Methoden Wetenschappelijk Onderzoek

Literature research
Today’s lecture

- To explain **how to use** scientific literature

- To introduce **assignment one**
  - A mini-literature study
Today’s specific questions

- How does scientific referencing work
- How does one refer to literature in a paper
- How does one make a reference list
- How does one do literature research
What is literature for?

• To get ideas
  – What are the hot topics?
  – What are the open questions?

• To get information
  – Background
  – Data
  – Techniques
Why literature references?

- To refer to work you used
  - To prevent repetition of questions/methods/results
  - To give credit to others
  - To avoid accusation of plagiarism
  - To make your work replicable and verifiable
- To show you are knowledgeable
  - Situate your work
    - Work leading up to it
    - Work comparable to it
    - Classics in your field (?)
A note on conventions

- Older papers generally use far fewer references
  - Turing 1952: 36 pages, 6 references
  - Shannon 1948: 54 pages, 7 references
  - Modern: at least 20–30 references
  - Perhaps related to easier access nowadays

- Different fields have different traditions
  - Engineering fields traditionally use few references
Where to find which research

• **Workshops**
  – Work in progress, discussion of ideas

• **Conferences** - proceedings
  – Work in progress, but more finished

• **Reports**
  – Finished work, to claim precedence

• **Journals**
  – Finished work, for dissemination

• **Books**
  – Large bodies of work, collections, textbooks, handbooks
Where do scientists publish?

- Different fields have different traditions
  - Computer science: many conferences
  - Linguistics: books have high status
  - Physics: preprint archives are important
  - Genetics: no discussion of results before journal publication

- Learn the traditions of your field!
Where do scientists publish?

• Because of online availability, and good quality control, journals are becoming the standard
  – But a journal is not a journal!

• Descending order of prestige:
  – Science, Nature
  – PNAS (Post Nature And Science)
  – Top journals in your field
    (examples: Cell, language, IEEE journals, JASA)
  – Less respected international journals
  – National journals
Impact factors (1)

- IEEE Transactions on Evolutionary Computation 4.4
- Evolutionary Computation 2.6
- Artificial Intelligence 2.5
- Artificial Life 2.1
- Journal of Artificial Intelligence Research 1.7
- Genetic Programming and Evolvable Machines 1.2
- AI Communications 0.8
- International Arab Journal of Information Technology 0.065
Impact factors (2)

• Average number of citations per article during the two preceding years
  – Calculated by ISI (The Institute for Scientific Information)
  – For journals indexed in Thomson Reuters’ Journal citation reports: www.webofknowledge.com
  – These are commercial firms!

• Problematic in many ways, but still a useful indication
  – But note: popularity attracts popularity!
Impact Factors (3)

• Journals with higher impact factors are more difficult to get into

• Scientists are partly judged by the impact factors of the journals in which they publish

• Hence, everyone wants to publish in journals with high impact factors…
References in papers (1)

• Scientific paper generally have a very predictable structure
  – Introduction
  – Methods
  – Results
  – Conclusion
  – Discussion

References to background, reasons for doing research, previous efforts
Methodological references
Fewer references, related to analysis methods etc.
Related work, impacted by this research, ongoing work
The mechanisms involved in the acquisition and processing of language are closely intertwined with the structure of language itself. Children routinely acquire language with little intentional tutoring by their parents and as adults use language with minimal effort. Indeed, our unique and nearly universal capacity to acquire and use language has even been cited as one of eight key transitions in the evolution of life (1). These features of species specificity and species universality, combined with the intimate fit between language structure and the mechanisms by which language is acquired and used, point to substantial genetic constraints. The nature and origin of the genetic basis for language remain the focus of much debate, however (2–4).

An influential line of thinking in the cognitive sciences suggests that the genes involved in language predetermine a highly specialized and species-specific language “module” (5), “instinct” (6), or “organ” (7). This module has been assumed to specify a number of domain-specific linguistic properties, including case marking, agreement, and conformity to highly abstract syntactic constraints, such as X-bar theory (8). Although some have argued that the genes encoding a language module arose through a sudden “catastrophic” genetic change (9), and others have remained agnostic on this point (10), “the default prediction from a Darwinian perspective on human psychological abilities” (ref. 11; p. 16) is the adaptationist view, that genes for language coevolved with human language itself for the purpose of communication (1, 8, 12–18).

To what extent are human behaviours a straightforward reflection of the underlying psychological characteristics of the individual? This is a key question in the cognitive sciences, and is central to the debate in linguistics over the relationship between the observed typological distribution of languages and psychological constraints on language acquisition (see e.g. Chomsky, 1965; Christiansen & Chater, 2008; Evans & Levinson, 2009): are the languages we see in the world a reflection of strong or even absolute constraints on possible languages imposed during acquisition, or might they also be a consequence of the interaction of multiple weaker constraints arising from acquisition and use?

To take a specific example: one property of human language is that variation tends to be predictable. In general, no two linguistic forms will occur in precisely the same environments and perform precisely the same functions (Givón, 1985). Instead, usage of alternate forms

Smith & Wonnacott 2010, Cognition

Chater et al. 2008, PNAS
References in papers (3)

• Notes:
  – Footnotes (rare nowadays, problematic)
  – Endnotes (used where space is at a premium)
    • In order of appearance, or alphabetical order

• Name and year
  – Takes a little getting used to
  – But most convenient – name and year are meaningful
The reference list (1)

- At the end of a paper, at the end of a chapter or at the end of a book

References


Chater et al. 2008, PNAS

Smith & Wonnacott 2010, Cognition
The reference list (2)

• The reference list contains:
  – enough information to locate the source
    • Name journal, book etc.
    • Issue number, volume etc.
    • Page numbers
  – the authors’ names
    • To give credit
  – the year
    • For establishing precedence
The reference list (3)

- Different information required for different types of sources
- Different formats used by different journals
  - Specified on journal web site (info for authors)
  - Compromises between space and information
  - But unfortunately diverse
- There are very useful tools to handle this:
  - EndNote
  - Reference Manager
  - BibTex
  - Zotero


Journal abbreviations

• In many styles, journal names are abbreviated
  – To save space

• There are lists for this!
  – http://library.caltech.edu/reference/abbreviations/
  – (and many others)

• Problems occur for older journals
  – Biol. Gabon. (= Biologia Gabonica)

14. Why women speak better than men (and its significance for evolution)

BART DE BOER

14.1 Introduction

One of the many striking differences between humans and our closest primate relatives is the shape of the vocal tract. Whereas orangutans, bonobos, chimpanzees, and gorillas all have similar vocal tracts with low flat tongues and high larynges, humans have a round tongue and a larynx that is positioned low in the throat (Fitch 2006; Negus 1949). The combination of a low larynx and a wide gap between the larynx and the velum makes it impossible for (adult) humans to swallow and breathe simultaneously. It is generally assumed (e.g. Fitch and Giedd 1999, section III C) that this increases the probability of choking on one's food. Even if the resulting reduction of fitness is not enormous, it must nevertheless be explained evolutionarily why the difference between humans and other higher primates has evolved.

One explanation that is sometimes mentioned (e.g. Pollick and de Waal 2007) is that the lowered larynx does not have an adaptive function, but that it is a direct side effect of bipedalism. This explanation is not really supported by evidence, and there is in fact some evidence that indicates that bipedalism does not influence the position of the larynx. As for the lack of positive evidence, the original sources that are usually referred to in this context (Aiello 1996; Dufresne 1958) do not really make a direct link between bipedalism and the anatomy of the human vocal tract. They only state that bipedalism removed some constraints on the function of the mouth and the larynx. The disappearance of these constraints allowed different functions of the larynx and vocal tract to evolve. However, these sources do not propose that the development of the human larynx did not serve an adaptive purpose.
References to a book


On author lists

• Different fields have different rules

• Alphabetical order
  – But not really fair to names with Z

• In order of contribution
  – First author is most important
  – Last author is supervisor

• In some fields author lists can be extreme
  (100 times as many authors as pages)
How to find literature (1)

- Scientists tend to be specialists
  - Huge knowledge about one’s own field
  - Acquired through years of interacting with colleagues and reading

- But what if you want to explore a new field?
  - Ask a colleague
  - But don’t be lazy!
Reasons for finding literature

- Learn new skills
- Get new ideas
- Find methods
- Find data
- Find older work

- Different reasons lead to different focus
How to find literature (2)

• Wikipedia!
  – This can lead to disappointment
  – Quality is very variable
  – Not necessarily representative
    • Used for “plugging” research

• Google!
  – Same problems
  – And which keywords should one use?
How to find literature (3)

• **Textbooks**
  – Russell & Norvig (Introduction to AI) contains loads of references
  – **Quality** is OK
  – But may not be entirely **representative**
  – Use **multiple** textbooks

• This can be the endpoint if your goal is to learn new **skills**
How to find literature (4)

• Conference proceedings
  – An overview of work that is state-of-the-art
  – As with journals, some conferences are more prestigious than others
    • Acceptance rate is sometimes given
  – Most conferences are peer-reviewed

• Different conferences have different foci
  – But the call for papers makes this clear

• Perusing conference proceedings gives you an idea of the players in a field
  – Who chairs sessions?
  – Who is on the program committee?
  – Who are the plenary speakers?
How to find literature (5)

• Peruse recent issues of the major journals
  – Journal publications represent more mature work

• Some journals are specialized in reviews
  – Trends in Cognitive Sciences
  – Annual Review of Computer Science

• But: it may be difficult to find relevant articles

• Fortunately, journals have search functions

• Or use RSS-feeds for when new issues appear
How to find literature (6)

• Read *introductory sections* of papers
  – (or *methods sections*)
  – To whom do they refer, and in what context?

• A great way to find *review articles*
• And to have *context* with references

• However, this only takes you *back in time!*
How to find literature (7)

• Citation analysis
  – Citeseer [http://citeseerx.ist.psu.edu/](http://citeseerx.ist.psu.edu/)
  – ISIWeb [http://apps.webofknowledge.com](http://apps.webofknowledge.com)
  – Google scholar [http://scholar.google.nl/](http://scholar.google.nl/)

• Finds articles that refer to a given article
  – Takes you forward in time

• Allows you to find oft-cited articles
  – Apparently important (?)
How to find literature (8)

• Citation analysis allows you to make a network analysis of scientific literature
  – Central articles within a field are important

• Not necessarily the same as many references
  – These may be from outsiders
  – Self-amplifying effect if everybody uses google scholar
Now you will have to use these skills!
The assignment (1)

• A short literature review on a given topic
  – Suggestions follow, but may choose your own

• Maximally three pages
• Minimally 30 references (1-2 pages)
• Max. 1 page in which you refer to the papers and briefly explain why they are important

• Answer the following questions:
  – Who appear to be the important players?
  – What are the hot topics?
  – Is your chosen field an active field?

• Refer with name and year, choose and specify style
  – Recommendation: use a reference management tool
The assignment (2)

• Work in groups of two

• Deadline: October 25
  – Each day late = 1 point less

• Please send pdf
  – b.g.deboer@uva.nl
  – Subject should contain: MWO 1
Suggested topics

- Topics of the AAAI 2010 conference call for papers
  - Agents
  - Cognitive modeling and human interaction
  - Commonsense reasoning
  - Constraint satisfaction and optimization
  - Evolutionary computation
  - Game playing and interactive entertainment
  - Information integration and extraction
  - Knowledge acquisition and ontologies
  - Knowledge representation and reasoning
  - Machine learning and data mining
  - Model-based systems
  - Multiagent systems
  - Natural language processing
  - Planning and scheduling
  - Probabilistic reasoning
  - Robotics
  - Search

- Choose a subtopic of one of these topics
  - Evolutionary computation: multi-objective GA’s
  - Multiagent systems – ant systems
  - Robotics – robot navigation

- Or choose your own favorite topic
  - Doesn’t even have to be AI
In a moment...

• Collecting data (part 1)
  – Ways in which different branches of science increase knowledge of the world

• A second step into being able to set up your own research project