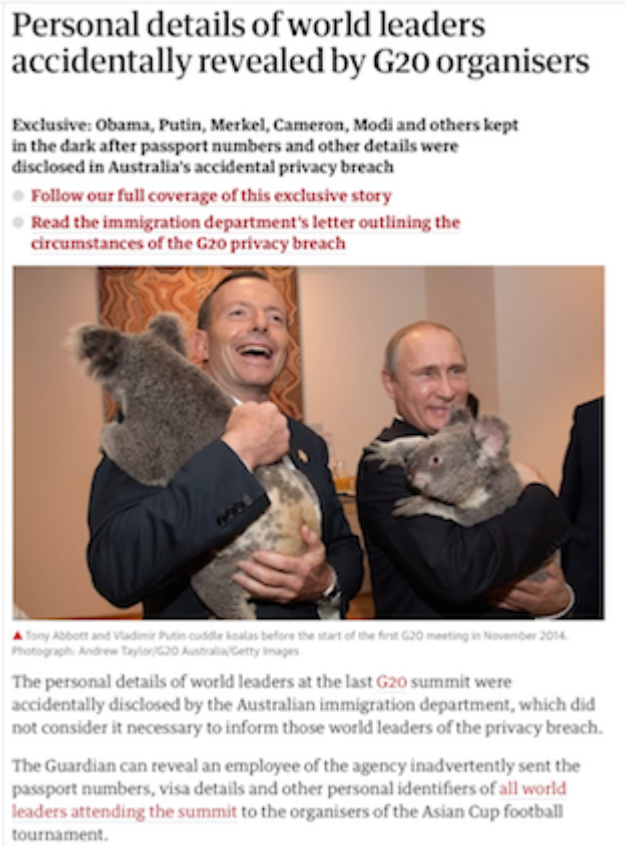
Data Ethics

Major incidents:

- TJ Maxx credit/debit card theft (2005/07)
  - Hackers gained access to accounts of over 100 million customers

    ⇒ Customers exposed to credit/debit card fraud ⇒ "You know, these are going to be sold off for a period of time in the future"
- Yahoo! data breach (2013-16)
  - Hackers gained access to all 3 billion user accounts
  - Details taken included names, DOBs, passwords, answers to security questions
    ⇒ Customers exposed to identity theft
    ⇒ Over 20 class-action lawsuits filed against Yahoo!
- Facebook-Cambridge Analytica data scandal (2018)
  - Millions of people's Facebook profiles used for political purpose without their consent
  - Cambridge Analytica was hired by President Trump's 2016 election campaign
    - gained access to data on 50 million Facebook users
      - aim: identify and manipulate their voting behavior.

    ⇒ Cambridge Analytica went bust as a consequence
- G20 world leaders data leak
  - an employee of the Aust. Immigration Department accidentally sent personal details (e.g. passport, visa) of all world leaders attending the summit to the organisers of the Asian Cup football tournament



Personal details of world leaders accidentally revealed by G20 organisers

Exclusive: Obama, Putin, Merkel, Cameron, Modi and others kept in the dark after passport numbers and other details were disclosed in Australia's accidental privacy breach

- Follow our full coverage of this exclusive story
- Read the immigration department's letter outlining the circumstances of the G20 privacy breach

▲ Tony Abbott and Vladimir Putin cuddle koalas before the start of the first G20 meeting in November 2014. Photograph: Andrew Taylor/G20 Australia/Getty Images

The personal details of world leaders at the last G20 summit were accidentally disclosed by the Australian immigration department, which did not consider it necessary to inform those world leaders of the privacy breach.

The Guardian can reveal an employee of the agency inadvertently sent the passport numbers, visa details and other personal identifiers of all world leaders attending the summit to the organisers of the Asian Cup football tournament.

Australian Privacy Act 1988

- outlines how personal information must be used/managed
- applies to government agencies, businesses and organisations with annual turnover of >$3 million, private health services, …
- *Individuals* have the right to:
    i. have access to their personal information
    ii. know why and how information is collected and who it will be disclosed to
    iii. ask to stop unwanted direct marketing
- *Businesses and organisations* must comply with the Australian Privacy Principles:
    i. how to collect personal information
    ii. how (not) to use personal information
    iii. how to secure personal information

In the event of a suspected or known data breach …

- contain breach where possible
- assess if personal information is likely to result in serious harm to affected individuals
    - individuals must be notified
    - Australian Information Commissioner must also be notified
- take action to prevent future breaches

Algorithm Ethics

Algorithms can be critical to life.

Uberlingen aircraft collision 1/7/02 at 11:35pm …

1. passenger jet A and cargo jet B on collision course at 36,000 feet
2. ground air traffic controller instructed **jet A pilot to descend**
3. seconds later, the automatic Traffic Collision Avoidance System (TCAS)
    - instructed **jet A to climb**
    - instructed **jet B to descend**
4. jet B's pilot followed TCAS, jet A's pilot followed the ATC instruction
5. all 71 people on board the two planes killed
   ⇒ Collision would not have occurred had both pilots followed TCAS

TCAS …

- builds 3D map of aircraft in the airspace
- determines if collision threat occurs
- automatically negotiates mutual avoidance manoeuvre
- gives synthesised voice instructions to pilots ("climb, climb")

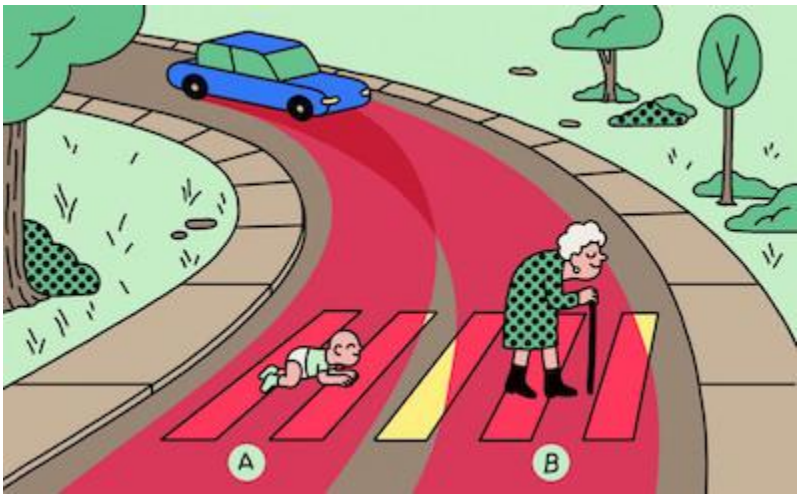
Autonomous automobiles …

- potential crash scenarios
    - if you have to choose between two actions, both harmful
        - do you choose the 'least harmful' action?
            - *how do you assess this?*
                - count the worst-case number of fatalities
                    - ... do you ignore age, ...?

Called the *runaway railway trolley problem* in AI

- Moral Machine Experiment, MIT Media Lab, 2014

    - example: the dilemma:
        - should a self-driving car continue straight ahead and possibly kill an elderly pedestrians, or
        - swerve into a barricade and possibly kill a crawling baby



- experiment open to public: millions played the game, 233 countries
- culture/country important factor, e.g.
    - 'individualistic' countries (Western democracies) often chose to spare the young
    - China/Japan/Korea were more even-handed

But it is very complicated:

- e.g. if the dilemma is choose between *spare your passenger and spare a pedestrian*
    - Japan: spare the pedestrian (strong effect)
    - Australia: spare pedestrian (weak)
    - U.S.A: neutral
    - France: spare passenger (weak)
    - China: spare the passenger (strong effect)

What if the dilemma is ...

- choose to spare the driver and spare the pedestrian?
- *would that effect whether you bought the car?*

German Government publishes world's first ethical guidelines for driverless cars (Sep 2017)

- trolley problem inapplicable
    - *All humans are considered equal for the purposes of harm minimisation.*
        - no discrimination between potential victims allowed
- must always be a human taking ultimate reponsibility

ACM/IEEE Software Engineering Code

Software engineers:

- shall ensure that their products meet the **highest professional standards** possible
- strive to fully **understand the specifications** for software
- ensure that specifications have been **well documented** and **satisfy users' requirements**
- ensure adequate **testing, debugging**, and review of software and related documents
- approve software only if it
    - is **safe**
    - **meets specifications**

- passes appropriate **tests**
- **does not diminish quality of life**, diminish privacy or harm the environment

How does this course contribute to this?

To become a competent Computer Scientists, you must be able to:

- choose effective **data structures**
- choose **algorithms** on these data structures
- analyse **performance** characteristics of algorithms (time/space complexity)
- package data structures & algorithms into an **abstract data type**
- implement data structures and algorithms using the **programming language C**

Examination

Approximately 50% code, 50% 'theory' (plus/minus 5%)

- 'theory' means provide a short explanation, provide a definition, show an example

Examinable material:

- Quiz & Assignments
- Quacks
- Linked Lists
- Program Analysis
- Heap Data Structures
- Priority Queues
- Graphs
- Weighted Graphs
- BSTs
- Splay Trees

In the code questions:

- you may need to read from *stdin* or from the command line
- if I give you an ADT, then I will ask you to:
  - write a missing function for the ADT,
  - or write a client that uses the ADT (e.g. quack question in mid-term)

Epilogue (2019-08-08 17:56:54由AlbertNymeyer编辑)