Talk at Edinburgh Filmhouse

Tom Leinster

For the event 'The Maths Behind *The Imitation Game*',
14 December 2014

Note for the public version Although some of the factual assertions made here may seem hard to believe, they are all carefully based on documentary evidence. Much of it comes from GCHQ's and the NSA's own internal documents, leaked by Edward Snowden.

As these notes were written for my own use when giving the talk, rather than for others to read, citations to that evidence are not included. However, most of the facts mentioned below also appear in blog posts I have made over the last couple of years, and links and citations are scrupulously provided there. An index of those posts is at https://golem.ph.utexas.edu/category/2014/07/math_and_mass_surveillance_a_r.html (that's a clickable link).

1 Preamble

We're here to discuss Turing's legacy, and I'm going to talk about the legacy of his work at Bletchley Park.

In 1946, the organization at Bletchley Park became GCHQ: Government Communications Headquarters, the UK's agency for communications intelligence, now based in Cheltenham.

In some ways, it has changed enormously since Turing's day. In some ways, it has stayed much the same.

Differences

- *Industrial scale*
  - In the film, the code-breaking team is characterized as 'half a dozen crossword enthusiasts in a tiny village in the south of England'.

  - It no longer relies on a few geniuses (if it ever did). Today, it's organized on an industrial scale.
  - It's part of a huge network of intelligence agencies: GCHQ not only works with MI5 and MI6, but also works extremely closely with the NSA (the American National Security Agency, the US equivalent of GCHQ), which in turn works closely with agencies such as the CIA.

- *Threat level*
  - In Turing's time, we were at war with the Nazis, who were exterminating millions. On an average weekend in World War Two, 60 British *civilians* (not soldiers) were killed.
  - Today, the usual justification for the agencies' activities is terrorism, which in Britain has killed fewer than 60 people in the last *ten years*.

- *Who they're spying on*
  - In Turing's time, it was the Nazi military.
  - Today, it's us. It's everyone, regardless of suspicion.
  - In the documents revealed by Edward Snowden, there's an NSA document describing their 'collection posture' as follows: 'Collect it all. Sniff it all. Know it all. Exploit it all.'
  - There's also a slide from GCHQ stating that they collect more than 50 *billion* communications every *day*. (That was in 2011; it won't have gone down since then.)

Similarities

- *They're still very interested in mathematicians*
  - The NSA is said to be the largest employer of mathematicians in the world. GCHQ is also a major employer of mathematicians in the UK.
  - Some mathematicians work there full time. Some work there over summers or on sabbaticals.
  - They may not know how their work will be used. I know mathematicians who have worked for GCHQ and came to sorely regret it after the Snowden revelations, not having known that they were working for an agency of mass surveillance.

- *Secrecy and detachment from the democratic process*
  - In the film, you see how much Bletchley Park is independent of democratic control. That's still the case for GCHQ now.
  - After the Snowden revelations began, on both sides of the Atlantic, senior politicians who were members of national security committees complained that they didn't even *know* about the mass surveillance programmes, much less *approve* them.
  - Of course, neither did we.

2 Code-breaking

I'm going to talk about two mathematical themes of Turing's work at Bletchley Park. The first is code-breaking.

Code-breaking is inherently mathematical.

Flavour

- Suppose I give you two 10-digit prime numbers and ask you to multiply them together. It's easy: just take out your phone or calculator, and it does it in an instant. You get a 20-digit answer.
- On the other hand, suppose I give you a 20-digit number, tell you that it's equal to one prime times another, and ask you to find out what those primes are. Then it's a very slow process: you can't do much better than going through all the possibilities (is it divisible by 2? by 3? and so on).
- When you hear that some code will take 10 years to break, that's roughly what it's about: the slowness of factorizing prime numbers.
- In the film, when Turing and Joan Clarke are having a picnic together, you see that someone's written '$n = pq$' on a piece of paper. This is what it's referring to. (The film must have had a mathematician as a consultant.)

One thing the Snowden documents make clear: it is the explicit aim of the NSA and GCHQ that no two human beings are able to communicate digitally without them being able to know the content.

- E.g. line from 2013 NSA budget request: they sought funding to 'Insert vulnerabilities into commercial cryptosystems'.

- This is done in several ways: e.g. by influencing industrial standards and by hacking software.
- Going back to the 1970s, there have been several well-documented instances of this.
- A major problem: *even if* — for some reason — you completely trust the secret, unaccountable intelligence agencies, weakening cryptographic systems makes them vulnerable to all attackers (fraudsters etc.), not just the agencies themselves.

Have GCHQ and the NSA broken internet encryption?

- A casual reading of journalism based on the Snowden leaks might lead you to think 'yes'.
- But actually, no. One ray of light from the Snowden revelations is that if it's *properly* implemented, encryption works. The agencies have enormous computers and smart employees, but apparently they haven't made any major mathematical breakthroughs enabling them to solve '$n = pq$' much more efficiently than anyone else.
- What the agencies actually do is 'cheat', e.g. by intercepting communications before they're encrypted or after they're decrypted. Example:
  - Yesterday, conclusive evidence was revealed that GCHQ had inserted very sophisticated malware (malicious software) into the systems of Belgacom, Belgium's largest telecommunications provider, which serves EU institutions and is a major international hub.
  - It infiltrated the system comprehensively enough that GCHQ were able to read communications *before* they were encrypted, so that they didn't actually need to break the encryption.

3 Algorithms

Second mathematical theme.

Turing was one of the first people to appreciate the power of algorithms executed by machine.

There's an obvious sense of that power which we're all highly aware of:

- Computing devices are everywhere in our lives

- Much of our communication is digital (e.g. email, mobile, text)
- On the other hand, we now know that if, for instance, an NSA analyst wants to read your email, all they have to do is sit at their desk, type your email address into a box, and click a mouse.

But there's also a more subtle sense: the power of algorithms to draw conclusions from large amounts of data.

- – Shortly after the first Snowden revelations, Barack Obama said 'No one is listening to your phone calls'. British spy chiefs have said similar things.
  - – But true or false, it's a red herring.
  - – With GCHQ intercepting more than 50 billion communications per day, the most useless thing possible would be to have a human being eyeballing the data.
  - – Algorithms are much more powerful, and have far greater invasive power.
- – For instance, there are algorithms tracking who visits the Wikileaks websites, who downloads privacy-guarding software such as Tor, and who reads certain technical journals. Together, these give a good idea of who might try to obstruct some of the intelligence agencies' aims.
  - – Similarly, having a person listen to the content of phone calls is very expensive. Having algorithms record and analyze who calls who when — 'metadata' — is vastly more effective.
- In case you're in any doubt as to the power of metadata:
  - – Stewart Baker, former senior lawyer at NSA, said: 'Metadata absolutely tells you everything about somebody's life'.
  - – Michael Hayden, former head of both NSA and CIA, added: 'We kill people based on metadata'.
  - – What he was probably referring to was the CIA's drone assassination programme, which the NSA supplies intelligence to. Some of its targets are killed based on a probabilistic assessment of phone metadata.

4 Ending

I want to finish by saying something about the end of Alan Turing's life.

- – 60 years ago, Turing was chemically castrated for having sex with another man. He's widely thought to have killed himself as a direct result.
  - – Today, two men can not only legally have sex, but even get married with the full endorsement of the state.
  - – That enormous change largely came about through persistent campaigning.
- – Campaigning to decriminalize anything puts you in a vulnerable position.
  - – By definition, you're siding with criminals against the power of the establishment.
  - – Over the years, there have been hundreds of documented instances of intelligence agencies disrupting legal campaigning and destroying the reputations of campaigners who have never committed any crime.
- – There are obvious dangers in having powerful state organizations interfere with perfectly legal activities, based on decisions made in secret.
  - – And it's an extraordinarily bitter irony that Turing's work has indirectly contributed to this.