COMP 516
Research Methods in Computer Science

Dominik Wojtczak

Department of Computer Science
University of Liverpool

COMP 516
Research Methods in Computer Science
Lecture 9: Research Process Models

Dominik Wojtczak

Department of Computer Science
University of Liverpool

Research Process Models

All definitions agree that research involves a systematic or methodical process

Dawson (2005), following Baxter (2001), identifies four common views of the research process:

- Sequential
- Generalised
- Circulatory
- Evolutionary

Research Process Models: Sequential (1)

Research process as

- Series of activities
- Performed one after another (sequentially)
- In a fixed, linear series of stages

Example:
Research process model of Greenfield (1996):

1. Review the field
2. Build a theory
3. Test the theory
4. Reflect and integrate

Research Process Models: Sequential (2)

Example:
Sharp et al (2002):

1. Identify the broad area of study
2. Select a research topic
3. Decide on an approach
4. Plan how you will perform the research
5. Gather data and information
6. Analyse and interpret these data
7. Present the result and findings

Research Process Models: Sequential (3)

Greenfield (1996):

1. Review the field
2. Build a theory
3. Test the theory
4. Reflect and integrate

Sharp et al (2002):

1. Identify the broad area of study
2. Select a research topic
3. Decide on an approach
4. Plan how you will perform the research
5. Gather data and information
6. Analyse and interpret these data
7. Present the result and findings

What do you think about this research process model?
What is wrong with it?

Research process models: Sequential (4)

Greenfield (1996):

1. Review the field
2. Build a theory
3. Test the theory
4. Reflect and integrate

Sharp et al (2002):

1. Identify the broad area of study
2. Select a research topic
3. Decide on an approach
4. Plan how you will perform the research
5. Gather data and information
6. Analyse and interpret these data
7. Present the result and findings

Problems with the sequential (and generalised) process model:

1. Stages not subject specific
2. No repetition or cycles
3. Starting point and order fixed

Research Process Models: Generalised (1)

- The generalised research process model recognises that the stages of the research process depend on the subject and nature of the research undertaken

  Example:
  Data gathering and data analysis play no role for research in pure mathematics and large parts of computer science
  Instead researchers make conjectures which they prove mathematically

- The generalised research process model provides alternative routes depending on the subject and nature of the research undertaken
- But each route is still sequential

Research Process Models: Generalised (2)

Example:

(1) Identify the broad area of study
(2) Select a research topic

In natural sciences:
(3) Decide on an approach
(4) Plan the research
(5) Gather data and information
(6) Analyse and interpret these data

In mathematics:
(3’) Make a conjecture
(4’) Prove the conjecture

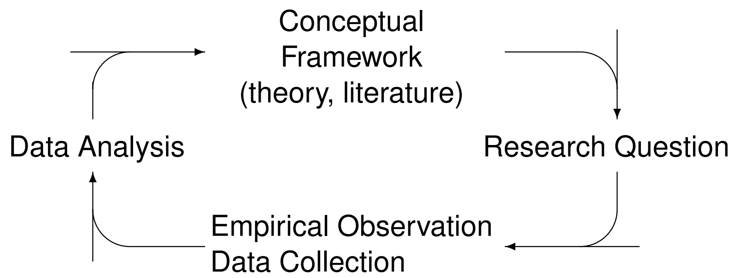
(7) Present the result and findings

Problems with the generalised process model:

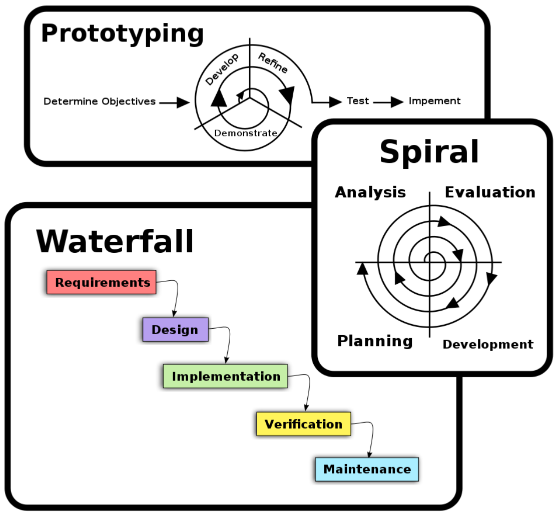
1. No repetition or cycles
2. Starting point and order fixed

Research Process Models: Circulatory

- The circulatory research process model recognises that any research is part of a continuous cycle of discovery and investigation that never ends
- It allows the research process to be joined at any point
- One can also revisit (go back to) earlier stages



Analogy to Software Development Patterns



Research Process Models: Evolutionary (1)

- The evolutionary research process model recognises that research (methods) itself evolve and change over time

  That is, over time our concept of

  - What research questions are admissible
  - What extent and methods of data collection are possible, necessary, ethical, or reliable
  - What methods are data analysis are available
  - What constitutes sufficient evidence for a hypothesis
  - What we mean by a systematic approach to research changes

Research Process Models: Evolutionary (2)

- The evolutionary research process model recognises that research (methods) itself evolve and change over time
- As an example, we can consider research in mathematics, in particular, its use of computers
- With respect to mathematical proofs we can make the following distinctions:
  (1) Proofs created solely by humans
    ⇝ typically 'sketchy', omitting steps that are considered 'obvious'
  (2) Computer-aided mathematical proofs
    ⇝ Structure and deductive steps still provided by humans, but certain computations are delegated to a computer
  (3) Fully formal, computer generated and validated proofs
    ⇝ Every step of a proof is conducted and validated by a computer, possibly under guidance by humans

Research Process Models: Evolutionary (3)

- The evolutionary research process model recognises that research (methods) itself evolve and change over time

Computer-aided mathematical proofs (1)

Four colour theorem

> *Any planar map can be coloured with at most four colours in a way that no two regions with the same colour share a border.*

Conjectured in 1852 by Guthrie. Proved in 1976 by Appel and Haken.
Proof involves a case analysis of about 10,000 cases for which the help of a computer was used

Proof seems generally accepted, but not by all mathematicians

Research Process Models: Evolutionary (4)

- The evolutionary research process model recognises that research (methods) itself evolve and change over time

Computer-aided mathematical proofs (2)

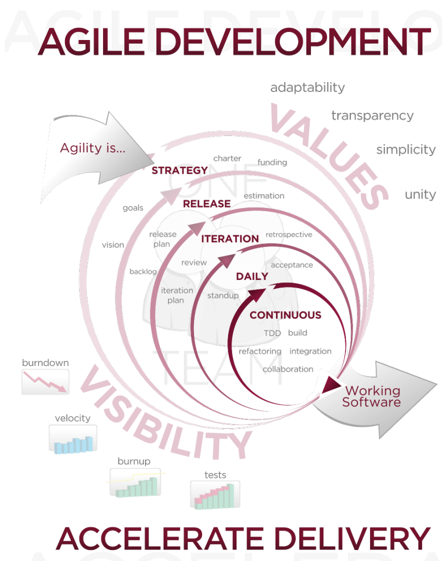
Sphere packing theorem

> *Close packing is the densest possible sphere packing.*

Conjectured in 1611 by Kepler. Hayes published a proof plan in (1997).
Execution of the plan involved solving about 100,000 linear optimisation problems using a computer. The computer files for the related programs and data requires more than 3GB of space

At one point it was suggested that the proof will be published with a disclaimer, saying that it is impossible for a human to check its correctness

Analogy to Software Development Patterns (2)



Research Process Models: Conclusion

- Among the four common views of the research process
  - Sequential
  - Generalised
  - Circulatory
  - Evolutionary

  the evolutionary research process model best describes the 'real' research process
- While the evolutionary research process model allows for the 'rules of the game' to change over time, this does not imply there aren't any rules
- For a young researcher it is best to follow the current established research process

Scientific Method

- Scientists use observations and reasoning to develop technologies and propose explanations for natural phenomena in the form of hypotheses
- Predictions from these hypotheses are tested by experiment and further technologies developed
- Any hypothesis which is cogent enough to make predictions can then be tested reproducibly in this way
- Once it has been established that a hypothesis is sound, it becomes a theory.
- Sometimes scientific development takes place differently with a theory first being developed gaining support on the basis of its logic and principles

Elements of a Scientific Method

The essential elements of a scientific method are iterations, recursions, interleavings and orderings of the following:

- Characterisations
  (Quantifications, observations and measurements)
- Hypotheses
  (theoretical, hypothetical explanations of observations and measurements)
- Predictions
  (reasoning including logical deduction from hypotheses and theories)
- Experiments
  (tests of all of the above)

Both characterisations and experiments involve data collection

Intellectual Discovery

- Knowing what the elements of a scientific method are does not tell us how to come up with the right instances of these elements
  - What predictions does a theory make?
  - What is the right hypothesis in a particular situation?
  - What is the right experiment to conduct?
- These are commonly derived by a process involving
  - Deductive reasoning
  - Abductive reasoning
  - Inductive reasoning

  Classification by Charles Sanders Peirce (1839-1914)
  See `http://plato.stanford.edu/entries/peirce/` for additional details

Intellectual Discovery: Deduction (1)

- Deductive reasoning proceeds from our knowledge of the world (theories) and predicts 'likely' observations

  Example:

  – Assume we know that A implies B.
  – A has been observed.
  – Then we should also obverse B.

- Useful for experiment generation for theories

  Example:

  Newton's theory of gravity versus Einstein's theory of relativity

  - Largely make the same predictions
  - Both predict that the sun's gravity should bend rays of light
  - However, Einstein's theory predicts a greater deflection
  - Correctness of Einstein's prediction confirmed by observation in 1919

Intellectual Discovery: Deduction (2)

- Deductive reasoning is often said not to lead to new knowledge
  (Note: This implies pure mathematicians largely waste their time)

  ⇝ Seriously underestimates the computational effort involved in deductive reasoning

  ⇝ Most theories are undecidable
  (There is no algorithm that even given infinite time could determine whether a statements follows from a theory or not)

  ⇝ Thus, establishing that a statement follows from a theory extends our knowledge

Intellectual Discovery: Abduction

- Abductive reasoning proceeds from observations to causes

  Example:

  – The phenomenon X is observed.
  – Among hypotheses A, B, C, and D,
  only A and B are capable of explaining X.
  – Hence, there is a reason to assume that A or B holds.

  ⇝ Requires a theory linking A, B, C, D to X

- Useful for hypothesis generation
- Hypotheses must then be confirmed / eliminated through further observation
- It is not easy from the outside to decide whether someone uses deduction or abduction
  ⇝ The two are often confused

Intellectual Discovery: Induction (1)

- Inductive reasoning proceeds from a set of observations to a general conclusion

  Example:

  – Tycho Brahe, a 16th century astronomer, collected data on the movement of the Mars.
  – Johannes Kepler analysed that data which was consistent with Mars moving in an elliptic orbit around the sun.
  – Inductive conclusion:
  Mars, and all other planets, move in elliptic orbits around the Sun, with the Sun at one of the focal points of the ellipse.

- Primary tool for theory formation

Intellectual Discovery: Induction (2)

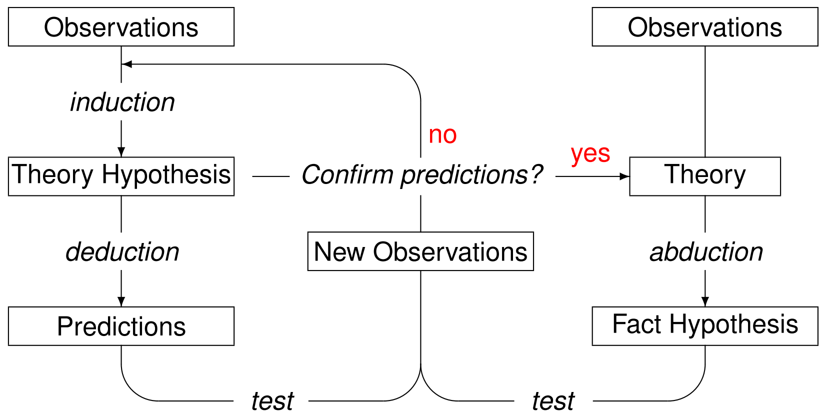
- An incomplete set of observations can easily lead to incorrect inductive conclusions

  Example:

  – All swans I've ever seen are white

  – Inductive conclusion: All swans are white

Scientific Method: A Model



Intellectual Discovery: Problems

- Deductive reasoning tells us that from 'A' and 'A implies B' we can conclude 'B'
  However, it cannot tell us whether 'A' or 'A implies B' holds, nor whether 'B' is what we want to show

- Abductive reasoning tells us that from 'B' and 'A implies B' we may conclude 'A'
  However, it cannot tell us whether 'B' or 'A implies B' hold, nor how to establish that 'A' is the case

- Inductive reasoning tells us that from '$A(o_1)$', ..., '$A(o_n)$' and '$B(o_1)$', ..., '$B(o_n)$' we may conclude '$\forall x. A(x) \Rightarrow B(x)$'.
  However, it cannot tell us what the properties '$A(\_)$' and '$B(\_)$' are (nor how large the number n needs to be)

To overcome these problems we need additional techniques.

Problem Solving

- Analogy: Look for similarity between one problem and another one already solved

- Partition: Break the problem into smaller easier sub-problems

- Random/Motivated Guesses: Guess a solution then prove it correct

- Generalise: Take the essential features of the specific problem and pose a more general problem

- Particularise: Look for a special case with a narrower set of restrictions than the more general case

- Subtract: Drop some of the complicating features of the original problem

Topic submission

- submission via the VITAL system `https://vital.liv.ac.uk`, resubmit if you have already submitted
- Essay Topic assessment for COMP516 (by 19th Oct, 6pm)
- a title for your presentation, which will also be the title of your essay
- a description of the intended research, which should clearly identify a well-defined research question (but not all the details)
- this description has a limit of 500 characters
- up to five keywords that highlight some of the most important themes, concepts, and issues related to your chosen topic (e.g. look up the keywords of the journal papers on that topic)
  `https://cgi.csc.liv.ac.uk/~dominik/teaching/comp516/submit.html`

Topic choices

- the topic for your COMP516 essay can be anything that interests your and is related to CS
- alternatively, pick some topic listed at the COMP516 webpage
- do not try to solve an open-problem as a topic (keep it for your MSc project)
- another possibility is to pick as your essay topic an MSc project was not picked last year
- `https://cgi.csc.liv.ac.uk/~comp702/` and use your CS login/password (not MWS) ....
- I will look through the topics and give feedback via VITAL
- submit your topic as soon as possible, the sooner you submit the sooner you will get feedback