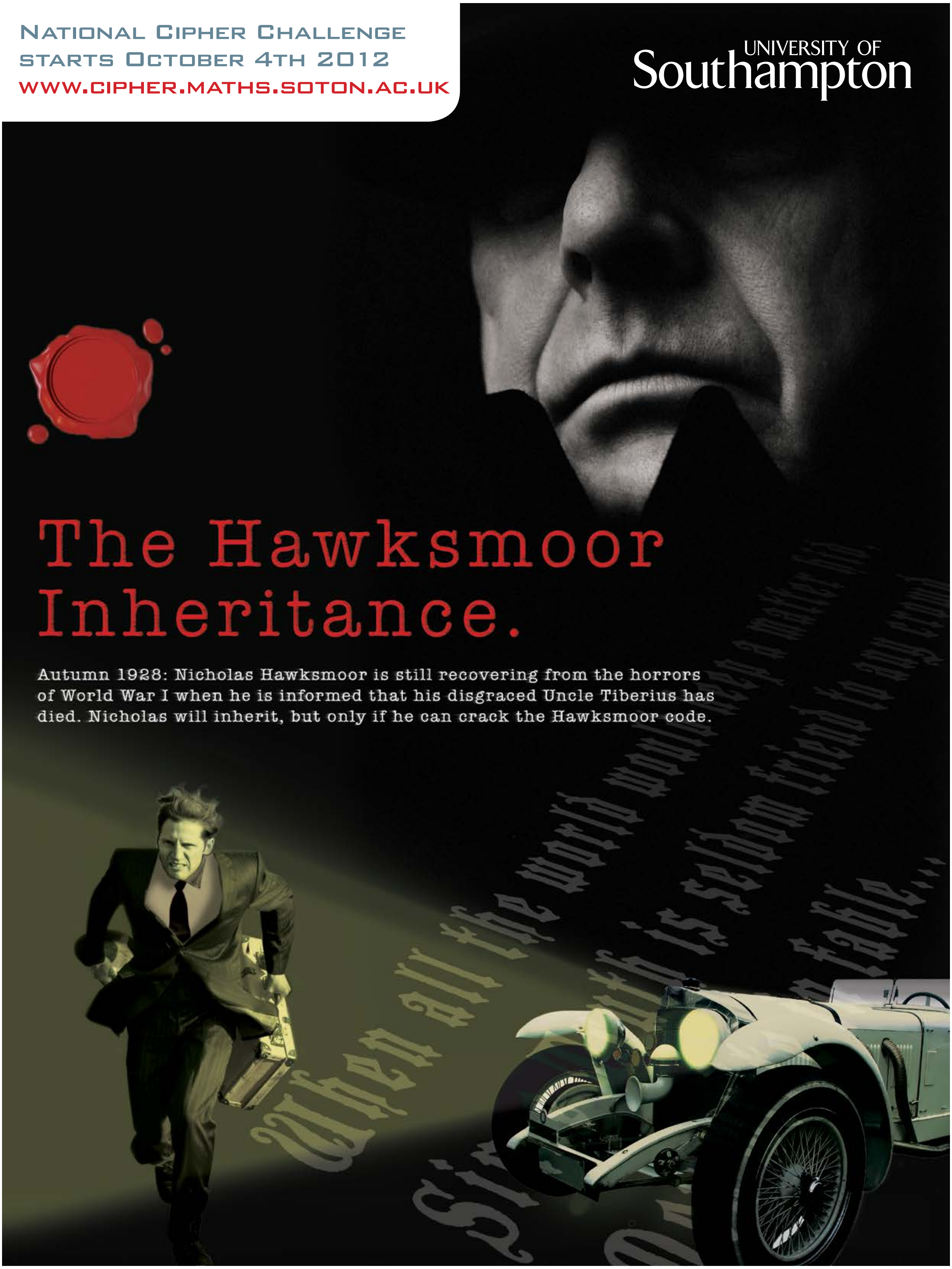
NATIONAL CIPHER CHALLENGE
STARTS OCTOBER 4TH 2012
WWW.CIPHER.MATHS.SOTON.AC.UK



The Hawksmoor Inheritance.

Autumn 1928: Nicholas Hawksmoor is still recovering from the horrors of World War I when he is informed that his disgraced Uncle Tiberius has died. Nicholas will inherit, but only if he can crack the Hawksmoor code.

Presented by the Mathematics Department at the University of Southampton. Competition designed and created by Harry.

TrinityCollegeCambridge · University of Southampton · GCHQ · IBM · NETCRAFT · bcs The Chartered Institute for IT

email:cipher@soton.ac.uk

Teachers’ notes and lesson plans

About the Challenge

Welcome to the National Cipher Challenge, a nationwide, online codebreaking competition introduce the competitors to the thrill of discovery experienced by professional mathematicians and computer scientists in their work.

The competition will run from October 4th 2012 to January 4th 2013 and will open for registration from September 21st. It is a great extension activity (or a fantastic maths club project) which can be tackled by students in teams or on their own.

Based on the fictional adventures of Nicholas Hawksmoor, the story will unfold in a series of short encrypted messages for the participants to crack, using all their skill and cunning. Together they will help Nicholas to decipher the Last Will & Testament of his disgraced Uncle Tiberius in an effort to uncover the secret of Tiberius’s downfall and to defeat his enemies. The lesson plans and notes contained at the end of this Teachers’ Pack, provide a good introduction to the skills needed by a successful code-breaker.

Entrants can take part in teams as small as a solo entrant, or as large as they want, but prizes have to be shared. Each team needs at least one registered account for the team captain and competitors can build teams by designating one of them as a team captain who can then use the Team Builder to invite others to join. That way anyone in the team can use their own account to submit solutions, check for feedback, resubmit and download certificates.

Challenges will be set periodically on the web-site www.cipher.maths.soton.ac.uk in the Challenges section, following this schedule:

Challenge	Publication date 15:15 on	Solution deadline 23:59 on
1	04/10/2012	10/10/2012
2	11/10/2012	17/10/2012
3	18/10/2012	24/10/2012
4	25/10/2012	07/11/2012
5	08/11/2012	21/11/2012
6	22/11/2012	05/11/2012
7	06/12/2012	19/12/2012
8	20/12/2012	04/01/2012

This year the first two rounds are a “warm up” exercise and will not contribute to the overall placings on the Championship leaderboard, which will be published from Round 3 onwards. You should try the early rounds as they are excellent practice and will count for the new Hut Challenge.

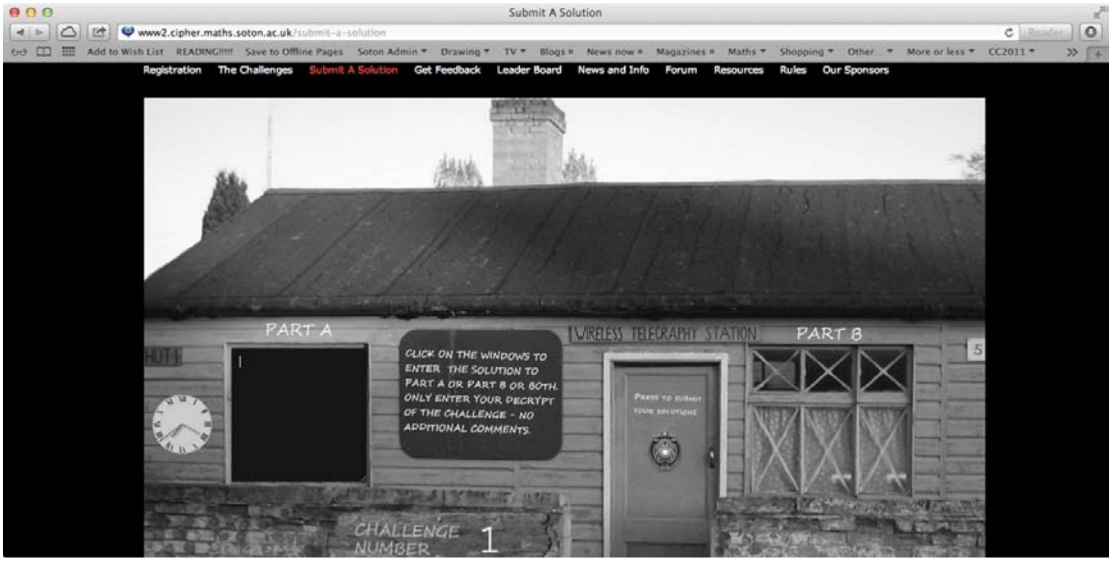
The structure of the competition

Each round of the competition will come in two parts, Part A (the Hut Challenge) and Part B (the Team Challenge). We will say more about Huts and Teams later in this introduction but for now think of them as the "easy" and the "hard" challenges (or the "hard" and "much harder" challenges if you prefer).

Part A challenges will consist of letters, notes etc. sent between our protagonists as they try to unravel the secret of Uncle Tiberius's will. You can expect these messages to be fairly lightly encrypted, at least at first, although in the latter stages of the competition security will be tightened and you will find the Part A ciphers harder to crack.

Part B consists of sections of Uncle Tiberius's Last Will and Testament, and these are encrypted in various ways. Again they will start simply but as Tiberius reveals more about his past the encryption in these challenges will become increasingly tricky. You may find that learning to use a spreadsheet or even to programme will be of particular value in tackling the later stages of the part B competition.

Submitting your solutions



You can submit solutions to either part A or part B (or to both parts together), at any time during a round, and if you need to resubmit you can use the same form. Just paste your entry as text in the appropriate box on the form. Click on the windows to reveal the boxes – the doorbell acts as a submission button and the clock tells you what time we think it is. It doesn't matter how you format your answer – with or without punctuation and spaces and whether or not you use capital letters, however you must only type or paste in the exact text of a decrypt of the message. So the texts "Tiberius, for no apparent reason, seems to have been in trouble with the authorities" could equally be submitted as "tiberiusfornoapparentreasonseemstohavebeenintroublewiththeauthorities".

It is a good idea to use a simple text editor to type up your solution (rather than something like Word) as the spell checker sometimes tries to change what you are typing and the "mistake" might be deliberate. Don't try to correct any errors you think we have made, always type in an exact decryption of the text.

Don't try to tell us what cipher we used, or to ask us a question, or to say how you solved the cipher in the entry form, we don't read it and it will be marked as an error in the solution. If you need to get hold of us you can post a message on the forum or send us an email at cipher@soton.ac.uk.

Scoring

Each of the two challenges in a round (part A and part B) are scored for accuracy in the same way. We strip out all the non-ascii characters, spaces and punctuation from your solution, convert it to lower case and compare that string of letters with our solution, which we have treated the same way. The more similar they are the higher the score you will get, and if they are identical you will score 100% for that challenge. If you spot a mistake in your answer you can submit again - we only ever take your most accurate answer into account and accuracy beats speed in every case, though speed is also important in the part B competition. In part B we look at all your submissions for the round and find those with the highest mark. We then take the first one of those that you submitted and award you points depending on how quickly you submitted it, according to a schedule which is published with each challenge. There are no speed points for part A, only for part B. You can find out your scores for each round in the feedback section of the site, and we will publish a leader board for each round. The first two rounds are a warmup so the points will not count for the overall leader boards but from round 3 we will publish a Championship leader board based on your total points from then in each of the competitions and we will also publish a score board for the points accumulated by each of the four Huts (see below).

Getting help

We offer online feedback on submissions during each round to help you if you make mistakes. The feedback is delayed so you will lose points if you rely on it rather than trying to correct your own errors quickly, but it can be useful if you are on the right track (and speed doesn't matter for part A challenges which are only scored for accuracy). The feedback consists of a score for accuracy, together with a copy of your submission with the first error highlighted. So for example, if you tried to submit a solution to the text above which missed the word "in" as below our feedback system would spot that and highlight the letter t as below. The letter t is only wrong marked here because the solution checker expected to see a letter i.

"tiberiusfornoapparentreasonseemstohavebeentroublewiththeauthoritiesforsometimebythen".

At the end of each round we will publish the official decrypts of part A and part B, which sometimes contain hints at how to tackle the next round. Participants often get stuck on a Challenge but, as in real life, sometimes a good night's rest is all you need. Other times you might need more practical help and can turn to the website for clues, either hidden in earlier rounds of the competition, or posted (by us) as comments on the forum. We ask you not to post hints of your own without checking them with us first as this will spoil the Challenge for others. Anyone posting solutions or links to solutions on our site or elsewhere may be barred from the site and disqualified from the competition - we do search for them and do find them!

Huts, Teams and solo entries

Every team needs at least one member (the team captain) and if you are entering on your own that is all you need. If several of you want to form a team to compete, whether you are in the same school or not, you can do that and again you only need one account belonging to the team captain who will then submit your entries. You can list the

members of the team as you want them to appear on the leader boards and on the certificates in the team captain's account.
If you want to all be able to post entries for the team, to read the feedback and to download the certificates then you should each sign up for an account and use the Team Builder tab in the team captain's account profile to issue invites to the other team members to join. The team members will then be able to accept the invitation by going to the same place in their own account profile. Individuals can leave a team at any time, but cannot be thrown out by anyone else, even the captain, so choose your teammates wisely. The captain can only leave a team once everyone else has. Team submissions (from anyone) belong to the team; you can't claim them as yours if you leave.

At the end of the second round of the competition (which will finish at midnight on Wednesday 17th October) we will assign each of the teams and solo entrants taking part in the Part A challenges to one of four Huts: Hut 3, Hut 4, Hut 6 and Hut 8, each named after one of the the famous Bletchley code breaking or sig-int divisions. The points won by a team or individual in each part A challenge will also be awarded to their Hut and the Huts will compete for National glory throughout the remainder of the competition. Total Hut points will be displayed on the website.

Prizes

The GCHQ prize of £1,000 will be awarded to the top codebreaker or codebreaking team in the country as measured by performance in the part B competition. IBM provides a second prize of £800 and Trinity College, Cambridge awards £700 to the third place on the leader board. The position on the leader board is determined by combining an accuracy score with a time score, and the full details of how this is done can be found on the rules section of the website.

www.cipher.maths.soton.ac.uk/rules

The University of Southampton awards a number of smaller prizes to teams taking part in each of the 8 part A Challenges and this year the top 50 individuals in teams on the part B Championship Leader board will win a laptop rucksack from our sponsor Netcraft, together with a Raspberry Pi computer donated by GCHQ.

The Prizegiving

Our sponsor BCS – The Chartered Institute for IT will host a prizegiving ceremony for selected participants at Bletchley Park on 12th April 2013. Some tickets will be available by lottery and you can apply for them online at
http://www2.cipher.maths.soton.ac.uk/index.php?option=com_cipher&controller=tickets

The afternoon will include a tour of the museum and tickets are always in high demand.

What do you need to do?

To take part you will need to register for the competition on our registration page:

www.cipher.maths.soton.ac.uk/registration

This will be open from September 21st, and you will need to provide the following information:

- Username: This will be the name you use to log on to the site to post comments, and also to submit your entries, check feedback and to print your certificate. Choose something memorable. You can share this info among the team, but do not let anyone else have it as they can use it to pose as you!
- Password: Again this is for logging on. Choose it carefully and keep it secret.
- Email address: This will be used to confirm your registration so it must be an active account you can check. We can only associate one email account to each team, so it might be an idea to set up an account (with Gmail or some other free provider) that the whole team can use for the competition. If we need to contact you this is how we will do it so add the account cipher@soton.ac.uk to the account address book to avoid sending our emails to your junk mail bin, make sure the account is not too full, and check it regularly.
- Team name: This will be used on the downloadable certificates and the leader boards to identify you. Try to choose something interesting but tasteful. (We reserve the right to refuse offensive names!)
- Team members for the web: Enter your names as you wish them to appear on the leader board. If you wish to remain anonymous online then leave this blank or use a nickname that does not identify you.
- Team members for the certificates: These are your names as you wish them to appear on the downloadable certificates and will not be used publically without your permission.
- Number in team: While we could be clever and write a script to try and compute the number in the team from the info about team member names this is not very accurate and far too much work, so please tell us how many of you are in the team so we get some idea of how many are taking part!
- Teacher contact: Give the name of a teacher we can write to if we need to check anything. You should get their permission first! We don't usually do this unless you win a prize. If you are home schooled give us a parent or carer's name here and write home schooled in the school name field. You will still need to give us contact details in the address fields below.
- School name: Again we will use this on the certificates and the leader boards so spell it right, and if it is a common name (like King Edward's) add a location (King Edward's, Southampton) to identify it.
- School address: If you win a prize this is how we will get it to you, so it is important that you get this right. We collect the postcode separately so we can easily map the participants to give us some idea of the spread of entries.
- School postcode: See above.

Frequently Asked Questions

How many can enter?

Teams of any size and composition may enter, and a school can enter as many teams as it wishes. Teams can be run from one or several individual accounts (see above) and inter-school teams are also allowed

Does everyone in a team need their own account?

No, but everyone can have one! The team is principally associated with one account holder (the team captain), but team captains can invite others to join the team at any point including when it is set up. You don't need to have an individual account to be associated with a team, as you can be listed as a member on the team captain's account. You do however need an account if you want to be able to post comments to the forum or to post solutions for the team. You also need one to log in to see feedback and get certificates.

What does it cost to take part?

At the moment we are lucky to have several generous sponsors, and the costs of the competition are covered by the School of Mathematics at the University of Southampton so there is no charge to take part.

When does the competition start?

Registration opens on September 21st. The first part of the competition will be published at 3.15 on Thursday October 4th. There is no need to rush to download it as you have one or two days in which to submit to achieve full speed marks. Often the website is overloaded for the first half an hour or so, and it probably pays to wait out the rush. The part B prizes will be allocated based on performance in rounds 3-8 so if you miss the first couple of challenges it won't matter too much.

When will the Challenges be published?

At or shortly after 3.15pm on each of the following Thursdays:
4th, 11th, 18th , 25th October; 8th, 22nd November; 6th, 20th December. These dates have been chosen to avoid the most common half term dates.

How long do we have to complete each challenge?

The deadline for each Challenge is 11.59pm on the day before the next Challenge is published, The number of points you score in part B depends on how early you submit your best attempt.

What age group is it aimed at?

Principally this is an extension activity for older pupils, but the early stages of the competition are aimed at a wide audience and there is something for everyone. We have had bright year 6 pupils do well on early rounds and this prepares them for further achievement in future competitions. Many pupils return year after year as they try to improve on the number of stages they can successfully complete, and the staged certificates give everyone an incentive to keep going as long as possible. There are also some small prizes awarded at random to a few participants each week as an added incentive. Finally we are introducing the national Hut Challenge this year, in which participants are assigned to national teams, Hut 3, Hut 4, Hut 6 and Hut 8 to compete for glory (not gold!). Even the smallest contribution can make a difference to the outcome of that competition so it is worth trying even if you can't complete the challenge.

Is this appropriate for a Math Club activity?

Certainly. The entire math club could enter as a team or you could divide up into smaller groups and use the math club meetings to discuss techniques and strategies.

Can pupils enter on their own?

Yes, we get many solo entrants and teachers do not have to be involved, but we do ask for the name of a teacher contact for prize administration.

Do team members have to all come from the same school?

No, in the past we have had several teams made up of members of different schools and colleges, and this is perfectly acceptable. However we do ask for the name of at least one of the schools and a teacher contact for prize administration, and, for now at least, the team captain's school will get all the credit!

What are the Huts and why have you introduced them?

Good question. This year we are trying something new. At the start of round 3 each team/solo entrant to part A of the Challenge will be assigned to one of four national teams, Hut 3, Hut 4, Hut 6 and Hut 8. The password to unlock your Hut membership will be published by a fan of the National Cipher Challenge at that time. We can't say more now, but all will become apparent later! The points awarded to the members of each Hut for their attempt at the part A challenges throughout the competition will be totted up to give a Hut total and these will be published on the website throughout the competition. Even if you don't get get full marks on a challenge your efforts will still play an important part in the score for the Hut. No prizes, but plenty of glory for the winning Hut!

Where can I find out more?

The rules page and the news and info pages give more information, and it is worth taking part in the forums, which we will monitor:

www.cipher.maths.soton.ac.uk/rules
http://www.cipher.maths.soton.ac.uk/news-and-info
http://www.cipher.maths.soton.ac.uk/discussion-boards

Now the heavy stuff:

Rules

1. The competition is only open to persons who are in full time school-level education in the United Kingdom (Home-schoolers also qualify as do those attending schools in the Channel Islands and the Isle of Man.).
2. The competition is only open to persons aged 18 or under on 31 August 2011.
3. Entries may be received from individuals or from teams. The teams may be of any size, but we reserve the right to restrict the number of team members listed on the honours board. If you are in a big team you should make sure we have a team name as well.
4. Teams must nominate a captain who we may contact via email.
5. The schedule of messages to be deciphered is given here, the list of prizes is given below.
6. Each challenge consists of two parts, part A and part B. You may submit solutions to either or both parts of the current challenge on the entry form.
7. For each of the challenges 1 to 8 there will be a small cash prize (this replaces the previously advertised book prize) awarded to 8 teams/competitors chosen at random from those who submit a correct entry to part A.
8. For each of the challenges 1 to 8 for which you submit an entry for part B you will be awarded a score, based on the accuracy of your best submission for that challenge and the order in which we receive the submissions. The scores will be used to determine the winners of the main Prizes, who will be chosen from among those achieving the highest total scores.
9. A solution will only be deemed to be correct if, disregarding the punctuation [and spacing], the deciphered plaintext (only involving the Roman characters A to Z [UPPER or lower case is fine]) is letter perfect as compared to our master solution.
10. A submission will only be deemed to be valid if it is submitted on the entry form and all the instructions on the entry form are adhered to.
11. The Challenge Committee may publish clues on the competition web-site if it considers it appropriate to do so.
12. If a correct solution of a challenge is not received before the deadline given on the schedule the Challenge Committee will have the discretion to not award the prize or award some or all of it to the entrant or entrants whom it judges to represent the best solution or solutions.
13. The competition will be judged by the Prize Committee, whose decision will be final in all matters regarding the competition including the award of prizes.
14. In order to qualify for any of the prizes all entrants, whether solo or part of a team agree to their school names being used in publicity associated with the competition including publication on the competition web-site. We will publish team member names only with your permission. When registering you will be asked to list the team member names, or appropriate nicknames for them that you are happy for us to publish on the website and in publicity associated with the competition.
15. In submitting an entry solo entrants vouch that it is solely their own work and teams warrant that it is solely their own collective work.
16. Entrants who do not abide by the rules will be disqualified from the competition and will not qualify for any of the prizes.
17. In submitting an entry to the competition, all entrants, be they individuals or members of a team agree to be bound by all the rules of the competition.
18. Winners and their schools will be notified as soon as possible after the solution deadline for each message. We will inform you on the website in the News and Info

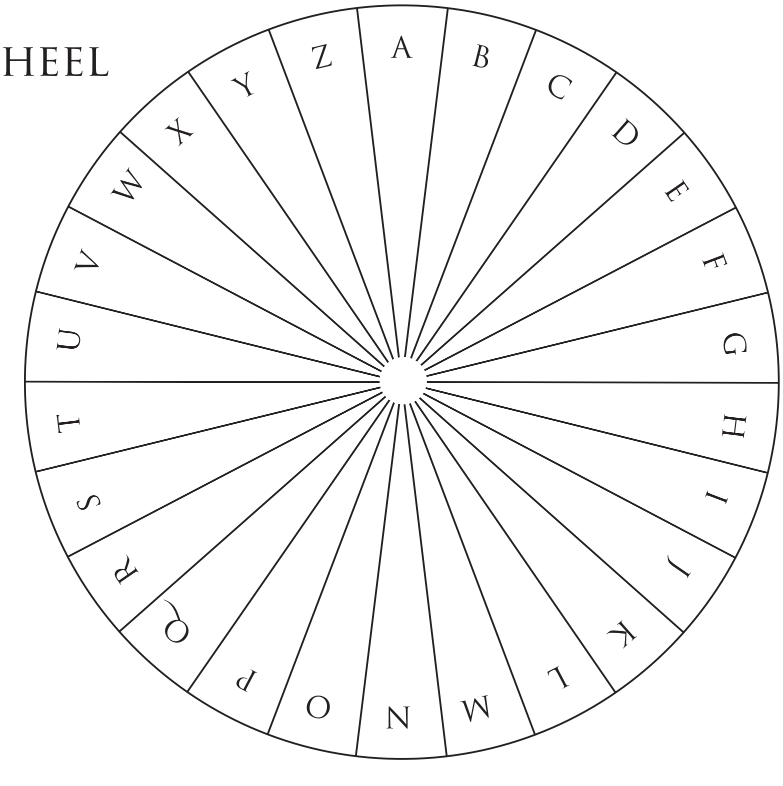
section and by email using the email address you provide. Please ensure that our messages to you are not filtered by your spam filters by adding our email address cipher@soton.ac.uk to your address book.
19. The organisers reserve the right to change any aspect of the competition at short notice and to split prizes where it is deemed appropriate. Such changes will be announced on the competition web-site as soon as practical.
20. Rapid fire multiple submissions put an unreasonable load on the servers and make it difficult for others to submit. Anyone who makes an unreasonable number of submissions will be open to disqualification. In the first instance this will mean that anyone who submits more than 20 times in 10 minutes will be disqualified.
21. Anyone posting a solution in any public forum before the deadline for the given Challenge will be disqualified and barred from the forums. Please do not publish hints without checking with us first as this may spoil the competition for others.

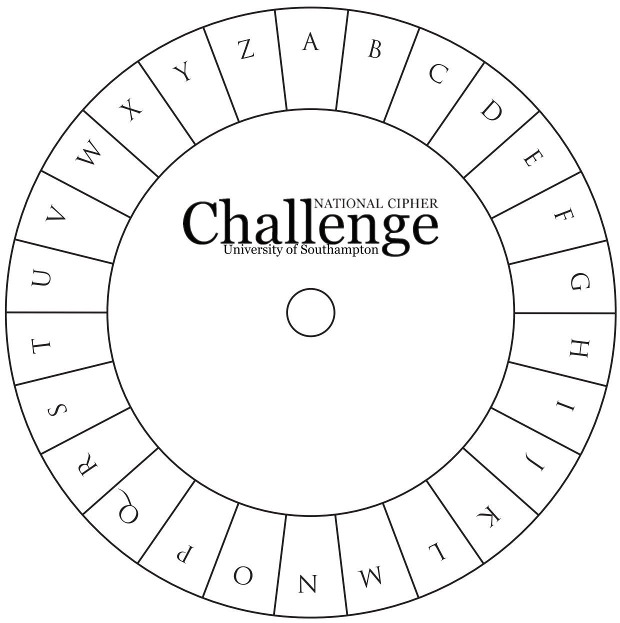
Terms and conditions

1. In accepting these Terms & Conditions you are agreeing to abide by the rules of the competition as printed on the competition website.
2. You are also agreeing that the challenge organisers are permitted to publish on the challenge website the content of those fields that are marked as visible on the registration page together with the contents of the field entitled "Team Members for Web".
3. You are further agreeing that we may use your contact details to contact you in relation to your entry in the Cipher Challenge competition and that we may contact your "Teacher Contact" if this is felt necessary.
4. The software we use to generate our webpages sometimes places small amounts of information on your device, for example, computer or mobile phone. These include small files known as cookies. They cannot be used to identify you personally but are used to allow you to login to our site and take part in the Cipher Challenges and read/post information on the forums. Under EU law, we need your consent to store these cookies on your device. Consequently, you are further agreeing to allow our software to store these cookies on your device.
5. We agree that we may not use any of the information obtained from the registration page or from your entry to the competition for any other purpose except administration of the competition and information about future Challenges unless we have your explicit permission to do so.

We also agree not to divulge any information about you except where specified in item (2) to any third-party.

Cipher Wheel





IN ASSOCIATION WITH
BCS
GCHQ
TrinityCollegeCambridge
IBM
NETCRAFT

On substitution ciphers

Graham A. Niblo
University of Southampton National Cipher Challenge 2008

These notes form a brief introduction to substitution ciphers, to accompany the lesson plans provided with the University of Southampton National Cipher Challenge, 2005. We would like to thank Hugh Evans of Sholing City Technology College for his assistance in the design of these teaching materials.

Caesar shift ciphers

The easiest method of enciphering a text message is to replace each character by another using a fixed rule, so for example every letter a may be replaced by D, and every letter b by the letter E and so on.

Applying this rule to the previous paragraph produces the text

```
WKH HDVLHVW PHWKRG RI HQFLSKHULQJ D WHAW PHVVDJH LV WR
UHSODFH HDFK FKDUDFWHU EB DQRWKHU XVLQJ D ILAHG UXOH, VR
IRU HADPSOH HYHUB OHWWHU D PDB EH UHSODFHG EB G, DQG
HYHUB OHWWHU E EB WKH OHWWHU H DQG VR RQ.
```

(Note the convention in these notes that ciphertext is written in capital letters, while plaintext is usually lowercase.)

Such a cipher is known as a shift cipher since the letters of the alphabet are shifted round by a fixed amount, and as a Caesar shift since such ciphers were used by Julius Caesar. To decode a Caesar shift it is enough to work out the amount of shift, which can be done, for example, by discovering which character has replaced the letter e. In the example above we might guess that the three-letter word starting the sentence is the and therefore that the letter e has been replaced by H. A quick check shows that the Caesar shift by 3 does indeed encode the word the as WKH, and it is easy to complete the decryption.

In fact there are only 26 Caesar shift ciphers (and one of them does nothing to the text) so it is not too hard to decipher the text by brute force. We can try each of the shifts in turn on the first word of the cipher text until we discover the correct shift. This process can be simplified by using a cipher wheel, a simple mechanical device that allows one to generate each of the Caesar shift ciphers, and to encode or decode messages using it. At the front of this pack you will find a sheet, which can be photocopied onto thin card in order to make a cipher wheel. Cut out the two discs, and fasten through their centres with a paper fastener to make the wheel. Use the convention that you read plaintext on the outer rim of the wheel and cipher text from the smaller wheel.

Keyword substitution ciphers

To increase the difficulty of deciphering the text we need a richer family of ciphers. A good example is furnished by the keyword cipher. In this we design an encryption table by choosing a keyword or phrase, which is used to jumble the alphabet as follows:

Write down the phrase, with no spaces between the letters, and omitting any repeated character. So if the phrase is "The Simpsons" we write down THESIMPON. Now we continue to go round the alphabet until every letter appears exactly once, and write the list under the standard alphabet:

```
abcdefghijklmnopqrstuvwxyz
THESIMPONQRUVWXYZABCDFGJKL
```

Of course if the key phrase is carefully chosen (for example "The quick brown fox jumps over the lazy dog") we may not need to complete the list as above, but such a choice is not necessary. The number of such ciphers is 26!, or approximately 10^{27}, and brute force cannot be used to attack the problem. However an attack is possible.

Consider the text

```
VEP HYXHLVHTP MO AWFJYFLT H RFNEPS HJNEHAPV FL VEFU ZHC
FU VEHV FV FU PHUC VM KPKMSFUP VEP IPCZMSY MS IPCNESHUP,
HLY EPLRP VEP RFNEPS HJNEHAPV. VEFU FU FKNMSVHLV, APRHWUP
FO VEP UPLYPS EHU VM IPPN VEP RFNEPS HJNEHAPV ML H NFPRP
MO NHNPS, VEP PLPKC RHL RHNVWSP VEP NHNPS, YFURMXPS VEP
IPC, HLY SPHY HLC RMKKWLFRHVFMLU VEHV EHXP APPL PLRSCNVPY
ZFVE FV. EMZPXPS FO VEP IPC RHL AP RMKKFVVPY VM KPKMSC FV
FU JPUU JFIPJC VM OHJJ FLVM PLPKC EHLYU.
```

As before we notice that the first word has three letters and, since it occurs several times, may well be the word the. This gives a strong hint that the letter e is enciphered as the letter P in the keyphrase cipher. Of course other three letter words are possible, *e.g.,* and or but. Nonetheless a quick check shows us that the letter P is the most common letter in the enciphered text, just as e is the most common letter in English so it is reasonable to assume that the correct decryption translates P to e. This also suggests that V stands for t and E for h, allowing us to begin to decipher the text. We will use the convention that uppercase letters denote enciphered letters and lowercase denotes plaintext characters:

```
the HYXHLtHTe MO AWFJYFLT H RFNheS HJNhHAet FL thFU ZHC
FU thHt Ft FU eHUC tM KeKMSFUe the IeCZMSY MS IeCNhSHUe,
HLY heLRe the RFNheS HJNhHAet. thFU FU FKNMStHLt, AeRHWUe
FO the UeLYeS hHU tM IeeN the RFNheS HJNhHAet ML H NFeRe
MO NHNeS, the eLeKC RHL RHNtWSe the NHNeS, YFURMXeS the
IeC, HLY SeHY HLC RMKKWLFRHtFMLU thHt hHXe AeeL eLRSCNteY
```

```
ZFth Ft. hMZeXeS FO the IeC RHL Ae RMKKFtteY tM KeKMSC Ft
FU JeUU JFIeJC tM OHJJ FLtM eLeKC hHLYU.
```

Reading carefully we see the single letter word H, and the four letter word thHt in the first line, and guess that H enciphers the letter a. Making that replacement we get:

```
the aYXaLtaTe MO AWFJYFLT a RFNheS aJNhaAet FL thFU ZaC
FU that Ft FU eaUC tM KeKMSFUe the IeCZMSY MS IeCNhSaUe,
aLY heLRe the RFNheS aJNhaAet. thFU FU FKNMStaLt, AeRaWUe
FO the UeLYeS haU tM IeeN the RFNheS aJNhaAet ML a NFeRe
MO NaNeS, the eLeKC RaL RaNtWSe the NaNeS, YFURMXeS the
IeC, aLY SeaY aLC RMKKWLFRatFMLU that haXe AeeL eLRSCNteY
ZFth Ft. hMZeXeS FO the IeC RaL Ae RMKKFtteY tM KeKMSC Ft
FU JeUU JFIeJC tM OaJJ FLtM eLeKC haLYU.
```

Now the two 2 letter words Ft FU are probably “it is” meaning that F enciphers “i” and U enciphers “s”. Hence we get:

```
the aYXaLtaTe MO AWiJYiLT a RiNheS aJNhaAet iL this ZaC
is that it is easC tM KeKMSise the IeCZMSY MS IeCNhSase,
aLY heLRe the RiNheS aJNhaAet. this is iKNMStaLt, AeRaWse
iO the seLYeS has tM IeeN the RiNheS aJNhaAet ML a NieRe
MO NaNeS, the eLeKC RaL RaNtWSe teh NaNeS, YisRMXeS the
IeC, aLY SeaY aLC RMKKWLiRatiMLs that haXe AeeL eLRSCNteY
Zith it. hMZeXeS iO the IeC RaL Ae RMKKitteY tM KeKMSC it
is Jess JiIeJC tM OaJJ iLtM eLeKC haLYs.
```

Continuing with appropriate guesses (haXe = have, easC = easy and so on) we decipher the text to get the following extract from Simon Singh’s excellent history of codes and ciphers, *The Code Book*:

```
“The advantage of building a cipher alphabet in this way
is that it is easy to memorise the keyword or keyphrase,
and hence the cipher alphabet. This is important, because
if the sender has to keep the cipher alphabet on a piece
of paper, the enemy can capture the paper, discover the
key, and read any communications that have been encrypted
with it. However if the key can be committed to memory it
is less likely to fall into enemy hands."
```

Frequency analysis

A more methodical attack is frequency analysis. One counts the number of occurrences of each character in the cipher text and compares it with an expected frequency for the standard English alphabet. In the cipher text above a character count gives us the following table of occurrences:

a	b	c	d	e	f	g	h	i	j	k	l	m
7	0	12	0	26	3	0	32	6	9	11	20	18
n	o	p	q	r	s	t	u	v	w	x	y	z
16	5	55	0	14	17	2	17	35	4	4	11	4

Compare this to a table of expected frequencies, taken from Simon Singh's "The Code Book":

a	b	c	d	e	f	g	h	i	j	k	l	m
8.2	1.5	2.8	4.3	12.7	2.2	2.0	6.1	7.0	0.2	0.8	4.0	2.4
n	o	p	q	r	s	t	u	v	w	x	y	z
6.7	7.5	1.9	0.1	6.0	6.3	9.1	2.8	1.0	2.4	0.2	2.0	0.1

Using this and information about common one, two and three letter words we have enough to begin to tackle the cipher.

Disguising the word structure

The chink in the armour of our ciphers so far has been the preservation of word structure. This allows one to spot common words. In order to avoid such weakness cryptographers usually remove punctuation and block the characters together in groups of four or five, so our previous cipher text looks like

```
VEPHY XHLVH TPMOA WFJYF LTHRF NEPSH JNEHA PVFLV EFUZH
CFUVE HVFVF UPHUC VMKPK MSFUP VEPIP CZMSY MSIPC NESHU
PHLYE PLRPV EPRFN EPSHJ NEHAP VVEFU FUFKN MSVHL VAPRH
WUPFO VEPUP LYPSE HUVMI PPNVE PRFNE PSHJN EHAPV MLHNF
PRPMO NHNPS VEPPL PKCRH LRHNV WSPVE PNHNP SYFUR MXPSV
EPIPC HLYSP HYHLC RMKKW LFRHV FMLUV EHVEH XPAPP LPLRS
CNVPY ZFVEF VEMZP XPSFO VEPIP CRHLA PRMKK FVVPY VMKPK
MSCFV FUJPU UJFIP JCVMO HJJFL VMPLP KCEHL YU
```

Usually the length of the text groups doesn't matter, however, in analysing a Vigenère cipher (see below) a carelessly chosen block length may make the length of the keyword more apparent, since it can reveal the repetitions more easily.

To attack cipher text that has been grouped in this way we have to work with letters not words. To do so we use the frequency analysis described above, together with a little judgement (or luck!). The process can be hard, but wars have been won or lost on the back of it, and so have fortunes.

“It was hard going, but Jericho didn’t mind. He was taking action, that was the point. It was the same as code-breaking. However hopeless the situation, the rule was always to do *something*. No cryptogram, Alan Turing used to say, was ever solved by simply staring at it.”
From *Enigma*, by Robert Harris.

Affine shift ciphers

Despite the advantages for an agent in using keyword substitution ciphers most modern ciphers are automated and rely on a mathematical encryption algorithm. Indeed the Caesar shift cipher can be viewed as just such a cipher:

We start by encoding each letter by its numerical position in the alphabet:

a	b	c	d	e	f	g	h	i	j	k	l	m
1	2	3	4	5	6	7	8	9	10	11	12	13
n	o	p	q	r	s	t	u	v	w	x	y	z
14	15	16	17	18	19	20	21	22	23	24	25	26

Next we shift the alphabet by adding 3 to each position:

a	b	c	d	e	f	g	h	i	j	k	l	m
4	5	6	7	8	9	10	11	12	13	14	15	16
n	o	p	q	r	s	t	u	v	w	x	y	z
17	18	19	20	21	22	23	24	25	26	1	2	3

Of course $24+3 = 27 \neq 1$, but here we are carrying out modular arithmetic, familiar as clock arithmetic, so that when we reach 26 we continue from 1.

Finally we replace the numbers with the letters they stand for:

a	b	c	d	e	f	g	h	i	j	k	l	m
d	e	f	g	h	i	j	k	l	m	n	o	p
n	o	p	q	r	s	t	u	v	w	x	y	z
q	r	s	t	u	v	w	x	y	z	a	b	c

This recovers the cipher table constructed in lesson plan 1 for the Caesar shift by 3.

There is a convenient shorthand for the Caesar shift by *n*, given by $x \rightarrow x+n$. It is confusing since here we are using *x* to stand for the position of a letter, and *n* to stand for the shift amount, *i.e.*, *x* and *n* are each one of the values 1 … 26. It is clear that since the shift is defined by the integer *n* there are only 26 Caesar shift ciphers.

There is a bigger class of shift ciphers which can be written in these terms known as the affine shift ciphers, and they exploit the fact that we can multiply as well as add integers in modular arithmetic. It is slightly complicated to set up formally but rather easy to do in practice so we will work through an example.

The affine shift $x \rightarrow 3x+5$

We start as before with the position table, but this time instead of replacing a position *x* with the number $x+3$ we will replace it by the number $3x+5$, where this number is interpreted appropriately. So for example $2 \rightarrow 3.2+5 = 11$, while $8 \rightarrow 3.8+5 = 29$ which is interpreted as 3 (29=26+3). Whenever the result of the computation is larger than 26 we keep subtracting 26 until it becomes smaller. More formally we compute $3x+5$ and then take the remainder after division by 26. This yields the table:

a	b	c	d	e	f	g	h	i	j	k	l	m
8	11	14	17	20	23	26	3	6	9	12	15	18
n	o	p	q	r	s	t	u	v	w	x	y	z
21	24	1	4	7	10	13	16	19	22	25	2	5

And from this we recover the encryption table as given on the handout for lesson 3:

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
H	K	N	Q	T	W	Z	C	F	I	L	O	R	U	X	A	D	G	J	M	P	S	V	Y	B	E

The affine shift ciphers can also be written in a shorthand form $x \rightarrow ax+b$ and the Caesar shift ciphers are special cases of the affine shift ciphers with $a=1$.

Now notice that in both the Caesar shift $x \rightarrow x+3$ and the affine shift $x \rightarrow 3x+5$ the letter y is enciphered as B, since $25+3 = 28 = 26+2$, and $3.25+5 = 80 = 3.26+2$. It follows that two different affine shift ciphers can encrypt a letter in the same way, so it is no longer sufficient to discover the letter substituting for e in order to decipher the message. Since there are two degrees of freedom in our choice of cipher we might hope that deciphering two letters is sufficient, and it is, since, if we know two values of the expression $ax+b$ we can solve the two corresponding simultaneous equations to find a and b.

We may be more familiar with this exercise when solving the equations over the real numbers, but the same method works for modular arithmetic, with the caveat that in general we cannot divide. This caveat has an interpretation in cryptography. In order for the rule $x \rightarrow ax+b$ to define a cipher it had better be the case that each of the numbers 1 … 26 appears exactly once in the list of numbers $ax+b$ as x ranges from 1 to 26. If we choose a carelessly (so that we can't divide by a mod 26) this might not be the case.

For example the rule $x \rightarrow 2x$ tries to encipher both m and z as Z, since 2.13=26 and 2.26 = 52 both of which are equal to 26 modulo 26. Such an encryption cannot easily be deciphered since the recipient of the message is unable to determine whether the sender intended Z to be read as m or z.

From a mathematician's point of view the enciphering rule defines a function from the alphabet to itself, and this needs an inverse if the cipher is to be decipherable in a deterministic way. In other words the number theory function $x \rightarrow ax+b$ needs to have an inverse in mod 26 arithmetic. It is a fact from elementary number theory that it will have such an inverse if and only if a is coprime to 26, that is, their only common divisor is 1.

There are 12 numbers less than 26 and coprime to it (those odd numbers not divisible by 13) so we have 12 possible choices of the number a, and 26 choices for the number b, yielding 312 affine shift ciphers. This makes a brute force attack, without frequency analysis, less practical than the much simpler situation for Caesar shift ciphers.

Polyalphabetic ciphers

The main weakness allowing us to tackle a substitution cipher is the irregularity in the distribution of letters in English text. Other languages demonstrate similar (though language specific) irregularities and you can find frequency tables for them on the web.

In order to remove this weakness from a cipher it is necessary to disguise the frequencies of letters in the plaintext and the easiest way to do this is by using a polyalphabetic cipher. In such a cipher each plaintext letter may be encoded in more than one way so that, for example, the letter e may be enciphered as both X and G within the ciphertext. One problem with this approach is that if X and G both encode for e we don't have enough letters left to encode the other 25 letters. One elegant solution to this problem is the famous French cipher known as the Vigenère cipher.

In a Vigenère cipher ANY letter might be encoded by any other; a given Vigenère cipher uses a subset of the 26 possible Caesar shift ciphers. Of course for a genuine recipient to have any hope of deciphering the message there has to be a way to determine for each cipher character which of the shifts has been used. The answer to this tricky problem is to choose a sequence of them known to both parties but to no-one else.

So the two parties might agree to use shifts of 22, 9, 7, 5, 14, 5 18, and 5 in that order and to continue repeating the pattern for the entire text: 22, 9, 7, 5, 14, 5 18, 5, 22, 9, 7, 5, 14, 5 18, 5, 22, 9 *etc.*.

In order to decode the cipher text the recipient shifts the first cipher character back by 22, the second back by 9 and so on to recover the cipher text. Of course the question remains how one can memorise the correct sequence, but here we borrow an idea from the keyword cipher. The shift numbers 1, …, 26 are taken to stand for the alphabet a, …, z, and then the pattern 22, 9, 7, 5, 14, 5 18, 5 spells the word vigenere.

To set up a Vigenère cipher the two parties agree in advance to use the shift pattern encoded by some agreed keyword or phrase; in our previous Golden Jubilee Cipher challenge we used a Vigenère cipher based on the keyword GOLD, so characters were shifted in turn by 7, 15, 12, 4. Such a cipher is very hard to crack.

The method we recommend is due to Babbage and Kasiski who independently discovered it, and is based on the regularity of the repetition. An analysis of repeated strings of letters is used to try to determine the length of the keyword, and once this is done a standard frequency analysis is applied to each part of the ciphertext encoded by a single cipher. A very good account of Babbage-Kasiski deciphering can be read in chapter 2 of Simon Singh's *The Code Book*.

On transposition ciphers

Sometimes when you carry out a frequency analysis you will find that each letter occurs with about the same frequency as you would expect in natural English text (or whichever language you are studying). This is a broad hint that the text is not enciphered using a substitution cipher, but rather by an anagram or transposition cipher, also known as an anagram cipher. In such a cipher the letters of the message are not replaced by substitutes, but rather jumbled using some rule which allows them to be untangled again to decipher the message.

Example

We will encipher the text:

```
The quick brown fox jumps over the lazy dog
```

We start with a keyword. Suppose our keyword is BAD. We write it at the head of a table with three columns, then enter the ciphertext in the boxes below. The last, empty, box is paddded with an X.

B	A	D
t	h	e
q	u	i
c	k	b
r	o	w
n	f	o
x	j	u
m	p	s
o	v	e
r	t	h
e	l	a
z	y	d
o	g	x

Next we rearrange the columns so that the letters in the keyword are now in alphabetic order

A	B	D
H	T	E
U	Q	I
K	C	B
O	R	W
F	N	O
J	X	U
P	M	S
V	O	E
T	R	H
L	E	A
Y	Z	D
G	O	X

Giving a ciphertext of

`HTEUQIKCBORWFNOJXUPMSVOETRHLEAYZDGOX`

If the keyword contains repeated letters then we delete them as we would if it were the keyword for a substitution cipher before constructing the grid. Hence if the keyword was TOFFEE we would use a grid of width 4 with header TOFE and we would rearrange the grid so that its header appeared as EFOT to encipher the message.

How do we tackle such a cipher?

Clearly the length of the keyword is quite crucial. You should be able to guess this from the length of the ciphertext, which will be a multiple of it. So in our example the ciphertext has length 36 which has factors 2,3,4,6,9 and so on. Hence we could try laying out the text in grids of width 2,3,4,6,9 respectively (a keyword of length 12 or more is unlikely) and examining the rows.

Of course a grid of width 2 would leave us just switching alternate letters so we probably don't need to lay it out that way to check it. Having checked and dismissed the idea of a keyword of length 2 the first grid we try looks like the second grid above. Having got to this point the best hope for a quick solution is to find a crib. If there is a word you think ought to appear in the cipher text then you could try looking for anagrams of that word. This is made difficult by the fact that in splitting the text into blocks (blocks of three in the example), If your crib word does not take up an entire block then even the characters from the crib that do appear will be jumbled with other nearby characters, so you need a reasonably long crib. On the other hand if it is too long only part of the word will appear in that block so you are looking for anagrams of parts of the crib.

In our example if we knew, for some reason, that the text was likely to contain the word "jumps" we could look for anagrams of "jum", "ump", "mps. Looking carefully you should see the anagram PMS in the text and we might guess that the first and second columns have been transposed while the third has remained fixed. Checking this we find have cracked the cipher.

Things are harder with longer keywords but the principle remains the same. Things get tougher if the plaintext is not in our own language, since it is harder to say what makes sense. Of course even in this case it may be that part of the message is in your language and the rest in another. In this case you might hope to crack the ciphertext corresponding to your native language, and apply the knowledge that gives you about the cipher to write down a decrypt of the entire message, even when the text is unfamiliar.

Other (subtle) cribs: In English the letters q and u occur together so if they are separated either you are not looking at English text or they should be brought back together by undoing the anagram.

Numbers often represent dates, so for example the letters/numbers 2, 1, s, t in proximity might represent 21^{st}, while 2,1,t,h might represent 12^{th}.

Cryptography Lesson Plan 1

Class: Cracking the Caesear shift ciphers.
Resources:

- **Leaflet "On substitution ciphers".**
- **Two handouts each with a plaintext and a cipher table**
- **Teachers' solutions for the handouts.**
- **One OHP slide with cipher text to crack, and partial decrypt and solution.**

Starter: (10 minutes approximately) Uses handouts for Groups A and B
Encryption exercise – split the class into groups A and B. Give each group the enclosed text to encipher using the given code. Encourage accuracy AND secrecy! Answers enclosed with handouts.

Main activity: (40 minutes approx) Uses OHP

- *Introduce the idea of a substitution cipher in general and the Caesar shift in particular.*
- *Suggest trial and error as a deciphering technique.*
- *Work through a very simple Caesar shift (by 3).*
- *Split the class again, swap over the ciphertexts from the starter exercise and get them to tackle them.*

Plenary (approx 10 minutes)

Discuss how to make the code harder to crack using a rule that is harder to determine, but remark on the need for an easy to remember rule (stressed agents must remember it and can't write it down!) Mention "keyword" substitution.

Handout for lesson 1.

GROUP A

Code: Caesar shift by 2

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B

Plaintext

There were plenty of schools in the world, but they were all run either by the various churches or the Guilds. Miss Butts objected to churches on logical grounds and deplored the fact that the only Guilds that considered girls worth educating were the Thieves and the Seamstresses. It was a big and dangerous world out there and a girl could do worse than face it with a sound knowledge of geometry and astronomy under her bodice.

From "Soul Music" by Terry Pratchett.

Handout for lesson 1.

GROUP B

Code: Caesar shift by 4

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D

Plaintext

The four houses are called Gryffindor, Hufflepuff, Ravenclaw and Slytherin. Each house has its own noble history and each has produced outstanding witches and wizards. While you are at Hogwarts, your triumphs will earn your house points, while any rule-breaking will lose house points. At the end of the year the house with the most points is awarded the House Cup, a great honour. I hope each of you will be a credit to whichever house becomes yours.
From “Harry Potter and the Philosopher’s Stone” by J.K. Rowling.

Teachers' solutions to encryption challenge

Ciphertext A

VJGTG YGTG RNGPVA QH UEJQQNU KP VJG YQTNF, DWV VJGA YGTG CNN TWP GKVJGT DA VJG XCTKQWU EJWTEJGU QT VJG IWKNFU. OKUU DWVVU QDLGEVGF VQ EJWTEJGU QP NQIKECN ITQWPFU CPF FGRNQTGF VJG HCEV VJCV VJG QPNA IWKNFU VJCV EQPUKFGTGF IKTNU YQTVJ GFWECVKPI YGTG VJG VJKGXGU CPF VJG UGCOUVTGUUGU. KV YCU C DKI CPF FCPIGTQWU YQTNF QWV VJGTG CPF C IKTN EQWNF FQ YQTUG VJCP HCEG KV YKVJ C UQWPF MPQYNGFIG QH IGQOGVTA CPF CUVTQPQOA WPFGT JGT DQFKEG.

Ciphertext B

XLI JSYV LSYWIW EVI GEPPIH KVCJJMRHSV, LYJJPITYJJ, VEZIRGPEA ERH WPCXLIVMR. IEGL LSYWI LEW MXW SAR RSFPI LMWXSVC ERH IEGL LEW TVSHYGIH SYXWXERHMRK AMXGLIW ERH AMDEVHW. ALMPI CSY EVI EX LSKAEVXW, CSYV XVMYQTLW AMPP IEVR CSYV LSYWI TSMRXW, ALMPI ERC VYPI-FVIEOMRK AMPP PSWI LSYWI TSMRXW. EX XLI IRH SJ XLI CIEV XLI LSYWI AMXL XLI QSWX TSMRXW MW EAEVHIH XLI LSYWI GYT, E KVIEX LSRSYV. M LSTI IEGL SJ CSY AMPP FI E GVIHMX XS ALMGLIZIV LSYWI FIGSQIW CSYVW.

OHP Slide for lesson 1

Ciphertext

WKH HDVLHVW PHWKRG RI HQFLSKHULQJ D WHAW PHVVDJH LV WR UHSODFH HDFK FKDUDFWHU EB DQRWKHU XVLQJ D ILAHG UXOH, VR IRU HADPSOH HYHUB OHWWHU D PDB EH UHSODFHG EB G, DQG HYHUB OHWWHU E EB WKH OHWWHU H DQG VR RQ.

Partial decrypt: Guess that the first word is "the" so that t is enciphered as W, h as K and e as H. This suggests a shift by 3:

the eDVLeVt PethRG RI eQFLSheULQJ D teAt PeVVDJe LV tR UeSODFe eDFh FhDUDFteU EB DQRtheU XVLQJ D ILAeG UXOe, VR IRU eADPSOe eYeUB OetteU D PDB Ee UeSODFeG EB G, DQG eYeUB OetteU E EB the OetteU e DQG VR RQ.

The word teAt could be tent, test or text, with text fitting with the shift by 3; the word OetteU which occurs twice, would decipher to "letter" confirming our guess.

Code: Caesar shift by 3

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Plaintext

The easiest method of enciphering a text message is to replace each character by another using a fixed rule, so for example every letter a may be replaced by d, and every letter b by the letter e and so on.

Cryptography Lesson Plan 2

Class: Cracking keyword substitution ciphers – emphasises letter frequency analysis and team work.

Resources:

- **Leaflet "On substitution ciphers".**
- **OHP 1 containing ciphertext**
- **OHP 2 Containing expected frequency table and incomplete actual frequencies.**
- **Handout summarising details of deciphering technique.**
- **OHP 3 With further thoughts on disguising the text.**

Starter: (10 minutes approximately) (Uses OHP 1)

Split the class into teams and get them to count the letter frequencies in the ciphertext. Emphasise the need for speed and accuracy. Maybe set the scene as a race against time.

Main activity: (30 minutes approx) (Uses OHP 1 and OHP 2 and handout)

- *Introduce the idea of a keyword cipher to make encryption more secure and more memorable (see "On substitution ciphers").*
- *Discuss the hunt for common words and letters and introduce frequency analysis – show a table of common frequencies and check it against the examples in lesson 1.*
- *Discuss the speed improvements given by parallel processing of the text. Split into 26 teams to do a frequency analysis of the given ciphertext on OHP 1. {It may be worth remarking that standard computer attacks on ciphers use this idea of parallel processing to speed up the attack.)*
- *Whole class session to construct frequency table, compare with expected frequencies (computed from percentages) and identify the letters "e"*

Plenary (20 minutes approx) (Uses OHP 2 and, time permitting OHP 3)

Draw together the intelligence gained by the groups and crack the cipher together. (You may wish to give out the handout summarising the technique after completing the exercise.)

If time permits (OHP 3):

- *Discuss how to make the code harder to crack by disguising the letter groups.*
- *Remark that the frequency table can mislead for non-standard or foreign language texts! Examine the extract from the book "A Void" by Georges*

OHP Slide 1 for lesson 2

Ciphertext

VEP HYXHLVHTP MO AWFJYFLT H RFNEPS HJNEHAPV FL VEFU ZHC FU VEHV FV FU PHUC VM KPKMSFUP VEP IPCZMSY MS IPCNESHUP, HLY EPLRP VEP RFNEPS HJNEHAPV. VEFU FU FKNMSVHLV, APRHWUP FO VEP UPLYPS EHU VM IPPN VEP RFNEPS HJNEHAPV ML H NFPRP MO NHNPS, VEP PLPKC RHL RHNVWSP VEP NHNPS, YFURMXPS VEP IPC, HLY SPHY HLC RMKKWLFRHVFMLU VEHV EHXP APPL PLRSCNVPY ZFVE FV. EMZPXPS FO VEP IPC RHL AP RMKKFVVPY VM KPKMSC FV FU JPUU JFIPJC VM OHJJ FLVM PLPKC EHLYU.

email:cipher@soton.ac.uk

OHP Slide 2 for lesson 2

Occurrences table

a	b	c	d	e	f	g	h	i	j	k	l	m
n	o	p	q	r	s	t	u	v	w	x	y	z

Expected Frequency table

a	b	c	d	e	f	g	h	i	j	k	l	m
8.2	1.5	2.8	4.3	12.7	2.2	2.0	6.1	7.0	0.2	0.8	4.0	2.4
n	o	p	q	r	s	t	u	v	w	x	y	z
6.7	7.5	1.9	0.1	6.0	6.3	9.1	2.8	1.0	2.4	0.2	2.0	0.1

This table was taken from "The Code Book" by Simon Singh, and gives expected frequencies as a percentage. To accurately compare it to the actual frequencies above you should compute the actual frequencies as percentages.

Actual Frequencies as percentages

a	b	c	d	e	f	g	h	i	j	k	l	m
n	o	p	q	r	s	t	u	v	w	x	y	z

Handout for lesson 2

STAGE 1 – P is the commonest letter in the ciphertext so could stand for e - maybe the first word is the:

the HYXHLtHTe MO AWFJYFLT H RFNheS HJNhHAet FL thFU ZHC FU thHt Ft FU eHUC tM KeKMSFUe the IeCZMSY MS IeCNhSHUe, HLY heLRe the RFNheS HJNhHAet. thFU FU FKNMStHLt, AeRHWUe FO the UeLYeS hHU tM IeeN the RFNheS HJNhHAet ML H NFeRe MO NHNeS, the eLeKC RHL RHNtWSe the NHNeS, YFURMXeS the IeC, HLY SeHY HLC RMKKWLFRHtFMLU thHt hHXe AeeL eLRSCNteY ZFth Ft. hMZeXeS FO the IeC RHL Ae RMKKFtteY tM KeKMSC Ft FU JeUU JFIeJC tM OHJJ FLtM eLeKC hHLYU.

STAGE 2 We see the single letter word H, and the four letter word thHt in the first line - guess that H encodes the letter a.

the aYXaLtaTe MO AWFJYFLT a RFNheS aJNhaAet FL thFU ZaC FU that Ft FU eaUC tM KeKMSFUe the IeCZMSY MS IeCNhSaUe, aLY heLRe the RFNheS aJNhaAet. thFU FU FKNMStaLt, AeRaWUe FO the UeLYeS haU tM IeeN the RFNheS aJNhaAet ML a NFeRe MO NaNeS, the eLeKC RaL RaNtWSe the NaNeS, YFURMXeS the IeC, aLY SeaY aLC RMKKWLFRatFMLU that haXe AeeL eLRSCNteY ZFth Ft. hMZeXeS FO the IeC RaL Ae RMKKFtteY tM KeKMSC Ft FU JeUU JFIeJC tM OaJJ FLtM eLeKC haLYU.

STAGE 3 The two 2 letter words Ft FU are probably it is meaning that F encodes i and U encodes s:

the aYXaLtaTe MO AWiJYiLT a RiNheS aJNhaAet iL this ZaC is that it is easC tM KeKMSise the IeCZMSY MS IeCNhSase, aLY heLRe the RiNheS aJNhaAet. this is iKNMStaLt, AeRaWse iO the seLYeS has tM IeeN the RiNheS aJNhaAet ML a NieRe MO NaNeS, the eLeKC RaL RaNtWSe teh NaNeS, YisRMXeS the IeC, aLY SeaY aLC RMKKWLiRatiMLs that haXe AeeL eLRSCNteY Zith it. hMZeXeS iO the IeC RaL Ae RMKKitteY tM KeKMSC it is Jess JiIeJC tM OaJJ iLtM eLeKC haLYs.

STAGE 4: haXe = have, easC = easy and so on - we get the following extract from Simon Singh's excellent history of codes and ciphers, *The Code Book*:

"The advantage of building a cipher alphabet in this way is that it is easy to memorise the keyword or keyphrase, and hence the cipher alphabet. This is important, because if the sender has to keep the cipher alphabet on a piece of paper, the enemy can capture the paper, discover the key, and read any communications that have been encrypted with it. However if the key can be committed to memory it is less likely to fall into enemy hands."

OHP Slide 3 for lesson 2

Obscuring a substitution cipher

1. We can disguise the word structure by regrouping the letters into blocks:

VEPHY XHLVH TPMOA WFJYF LTHRF NEPSH JNEHA PVFLV EFUZH CFUVE HVFVF UPHUC VMKPK MSFUP VEPIP CZMSY MSIPC NESHU PHLYE PLRPV EPRFN EPSHJ NEHAP VVEFU FUFKN MSVHL VAPRH WUPFO VEPUP LYPSE HUVMI PPNVE PRFNE PSHJN EHAPV MLHNF PRPMO NHNPS VEPPL PKCRH LRHNV WSPVE PNHNP SYFUR MXPSV EPIPC HLYSP HYHLC RMKKW LFRHV FMLUV EHVEH XPAPP LPLRS CNVPY ZFVEF VEMZP XPSFO VEPIP CRHLA PRMKK FVVPY VMKPK MSCFV FUJPU UJFIP JCVMO HJJFL VMPLP KCEHL YU

2. We can distort the frequency table – this text was adapted for last years cipher challenge!

Augustus, who has had a bad night, sits up blinking and purblind. Oh what was that word (is his thought) that ran through my brain all night, that idiotic word that, hard as I'd try to pin it down, was always just an inch or two out of my grasp - fowl or foul or Vow or Voyal? - a word which, by association, brought into play an incongruous mass and magma of nouns, idioms, slogans and sayings, a confusing, amorphous outpouring which I sought in vain to control or turn off but which wound around my mind a whirlwind of a cord, a whiplash of a cord, a cord that would split again and again, would knit again and again, of words without communication or any possibility of combination, words without pronunciation, signification or transcription but out of which, notwithstanding, was brought forth a flux, a continuous, compact and lucid flow: an intuition, a vacillating frisson of illumination as if caught in a flash of lightning or in a mist abruptly rising to unshroud an obvious sign - but a sign, alas, that would last an instant only to vanish for good.

From “A Void” by Gilbert Adair. The letter "e" does not appear even once in the book!

Cryptography Lesson Plan 3

Class: Affine shift ciphers – emphasises clock arithmetic and gives more practice at frequency analysis.

Resources:

- **Leaflet "On substitution ciphers".**
- **OHP 1, giving partial encryption table for the 3x+5 affine shift cipher together with teachers' solution.**
- **OHP 2-4, with cipher text to crack, method and solution.**

Starter: (10 minutes approximately) Uses handout

Complete the encryption table on the OHP (the affine shift cipher $x \rightarrow 3x+5$ *is discussed in the teachers' notes).*
Encourage them to try to spot the pattern and guess the rule which should be concealed.

Main activity: (40 minutes approx) Uses OHP

- *Introduce the class of affine shift ciphers mentioning "clock arithmetic" mod 26*
- *Show them that the cipher table arises from the affine shift* $x \rightarrow 3x + 5$.
- *Discuss the fact that you only need to know the value of two letters to deduce the affine shift (solving two simultaneous equations mod 26).*
- *Use frequency analysis and modular arithmetic to decipher an affine shifted text together or in groups.*

Plenary (approx 10 minutes)

Discuss generalisations to modular arithmetic mod n.

OHP slide 1 for lesson 3.

Spot the pattern?

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
H	K	N					C																		

$x \rightarrow 3x + 5$

1	2	3	4	5	6	7	8	9	10	11	12	13
8	11	14	17	20	23	26	3	16	9	12	15	18
14	15	16	17	18	19	20	21	22	23	24	25	26
21	24	1	4	7	10	13	16	19	22	25	2	5

Encryption table

a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
H	K	N	Q	T	W	Z	C	F	I	L	O	R	U	X	A	D	G	J	M	P	S	V	Y	B	E

email:cipher@soton.ac.uk

OHP Slide 2 for lesson 3

Ciphertext

LMYFU BKUUS DDYFA XWCLA OLPSF AOLMJ FASDS NSFGJ
FAOEL SOMYT DJLAX EMHJM BFMIB JUMIS HFSUL AXUBA
FKJAM XLSKF FKXWS DJLSO FGBJM WFKIU OLFMX MTMWA
OKTTG JLSXL SKFFK XWSDJ LSIZG TSXWJ LJLSX LSUMF JSDJL
SIZGH FSQYS XOGLS DMMDT SDMXJ LSBAT SMHBK BSFLS
BFMCT SDKFM YXDJL SLYJM ZTANA MYUXM CJMCL MCKUT
MMEAX WKJLA IKXDC LMCKU XJJLA UCKUC LKJAJ LKDZS SXTAE
SHMFJ SXAXJ SFIAX KZTSI MXJLU TKUJG SKFXM CMXDS FLSLK
DWMXS IKDJL SOLMF YUTAX SMHIS KXAXW TSUUT SJJSF
UDKXO SDZSH MFSLA USGSU ZYJJL SGCSF SXMJI SKXAX WTSUU
JLSGC SFSTM KDSDC AJLJL SIMUJ NAJKT ISKXA XWAIK WAXKZ
TSAHM XTGLS OMYTD HAXDA JZYJC LSFSC KUJLS BKJJS FXCLS
FSCKU JLSBK JJSFX CLSFS CKUJL SBKJJ SFXHF MISXA WIKZG
FMZSF JLKFF AU

Occurences table:

A	B	C	D	E	F	G	H	I	J	K	L	M
34	12	19	23	4	41	12	10	16	50	38	46	42
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
3	11	1	1	0	78	22	28	0	13	41	11	11

Use frequency analysis to guess that S enciphers for e, and J for t.

This tells us that for an affine shift cipher

$x \rightarrow ax + b$

$a.5 + b = 19$ (e → S)
$a.20 + b = 10$ (t → J)

Solving mod 26 we see that $15.a = -9$ mod 26. Now 7.15 is congruent to 1 mod 26 since 7.15 = 105 = 104 +1= 4.26 +1. It follows that $7.15.a = 7.-9$, or a is congruent to -63.

Now -63 = -52 - 11, so a is congruent to -11, or equivalently to 15 mod 26. Hence $a = 15$. Now from $a.5 + b = 19$ we get $75 + b$ is congruent to 19, or b is congruent to –56 mod 26.

Since -56 = -2.26 - 4, b is congruent to -4 mod 26 so $b = 22$.

To check this $20.a + b = 300 + 22 = 322 = 12.26 + 10$, so $a.20 + b = 10$ as required. So the affine shift is $x \rightarrow 15x+22$ and the decrypt is given by the inverse function $y \rightarrow 7(y-22)$

[It might look strange but "dividing by 15" is the same as multiplying by 7 in mod 26 arithmetic.]
Equivalently the decryption is achieved by the affine shift $y \rightarrow 7y+2$.

Encryption table:

a	b	c	d	e	f	g	h	i	j	k	l	m
I	P	W	D	K	R	Y	F	M	T	A	H	O
n	o	p	q	r	s	t	u	v	w	x	y	z
V	C	J	Q	X	E	L	S	Z	G	N	U	B

OHP Slide 4 for lesson 3

Decrypt

hours passe dduri ngwhi chjer ichot riede veryt rickh ecoul dthin kofto promp tsome fresh inspi ratio nhear range dthec rypto grams chron ologi cally thenh earra ngedt hemby lengt hthen hesor tedth embyf reque ncyhe doodl edont hepil eofpa perhe prowl edaro undth ehuto blivi ousno wtowh owasl ookin gathi mandw howas ntthi swasw hatit hadbe enlik efort enint ermin ablem onths lasty earno wonde rheha dgone madth echor uslin eofme aning lessl etter sdanc edbef orehi seyes butth eywer enotm eanin gless theyw erelo adedw ithth emost vital meani ngima ginab leifo nlyhe could findi tbutw herew asthe patte rnwhe rewas thepa ttern where wasth epatt ernfr omeni gmaby rober tharr is