XX10190: Feedback on Section B of the May exam

Let us begin with pink paper.
Most of you can handle the pink paper, but 34 out of the first 100 scripts I marked did not have the questions you had attempted ticked. I specifically asked you to do that, but you were all anxiously filling in your calculator number which, whatever the pink paper may say, is of no importance at all.
Three of you stuck the flap down in the wrong place, which leaves your name visible, thus exposing you to whatever danger it is that having your name known to us is supposed to present. We have over 300 scripts to mark: we don't stop to read people's names.
More seriously, eight of you handed in the two sections in one pink sheet, instead of separately as you were told. This is very inconvenient. Only one of you had actually written a Section A question and a Section B question on the same sheet of paper: six of the other seven simply ignored the instructions. The seventh had retired ill and understandably just handed in whatever was available.
The actual exam was rather mixed. There was a sharp divide between those who had come to the problems classes and those who hadn't. The former had seen examples showing them how to do the questions. This is an advantage. By watching us write mathematics, you also learn how to write mathematics, which is not something you are likely to be naturally good at. In particular, what you write should make sense when read out. For instance, $\forall$ means "for all", not just "all". The abbreviation for "all" is "all". And – although this is nothing to do with mathematics – "it's" means "it is" or "it has" but not "of it", which is "its". Students whose native language is English were if anything more likely to get this wrong.

Q3. Most of you know what a cyclic group is and what a generator is, but not all. It is not quite enough to take $G = \{g^a \mid a \in \mathbb{N}\}$ because if G might be infinite and then you don't have inverses: you have to put $\mathbb{Z}$ rather than $\mathbb{N}$. More generally, if you are asked what it means for a group G to be cyclic, your answer should begin "G is cyclic if…", not "If G is cyclic then…".
You were less clear about the meaning of $\mathbb{F}_p$. Some of you wrote phrases like "a field in modulo p", which certainly wouldn't explain what $\mathbb{F}_p$ is to anybody who didn't already know is, and in fact doesn't mean anything at all. Some of you said it was the numbers $1, \ldots, p$ but didn't say any more, not bothering to mention what the operations are. $\mathbb{F}_p^*$ caused even more problems, but mostly later: however, you must not call $\mathbb{F}_p^*$ a field. It isn't a field: it has no zero. Generally, the words "group", "ring" and "field" are protected and mustn't be used loosely. The same goes for notation. A lot of you wrote $\mathbb{F}_P^* = \mathbb{F}_p/\{0\}$, but the notation for set difference isn't /: it's $\setminus$ (or even – but that already means so many other things that I personally avoid it). You are no more free to change that than you are free to decide that $\cup$ means intersection.
You don't know what Fermat's Little Theorem is. A lot of you just wrote $a^p = a \bmod p$ without bothering to say what a or p were. This doesn't get you any marks. What if I take $a = \sqrt{-1}$? Or $p = 100$? Those who did know mostly couldn't prove it. Often you simply quoted the more general Euler-Fermat theorem, which is no help: I want a proof, not a stronger statement.
The next part asks about the key exchange version of DH and a lot of you wrote about the messaging version regardless. This mostly didn't matter very much, but it did cost marks: you should read the question. More serious were incomplete explanations, saying things like "they choose the message from a public list". A list of what? "A public list" could be some junk on Buzzfeed. Even commoner was not saying what a and b are: just "Alice chooses a".
But the most common, and most serious, mistake, was to say that $a \in \mathbb{F}_p^*$. No, it is definitely not. That means a non-zero integer mod p, and what you want is an integer mod $p-1$. It is true that there are $p-1$ of each (though actually you want a non-zero integer mod $p-1$, and some of those are no good either) but that doesn't mean they are the same things. There are seven days of the week but if I gave you an appointment for 10:15 on Sneezy or half-past one on the Mausoleum of Halicarnassus you would complain, even though there are seven dwarves and seven wonders of the ancient world too.
Still, most of you did more or less know how DH works. Rather fewer understand the discrete log problem. I didn't directly ask you say what it is, but if your explanation of why the system would be insecure if Eve could solve it quickly left me doubtful about whether you even knew what it was, you did not get full marks. It's no use just saying "if Eve could find dlogs it would be easy". I've just told you that. I asked you why it would be easy.

One detail that a lot of you got wrong, although I didn't dock a mark for it, is that Eve doesn't *have* to solve the discrete log problem. She's really after the shared secret, and she might be able to get that without solving the dlog problem – either mathematically, or by waiting till Fred tells her. But if she can solve the dlog problem, then she gets the secret. Another thing (which again I didn't dock a mark for) is that Alice shouldn't choose m, at least, not if Bob has any sense. She and Bob should choose m at random from a list. Mathematically it all works perfectly well if Alice chooses m, but Bob doesn't trust her so he shouldn't let her take control of the process. She might be able to pick an m that would make the dlog problem in base m easy for her, and so when Bob sends her m^b she can compute b.
The computation at the end was done pretty well. You should use a repeated squares method (most people did) and you could have saved some time if you had noticed that $65 = -2$, which is rather easier for humans to compute with. Some of you forgot the slogan and worked mod 66 instead of mod 67, which was disastrous. A few pointed out that $31 \times 13 = 7 \bmod 66$ (these are keys, so I do mean 66 this time) and then simply computed $10^7 \bmod 67$, which was very quick and perfectly correct since I didn't actually ask you to perform Alice's and Bob's calculations.

Q4. You do know what $a|b$ means. This is progress. In previous years the idea that it is a number has been very widespread. This time only a few people were inattentive enough to think that.
You were vaguer about the totient function. I didn't drop a mark for giving the range $0 \leq k < n$, although it should be $0 \leq k \leq n$ – it only makes a difference for $n = 1$. I even allowed $1 \leq k < n$, which doesn't even make sense for $n = 1$. Still, you ought to be alert to such things, especially as I did point this out in the lectures. Most attempts at a quick explanation for the product formula were successful, but many people made no attempt.
Similarly, there were people who could remember a proof of the Möbius formula and a much larger number who couldn't and didn't try. A misprint at this point confused almost nobody.
Most of you can explain RSA. There was some confusion about what the valid choices of a are, but not much. The number of people who can't distinguish between RSA and DH was lower this year. However, any mention of the discrete log problem at this point cost marks. The discrete log problem has nothing to do with RSA.
The last part got rather few fully correct answers. A lot of you thought that everything was fine. The mention of Fred was supposed to be a hint that this is not the case. Another large group of you extracted m^2 from the data (not spotting how easy it is to extract m) and then advised Eve to take square roots in $\mathbb{Z}/N$. But how can she do that? If she could take cube roots in $\mathbb{Z}/N$ she would have everything she needs anyway, because that's what Alice does when she gets Bob's message: why should square roots be any easier? It's as if Alice has taken $a = 2$, and in fact she could almost do that: the only thing that stops her is that $\gcd(2, \varphi(N)) = 2$, which reflects the fact that asking Bob to square his message would lose the distinction between m and $-m$. As long as you are confident that m and $-m$ aren't both going to make good sense you could even permit that.
A few of you answered that since Fred has apparently changed the value of a without informing Alice, the result will be that when Alice tries to decrypt Bob's message sent with $a = 5$ she will do so assuming that $a = 3$, and therefore get rubbish. I gave this some credit, because it's true; but the situation is worse than that.

Q5. You mostly know what a code is. A lot of people couldn't remember what the rate is, which is perhaps not surprising as it didn't figure very much in the course. Many of you cannot remember what Hamming distance is, which is more surprising because it is really very easy, but most of you can get from there to the minimum distance.
Some people managed to give completely satisfactory one-line answers to (iii), which was generally done well by those who attempted it; but evidently a significant minority simply have no idea about this.
Most of you wrote down the correct definitions in (iv). Again, a minority (quite small now) had no idea, which means you didn't learn about this topic and can have no complaints at not getting any marks for it.
Part (v) floored a few people by being over $\mathbb{F}_3$ rather than $\mathbb{F}_2$, but a more common error was not realising that there was anything to check at all. A few people simply did part (vi) first and then pointed out that since a decoder matrix existed (they've just found one) everything must be all right. Completely true, and full marks. But most people who did this well found the rank of G by row reductions and then moved on

to (vi). This was where non-attendance at the examples classes did most damage. A remarkable number of you, though, wrote down the matrix consisting ofthe first three rows of G and defiantly tried to find its inverse, ignoring the fact that as its last column is zero it obviously doesn't have one. A few even found something that they claimed was its inverse. Sometime the purported inverse invoved thirds, which looks all right until you realise that $3 = 0$ in $\mathbb{F}_3$ so you mustn't go dividing by 3. Some people made that mistake without trying to take the impossible inverse matrix. Some realised that you can proceed by changing the order of the rows (that's not the only way) and then forgot to change back at the end: this still got most of the marks. But actually, most people who got this far gave correct answers. Many of you gave H as a 3×5 matrix of rank 2 rather than as a 2×5 matrix, but that is still correct and you didn't lose any marks for the slight inefficiency.

GKS, 6/6/16