. These guidelines also include procedures for the possible special treatment of candidates who have been absent for part of the examination. Again, such cases are rare.

Candidates are not required to pass each written paper separately and, in Part II, are not required to pass the dissertation. There is a reasonable correlation between written-paper marks and dissertation scores and no policy requiring the separate passing of different components has ever been implemented in Part II. The rare candidate who submits an inadequate dissertation almost invariably performs very badly in the written papers too.

6 — Classing Convention: Computer Science Tripos (Part III)

Part III of the Computer Science Tripos is available to those students who attain a first in Part II (or those who are ranked above the median of students attaining honours in Part II, and who also satisfy one of the exceptional conditions relating to attainment in their dissertation, written papers, or the first two parts of the Computer Science Tripos). There is no cohort tracking or normalisation used for Part III; rather students obtain an

overall percentage score for the year, with 60% being the passing grade, 67% being "pass with merit" and 75% corresponding to "pass with distinction".

These scores are calculated by combining raw scores from individual elective modules with the score attained for the research project. Each of the six taught modules contributes 1/9th of the overall grade, while the project accounts for 1/3rd. In addition to attaining a passing grade overall, students are also expected to attain a passing grade for their research project.

In the process of Research Project selection, Part III students fix a *Project Supervisor* in conjunction with their Director of Studies. Because the Project Supervisor is an Assessor for the purposes of examining (i.e. provides a project mark to the examining system), he or she must be a University Teaching Officer at the Computer Laboratory or otherwise approved by Head of Department.

The project dissertation is dual-marked, with one assessor being one of the Part III examiners, and the other being the project supervisor. Each assessor produces a percentage score, and these are averaged to provide a provisional mark. Should the individual scores be widely discrepant, a third assessor may be used. In addition, students may be called for a viva voce examination, which may lead to adjustment of the provisional mark.

7 — Errors and Some Concluding Notes

Considerable care is taken over the processing of marks to reduce error. The following paragraphs explain the procedures employed in Parts I and II.

Candidates make out a separate cover sheet for each question attempted and the scripts for a given question are marked by a single assessor. The total mark awarded for a given script is written on the associated cover sheet, each total being the sum of the sub-marks awarded for the attempts for the separate parts of the question. The addition of these sub-marks is checked independently and cases of doubt are referred back to the assessor for clarification.

Each total is then transcribed twice: once into a spreadsheet and, separately, into a mark book. Two people attend to the bundle of scripts for each question. One of them calls out the candidate numbers and associated marks to the other who keys the marks into a spreadsheet. The two then swap rôles and the marks are written into a mark book.

In this way two pairs of eyes look at each hand-written mark. There is occasional conferring about doubtful digits. A common difficulty is distinguishing the digits 0 and 6. Such problems are often resolved by retotalling the sub-marks made on the script itself.

When all the marks for a paper have been transcribed into a spreadsheet and a mark book the spreadsheet is printed out and the hard-copy is compared with the mark book.

There are few other examinations in Cambridge where such care is taken!

The arrangements for conducting Computer Science examinations are believed to be sound. The procedures are subject to regular internal and external reviews and refinements are introduced when appropriate.

Dissertations are inherently more difficult to assess than written papers and assessed exercises and great care has been put into developing the procedures currently in use.

At the end of the examination period the Chairman of Examiners for Part Ia writes a report on the Part Ia examination. Additionally, the Chairman of Examiners and the External Examiner for the other Parts of the Tripos each writes a report on these examinations.

These reports are seen by the Faculty Board of Computer Science and Technology and the Tripos Management Committee. Recommendations are discussed and, where appropriate, procedures are modified.

Each Chairman's report includes a summary of attempts and marks, together with assessors' comments on how the candidates dealt with each question. These reports are available via the Computer Laboratory's web site (see Appendix A).

As a final thought, one may note that the guidelines which are issued by the examiners for Part Ia of the Natural Sciences Tripos have remained essentially unchanged for several decades, though minor adjustments to the text are made annually. This year-on-year consistency is laudable and it is intended that the framework presented in this document will achieve equal stability.

Alan Mycroft, TMC chair

February 2012

[Minimally editted from the 2010 Marking and Classing document of F.H. King.
Tripos Management Committee is reviewing the structure and content of this document.]

Accompanying Documents

Form and Conduct Notice:
http://www.cl.cam.ac.uk/teaching/exams/formcond.pdf

Structure of Papers 1 and 2:
http://www.cl.cam.ac.uk/teaching/exams/partiastruct.pdf

Structure of Papers 3 to 6:
http://www.cl.cam.ac.uk/teaching/exams/partibstruct.pdf

Structure of Papers 7 to 9:
http://www.cl.cam.ac.uk/teaching/exams/partiistruct.pdf

Head of Department Announcements:
http://www.cl.cam.ac.uk/teaching/exams/headofdeptnotices.pdf
http://www.cl.cam.ac.uk/teaching/exams/HofD-notice-p3-2011.pdf

Part II Project Briefing Document:
http://www.cl.cam.ac.uk/teaching/projects/pinkbook/

Guidelines for Assessors — Part II Project Dissertations:
http://www.cl.cam.ac.uk/teaching/projects/pinkbook/node19.html

Examination Questions from Previous Years:
http://www.cl.cam.ac.uk/teaching/exams/pastpapers/

Examiners' Reports (including Comments and Summaries):
http://www.cl.cam.ac.uk/teaching/exams/reports/

Example Examination Question

The following specimen question is taken from a Part II Computer Science Tripos paper in 2002. The heading shows the course on which the question is principally based and the marking scheme for the question is explicitly indicated by showing how the 20 marks are distributed.

Almost every question in Parts IA, IB and II of the Computer Science Tripos follows the pattern implied by this example.

---

8 Advanced Algorithms

(*a*) Explain how to check a large number for primality using a probabilistic method that gives you a bound of the probability of getting an incorrect judgement. [7 marks]

(*b*) Give an asymptotic formula predicting the number of computer operations needed to verify that a number with n bits is prime, supposing that multiplication, division and remaindering are done using $O(n^2)$ methods and that you want to achieve a probability of error bounded by 1 in 2^{60}. You do not need to prove that the algorithm you describe works, but you should nevertheless explain it carefully and completely. [7 marks]

(*c*) The gap between adjacent primes near the integer N is roughly $\log(N)$. Estimate roughly the number of computer operations you would expect to be needed to find a 2000-bit prime that is just slightly larger than some given 2000-bit random number. [6 marks]

Examination Question Assessment — June 2010

Dear Assessor,

Here are some scripts which we should like you to mark, together with the associated solution notes.

Computer Science Tripos Part II 2010

Paper 9 Question 7

NAD – Advanced Graphics

Marking the Scripts

- Please put the total mark on the cover sheet. It should be a whole number in the range 0 to 20.
- Please make it easy to associate the part marks on the script with the marks in the question paper (e.g. "4/6", "1/5" rather than simply "4", "1").
- It is essential to put some indication of marking on every page (including blank ones) to make it clear that no page has been missed (a mark, or a tick, or a line, for example).
- No comments should be written on the scripts.

Returning the Scripts

- During the working day there will often be someone in the Examination Office who will be pleased to accept marked scripts. Alternatively, please return the marked scripts to Chris or Reception. We should appreciate receiving marked scripts as soon as possible and in any case before

12 noon on Thursday 10 June

- It would be greatly appreciated if the returned scripts were sorted by ascending order of candidate number.
- Please return this letter with your scripts.
- You are invited to return your solution notes, updated and annotated as you feel appropriate. These notes will be placed in the course Supervisors' Guide in the Student Administration office unless you request otherwise.

Comments on the Candidates' Performance for the Examiners' Report

- Please supply a few comments by e-mail to `chn2@cam.ac.uk` on how the candidates dealt with this question. These comments will be included in the Examiners' Report and will be placed on the Web.

If you need any guidance please consult an Examiner or Chris.

Thank you,

`John.Daugman, Andrew.Moore, Mike.Gordon, Christine.Northeast`

Part II Dissertations — Processing the Raw Marks

This Appendix gives details of how the raw marks supplied by the assessors are processed. The process is described by referring to an outline of the associated spreadsheet:

Candidate	a_1	a_2	a_3	a_4	a'_1	a'_2	a'_3	a'_4	c	d
Bloggs, A.	32	29	31		55.8	58.5	59.2		57.8	57.8
Mean:	m_1	m_2	m_3	m_4	70	70	70	70	m_c	m_d
St. Dev:	s_1	s_2	s_3	s_4	12	12	12	12	s_c	s_d

Apart from the candidate's name, the first line is taken directly from the 1997 spreadsheet, when there were four assessors because of the very large number of candidates. Other lines are shown schematically. The columns headed a_1, a_2, a_3 and a_4 show the raw marks supplied by four assessors, each marking out of 50. Each assessor read only three-quarters of the dissertations which accounts for the blanks in the columns.

At the foot of these columns, the mean and standard deviation for each assessor are shown as m_i and s_i, where $i = 1, 2, 3, 4$. Typically, m_i is in the mid 30s and s_i is about 7.

The raw marks are then normalised via the formula:

$$a'_{ij} = 70 + (a_{ij} - m_i).12/s_i$$

The normalised marks are shown in the columns headed a'_1, a'_2, a'_3 and a'_4. The purpose of the formula is to force the mean and standard deviation of the normalised marks to be 70 and 12 respectively.

These values ensure that *very* few entries are outside the range 46 to 94 (roughly the range for a typical written paper). Although the formula can lead to the occasional mark being negative or greater than 100 such an occurrence is rare and is discussed later.

The procedure allows different assessors to mark to different means, so those who like to think of 63 (say) as a mid II.1 are free to treat that as a reference mark.

The average of the three normalised marks is then calculated for each candidate and the results are shown in the column headed c (for Calculated mark). For the majority of candidates (including Bloggs) the same mark is shown in the column headed d (for Decided mark).

At their deliberation meeting, the examiners work from a list sorted on the calculated marks rather than a list in alphabetical order. Each candidate is considered carefully, starting with the candidate with the highest calculated mark.

When Bloggs is considered, it would be noticed that the three raw marks and the three normalised marks show a narrow spread so there is no *prima facie* reason for changing the calculated mark.

Nevertheless, Bloggs's dissertation will additionally be compared with other dissertations with neighbouring marks and such consideration may lead to the d mark being higher or lower than the c mark.

When the normalised marks of a candidate show a wide spread or when a candidate's calculated mark is close to the upper or lower extreme, the dissertation will be discussed especially carefully. The outcome is recorded as the d mark which may be several marks above or below the c mark.

The very rare case of a negative normalised mark would, of course, lead to such a discussion, as would a mark greater than 100 though this has never yet happened.

The mean of the c values is shown as m_c which is necessarily 70 but the standard deviation s_c will be somewhat less than 12. If the calculated marks were independent s_c would be roughly $12/\sqrt{3}$ but the marks are *not* independent. If the assessment process is sound, the marks for a given candidate should correlate fairly closely. Accordingly, the value of s_c is a crude measure of goodness. The closer it is to 12 the better.

The mean and standard deviation of the d values are shown as m_d and s_d and these values are usually close to m_c and s_c. For many candidates the d mark is the same as the c mark and, where the marks differ, upward adjustments are typically matched by downward adjustments.

The values shown in the column headed d are the published marks which appear in the final mark book.

The arguments for selecting a mean mark of 70 are both psychological and pragmatic. Viewed as a score for a written paper, 70% is generally regarded as a first-class mark. The same percentage for an individual question in a written paper, 14 out of 20, is generally regarded as respectable but hardly a model answer.

Even weak candidates usually turn in at least one question which merits 14 out of 20 and since most candidates put very considerably more effort into their projects than into any individual course a 70% mark is not unreasonable as a mean.

In the days when the mean published mark was 63 there were frequent complaints from Directors of Studies, Supervisors and candidates themselves, that the marks did not reflect the effort put in or the quality of the work. There have been no such complaints since the mean mark was changed to 70. Given the use of Baxter Averages, this inflation in project marks has no effect on the distribution of classes in the class list.

For some years a non-linear transformation was used to adjust the mean to 70. This sometimes led to difficulties at the extremes. The practice of scaling the marks of all assessors to a mean of 70 at an early stage dispenses with any need for such transformation.

Determining Class Boundaries — Cohort Tracking and Baxter

This Appendix describes the steps involved in converting the final order-of-merit table into a class list. The following outline description is restricted for simplicity to Part Ib but it is readily adapted to Part II:

1. Well in advance of the examination, the Examinations Officer (who manages the computer processing of the marks in Parts Ib and Part II but who is not an examiner) identifies those Part Ib candidates who took Part Ia in the previous year and deems such candidates to form a cohort. To be consistent with the principles of anonymous marking, the identities and Baxter scores of the members of the cohort are hidden from the examiners.

2. Taking the Baxter scores of 5 for a First, 3 for a II.1, 2 for a II.2, 1 for a III and 0 for an unclassed candidate, the Part Ia Baxter Average for the cohort is computed and this value is *not* hidden. It may be *appreciably higher* or *appreciably lower* than the overall Baxter Average in Part Ia depending on whether a predominance of weaker or stronger Part Ia candidates chose *not* to continue into Part Ib. These lost candidates do not contribute to the cohort.

3. An order-of-merit table is prepared. In this, candidates are sorted on the primary key of overall grand total mark (the four written papers less any penalties).

4. Provisional border lines are then drawn such that the distribution of candidates in the cohort is as it was in Part Ia. Of course an individual candidate will not necessarily be in the same class as in Part Ia. The Baxter Average for the cohort is, at this stage, exactly as it was in Part Ia.

5. A decision is made as to whether any candidates at the bottom of the table should not be classed. If there are any, another border line is drawn separating the IIIs from the unclassifieds.

6. Using the provisional border lines, a Baxter score is assigned to all Part Ib candidates (including any who did not take Part Ia in the previous year). The spreadsheet then determines a second Baxter Average, the overall Baxter Average for the examination.

7. The overall *Baxter Average for the examination* may be appreciably above or below the Part Ib *Baxter Average for the cohort* who took Part Ia, depending on whether those who joined the class are stronger or weaker than those in the Part Ia cohort.

8. The examiners must now consider the border lines carefully. They may decide to adjust one or more border lines slightly. This stage cannot be too prescriptive but the candidates in the Part Ia cohort should not, in Part Ib, be distributed significantly differently from the way they were in Part Ia. In consequence, the Part Ib Baxter Average of the cohort should not be significantly lower than it was in Part Ia and it should not be significantly higher.

Part IB Case Study

The following case study is based on the actual data for a particular year. Of 77 Part IB candidates, 65 had taken Part IA in the previous year when they had been distributed as follows: Class I – 12, Class II.1 – 22, Class II.2 – 27, Class III – 4.

This led to a Baxter Average for the cohort of:

$$\frac{12 \times 5 + 22 \times 3 + 27 \times 2 + 4 \times 1}{65} = 2.83$$

Provisional border lines were then drawn such that the members of the cohort were distributed in the same way. Using these border lines for the whole class led to an overall Baxter Average of:

$$\frac{14 \times 5 + 26 \times 3 + 32 \times 2 + 5 \times 1}{77} = 2.82$$

The bottom candidate was declared unclassed so the overall Baxter Average at this stage was:

$$\frac{14 \times 5 + 26 \times 3 + 32 \times 2 + 4 \times 1 + 1 \times 0}{77} = 2.81$$

and the Baxter Average for the cohort was

$$\frac{12 \times 5 + 22 \times 3 + 27 \times 2 + 3 \times 1 + 1 \times 0}{65} = 2.82$$

The Examiners then considered marks at the provisional borderlines in detail and after discussion the partitioning was revised to $14, 28, 30, 4, 1$ which gave a revised overall Baxter Average of:

$$\frac{14 \times 5 + 28 \times 3 + 30 \times 2 + 4 \times 1 + 1 \times 0}{77} = 2.83$$

and a Baxter Average for the cohort of:

$$\frac{12 \times 5 + 24 \times 3 + 25 \times 2 + 3 \times 1 + 1 \times 0}{65} = 2.85$$

The Baxter Average of the cohort thus exceeded its Part IA value, but only slightly.

Part II Case Study

The steps for Part II are essentially the same except that for some cohort the Baxter Averages are known for both Part IA and Part IB.

In another case based on actual data, there were 72 candidates of whom 68 had taken Part IA and Part IB in the two previous years when their Baxter Averages had been 3.00 and 3.13 respectively. In drawing up the provisional border lines in Part II, the examiners used the distribution of the cohort in Part IB $(19, 22, 25, 2)$, leading to a Baxter Average for the cohort of:

$$\frac{19 \times 5 + 22 \times 3 + 25 \times 2 + 2 \times 1}{68} = 3.13$$

Using these border lines for the whole class led to an overall Baxter Average of:

$$\frac{20 \times 5 + 22 \times 3 + 27 \times 2 + 3 \times 1}{72} = 3.10$$

The bottom candidate was declared unclassed so the overall Baxter Average at this stage was:

$$\frac{20 \times 5 + 22 \times 3 + 27 \times 2 + 2 \times 1 + 1 \times 0}{72} = 3.08$$

and the Baxter Average for the cohort was

$$\frac{19 \times 5 + 22 \times 3 + 25 \times 2 + 1 \times 1 + 1 \times 0}{68} = 3.12$$

The Examiners then considered marks at the provisional borderlines in detail and after discussion the partitioning was revised to $21, 21, 27, 2, 1$ which gave a revised overall Baxter Average of:

$$\frac{21 \times 5 + 21 \times 3 + 27 \times 2 + 2 \times 1 + 1 \times 0}{72} = 3.11$$

and a Baxter Average for the cohort of:

$$\frac{20 \times 5 + 21 \times 3 + 25 \times 2 + 1 \times 1 + 1 \times 0}{68} = 3.15$$

The Baxter Average of the cohort thus slightly exceeded its Part Ib value.