Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

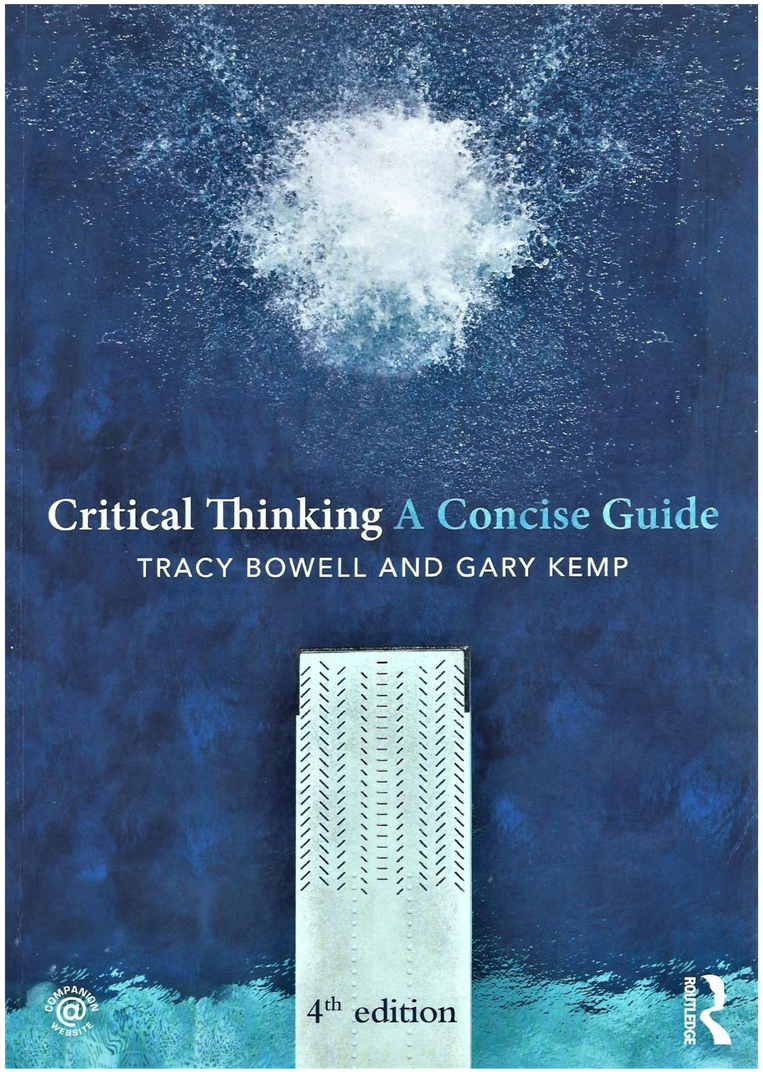
Precise Thinking 3.1
Defeasible Arguments

Basics and Motivation

Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

PRECISE THINKING



We *very* loosely follow this book:

> Bowell, T., & Kemp, G. (2015). *Critical thinking: A concise guide (4th ed.)*. Routledge.

However, *lots* of things deviate from the book. What is said in the lecture has precedence over the book.

You do **not** have to buy the book. If you want to have an inexpensive look, you can find a digital copy of the 2nd edition online at:

http://www.academia.edu/download/46383480/_Tracy_Bowell__Critical_Thinking_A_Concise_Guide_BookFi.org.pdf

MORE THAN VALIDITY?

P1:	There are very few right and millions of wrong combinations in lotto.
P2:	In order to win money in lotto, you need a right combination.
C:	You won't win money in lotto. ✗|✓

I have to get up at 6:00 am tomorrow, and set my alarm clock to 5:50 am. I never snooze for longer than 5 minutes and my alarm clock works very reliably. So, I will wake up in time tomorrow.

Yes, you for sure *can* try to cure Covid-19 with herbal essences, but doing that is like trying to fix your broken car by giving it a big hug – maybe it gives you a good feeling, but it won't really work.

P1:	The patient is coughing, has a sore throat and is overall tired.
P2:	Before getting the above symptoms quite abruptly, the patient was healthy.
C:	The patient has ~~a cold~~ Covid-19. ✗|✓

P1	The sun rose every morning so far.
C:	The sun is going to rise tomorrow. ✗|✓

DEFEASIBLE REASONING

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/entries/reasoning-defeasible/

Defeasible reasoning

Reasoning is defeasible iff the corresponding argument is rationally compelling but not deductively valid.

P1: Tim is a bird.

C: Tim can fly. ✗|✓

P1: Tim is a bird.

P2: Tim is a penguin.

C: Tim cannot fly. ✗|✓

Defeasible reasoning

P1: Anne is Bob's grandmother.

C: Bob loves Anne. ✗|✓

Not truth-preserving: there are cases in which all premises are true, but the conclusion is false.

VS

Deductive reasoning

P1: Anne is Bob's grandmother.

C: Bob is Anne's grandchild. ✓| -

Truth-preserving: there are no cases in which all premises are true, but the conclusion is false.

DEFEASIBLE REASONING

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/entries/reasoning-defeasible/

P1: The patient is coughing, has a sore throat and is overall tired.

P2: Before getting the above symptoms quite abruptly, the patient was healthy.

C: The patient has ~~a cold~~. covid-19 ✗ | ✓

Defeasible reasoning

P1: Anne is Bob's grandmother.

C: Bob loves Anne. ✗ | ✓

Not truth-preserving: there are cases in which all premises are true, but the conclusion is false.

VS

Deductive reasoning

P1: Anne is Bob's grandmother.

C: Bob is Anne's grandchild. ✓ | -

Truth-preserving: there are no cases in which all premises are true, but the conclusion is false.

DEFEASIBLE REASONING

Example

P1: My neighbour told me that he just bought a new Mercedes.

P2: There is a brand-new Mercedes in front of his house.

C: The Mercedes in front of my neighbour's house is his. ✗ | ✓

DEFEASIBLE REASONING

Example

P1: My neighbour told me that he just bought a new Mercedes.

P2: There is a brand-new Mercedes in front of his house.

P3: His sister, who is usually a very reliable source of information, says that he just brags about it everywhere, but that he did not actually buy a new car and that he could not afford it.

C: The Mercedes in front of my neighbour's house is not his. ✗ | ✓

DEFEASIBLE REASONING

Example

P1: My neighbour told me that he just bought a new Mercedes.

P2: There is a brand-new Mercedes in front of his house.

P3: His sister, who is usually a very reliable source of information, says that he just brags about it everywhere, but that he did not actually buy a new car and that he could not afford it.

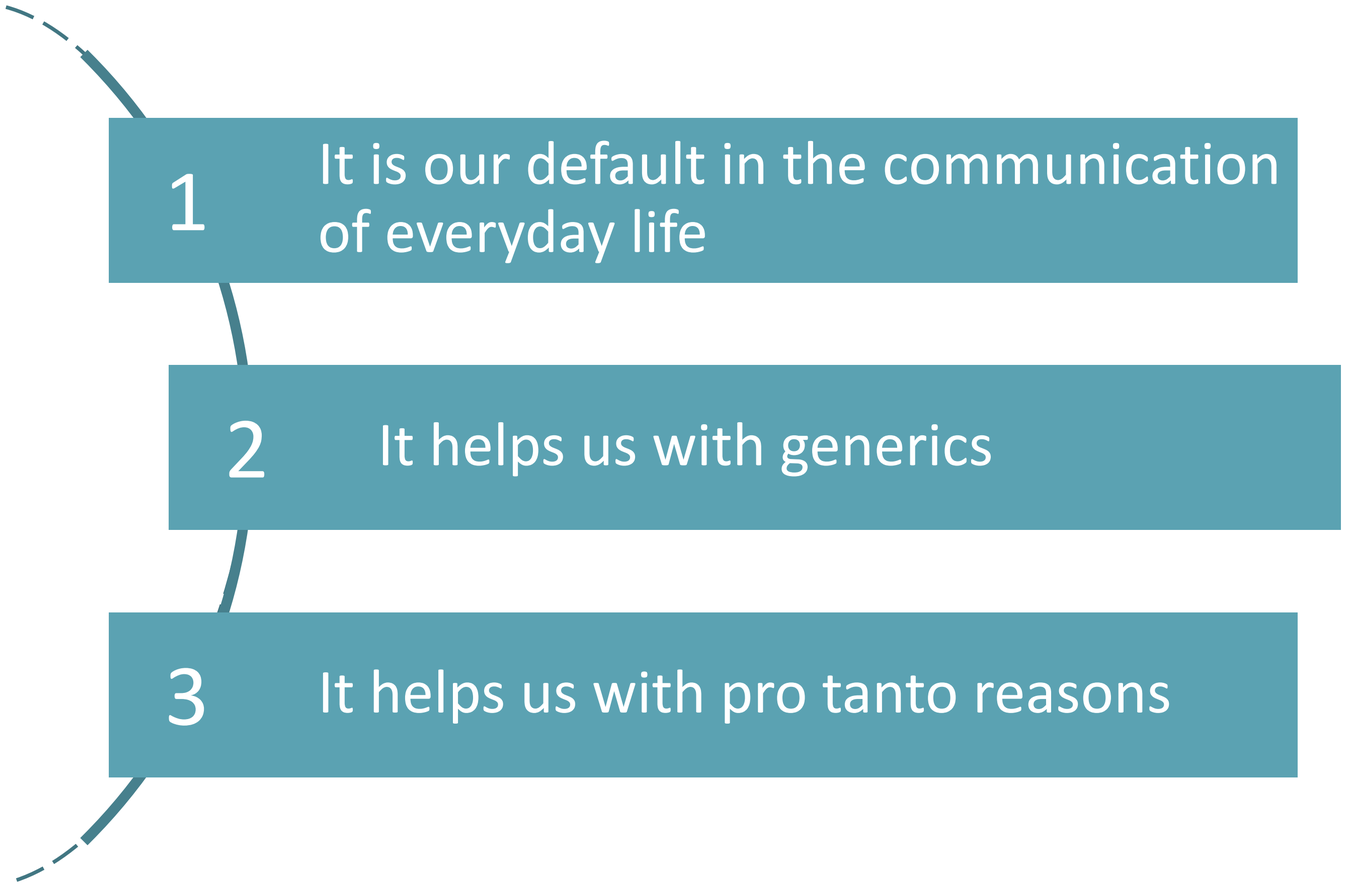
P4: His wife, who has no incentive of lying, says that his sister is always jealous and often tells lies about the family.

C: The Mercedes in front of my neighbour's house is his. ✗ | ✓

WHY DEFEASIBLE REASONING?

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/.

Reasons for defeasible reasoning (selection)



WHY DEFEASIBLE REASONING?

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/.

1. It is our default in the communication of everyday life

- We seldomly argue deductively in everyday life, because
  - this sometimes is impractical
  - most people do not know how to do this
  - often we do not have enough information to make a sound argument
- We often make assumptions based on what is and is *not* said and draw conclusions

  Example 1:

  Your doctor tells you: “Take one pill in the morning and one in the evening.”
  This leaves some logical room, e.g. for taking another pill at noon. But what he means is: “Take exactly one pill in the morning and exactly one in the evening and no third pill at any time of the day.”
- We update our conclusions if new information comes in

WHY DEFEASIBLE REASONING?

1. It is our default in the communication of everyday life

Example 2: you want to take a train from Saarbrücken to Berlin. You do not want to change trains more than two times and you want arrive before 8 pm. You can be at the train station at 11 am. Which train do you take?

Ihre Fahrtmöglichkeiten

Bahnhof/Haltestelle	Datum	Zeit	Dauer	Umst.
Saarbrücken Hbf	Do, 12.09.19	09:47	6:42	1
Berlin Hbf (tief)	Do, 12.09.19	16:29		
Saarbrücken Hbf	Do, 12.09.19	09:51	8:07	3
Berlin Hbf (tief)	Do, 12.09.19	17:58		
Saarbrücken Hbf	Do, 12.09.19	10:58	7:05	1
Berlin Hbf (tief)	Do, 12.09.19	18:03		
Saarbrücken Hbf	Do, 12.09.19	11:51	8:00	3
Berlin Hbf (tief)	Do, 12.09.19	19:51		
Saarbrücken Hbf	Do, 12.09.19	12:47	7:10	1
Berlin Hbf (tief)	Do, 12.09.19	19:57		
Saarbrücken Hbf	Do, 12.09.19	13:47	6:39	2
Berlin Hbf (tief)	Do, 12.09.19	20:26		

1 Erwachsener, 2. Klasse

generated on www.bahn.de

departs before 11 am
more than two changes
arrives after 8 pm

so will you take this one?
Why not this one?

Saarbrücken Hbf	Do, 12.09.19	11:47	6:42	1
Berlin Hbf (tief)	Do, 12.09.19	18:29		

Because it wasn't in the list!

WHY DEFEASIBLE REASONING?

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/.

2. Helps us with generics

pexels.com

Is the poem wrong?

WHY DEFEASIBLE REASONING?

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/.

2. Helps us with generics

Defeasible reasoning makes it easier for us to handle generics:

"roses are red"	"violets are blue"	"birds can fly"	"children like sweets"
shorthand for	shorthand for	shorthand for	shorthand for
"**typically**, roses are red"	"violets **normally** are blue"	"**usually**, birds can fly"	"**most** children like sweets"

P1:	Rurak is a raven.		P1:	Rurak is a raven.
P2:	Ravens can fly.	we mean →	P2:	Usually, ravens can fly.
C:	Rurak can fly. ✗ | ✓		C:	Probably, Rurak can fly. ✗ | ✓

But it still may turn out that Rurak cannot fly.

WHY DEFEASIBLE REASONING?

Koons, Robert, "Defeasible Reasoning", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2017/entries/reasoning-defeasible/.

3. Helps us with pro tanto reasons

Pro tanto reason

A reason that can be outweighed by other reasons.

Example: You promised to pick up your parents at the airport. To be there in time, you have to leave your house at 6pm. This promise gives you a pro tanto reason to actually pick up your parents.

P1:	You promised to pick up your parents at the airport in time.
P2:	To be there in time, you have to leave for the airport at 6pm.
C:	You should leave for the airport at 6pm. ✗ | ✓

P1:	You promised to pick up your parents at the airport in time.
P2:	To be there in time, you have to leave for the airport at 6pm.
P3:	Your house is on fire at 6pm.
C:	You should not leave for the airport at 6pm. ✗ | ✓

Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

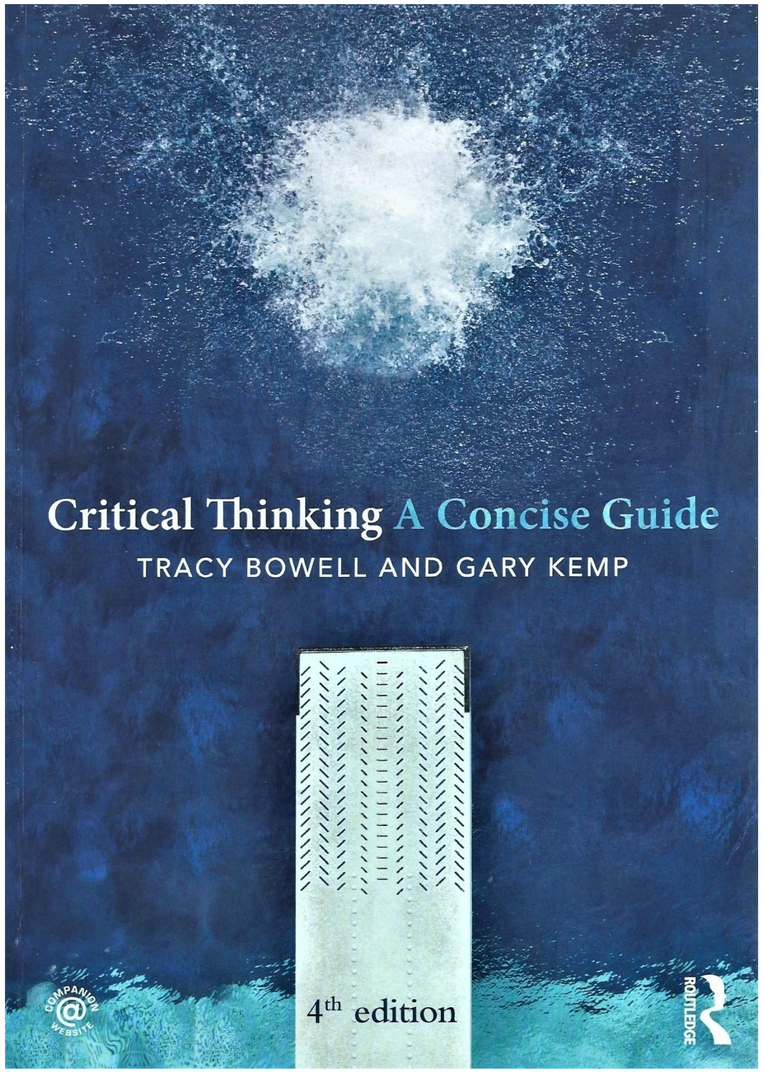
Precise Thinking 3.2
Defeasible Arguments

Kinds of Defeasible Reasoning

Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

PRECISE THINKING



We *very* loosely follow this book:

> Bowell, T., & Kemp, G. (2015). *Critical thinking: A concise guide (4th ed.)*. Routledge.

However, *lots* of things deviate from the book. What is said in the lecture has precedence over the book.

You do **not** have to buy the book. If you want to have an inexpensive look, you can find a digital copy of the 2nd edition online at:

http://www.academia.edu/download/46383480/_Tracy_Bowell__Critical_Thinking_A_Concise_Guide_BookFi.org.pdf

COMMON KINDS OF DEFEASIBLE REASONING

Common kinds of defeasible reasoning

(general) defeasible arguments

inductive arguments

abductive arguments

analogies

ABDUCTIVE ARGUMENTS

Abductive Arguments
you have an observation and try to find the best explanation

P1:	The patient has a runny nose, a headache, and is overall tired.
P2:	Before getting the above symptoms quite abruptly, the patient was healthy.
P3:	It is winter.
C:	The best explanation is: The patient has a cold. ✗ | ✓

Examples: Diagnostic systems (symptoms explained by illness),
image recognition (pattern of pixels explained by object in the image)

Keyword: “The best explanation is:”

ABDUCTIVE ARGUMENTS

Sherlock Holmes – the master of ~~deduction~~ abduction

https://www.bbc.co.uk/programmes/b018ttws

COMMON KINDS OF DEFEASIBLE REASONING

Common kinds of defeasible reasoning

(general) defeasible arguments

inductive arguments

abductive arguments ✓

analogies

ANALOGIES

Analogy
you argue for a case by pointing to similarities to another case

(often more rhetoric than a form of argument, however it is often used in an argument-like way)

General form:

P1:	X has property P.
P2:	X and Y are similar in relevant aspects.
P3	P and R are similar in relevant aspects.
C:	Y has property R.

ANALOGIES

Example

"... robots aren't merely replacing human drivers, just as human drivers in the first automobiles weren't simply replacing horses: The impact of automating transportation will change society in radical ways, and ethics can help guide it."

(from Patrick Lin (2013). The Ethics of Autonomous Cars. Newspaper article in *The Atlantic.* http://www.theatlantic.com/technology/archive/2013/10/the-ethics-of-autonomous-cars/280360/)

X: the process of human drivers replacing horses

Y: the process of robots replacing human drivers

P: changed society in a radical way (and ethics could help guide it?)

R: will change society in a radical way (and ethics can help guide it?)

P1:	X has property P.
P2:	X and Y are similar in relevant aspects.
P3	P and R are similar in relevant aspects.
C:	Y has property R.

Note: this analogy is flawed in several ways. → Exercises

COMMON KINDS OF DEFEASIBLE REASONING

Common kinds of defeasible reasoning

(general) defeasible arguments

inductive arguments

abductive arguments ✓

analogies ✓

DEFEASIBLE ARGUMENTS

Defeasible Arguments
reasoning to a probable conclusion

P1: John is from Scotland.

C: John has eaten Haggis before. ✗ | ✓

Not valid.

P1: John is from Scotland

P2: Everybody who is from Scotland has eaten Haggis before.

C: John has eaten Haggis before. ✓ | –

Valid, but not sound.

Not valid, but defeasibly forceful!

P1: John is from Scotland.

P2: Most of the people who are from Scotland have eaten Haggis before.

C: John has eaten Haggis before. ✗ | ✓

Defeasible Force

An argument is defeasibly forceful iff it is not deductively valid, but given the premises and only the premises, it is sufficiently probable that the conclusion is true.

Two Different Conceptions of Defeasible Arguments

Strict Conception

P1: John is from Scotland.

P2: Most of the people who are from Scotland have eaten Haggis before.

C: John has eaten Haggis before. ✗ | ✓

- You cannot take context into account to a large extent
- What makes the conditional probability large has to be explicit as a premise (see P2)

→ Connection between premises and conclusion is relatively tight

Loose Conception

P1: John is from Scotland.

C: John has eaten Haggis before. ✗ | ✓

- You can take context into account to a relatively large extent
- What makes the conditional probability large does not to be explicit as a premise

→ Connection between premises and conclusion is relatively loose

DEFEASIBLE ARGUMENTS

Note 1: Often, we cannot assign specific probabilities, but that's ok.

Phrases like "probably", "likely", "oftentimes" etc are vague

We can**not** say that something is probable iff it's probability is > 0.5 or the like.

Depending on the context, these words can mean different things, e.g.

- "You shouldn't buy a ticket for the lottery, you probably won't win anyway"
- "I'm taking the bus, so, probably I'll be late"

Note 2: Certainty and truth/falsity are two different things.

“It is **certain** that there will be schnitzel in the mensa tomorrow.”	≠	“It is **true** that there will be schnitzel in the mensa tomorrow.”
“It is **certain** that there will **not** be schnitzel in the mensa tomorrow.”	≠	“It is **false** that there will be schnitzel in the mensa tomorrow.”
“It is **neither certain** that there will be schnitzel in the mensa tomorrow, **nor** is it **certain** that there will **not** be schnitzel in the mensa tomorrow.”	≠	“It is **neither true** that there will be schnitzel in the mensa tomorrow, **nor** is it **false** that there will be schnitzel in the mensa tomorrow.”

Note 3: Probability is, naturally, a matter of degree, and so is defeasible force.

We can say that one argument is more defeasibly forceful than another.

P1: In lottery A you draw exactly one lottery ticket from a box with 1,000 tickets of which exactly one is a winning ticket.

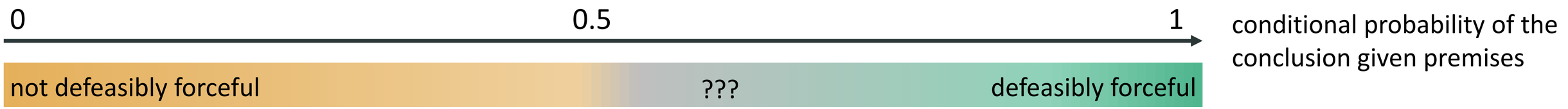
C1: You won't draw the winning ticket in lottery A. ✗|✓

P2: In lottery B you draw exactly one lottery ticket from a box with 10,000 tickets of which exactly one is a winning ticket.

C2: You won't draw the winning ticket in lottery B. ✗|✓

DEFEASIBLE ARGUMENTS

Note 4: There is no cutoff point to defeasible force. It, again, is a vague term.



DEFEASIBLE ARGUMENTS

Defeasible Cogency

An argument is defeasibly cogent iff it is defeasibly forceful and all the premises are true.

deductive arguments	defeasible arguments
(deductively) valid	defeasibly forceful
(deductively) sound	defeasibly cogent
indicator words: “therefore”, “thus”	indicator words: “probably”, “likely”, …

COMMON KINDS OF DEFEASIBLE REASONING

Common kinds of defeasible reasoning

(general) defeasible arguments ✓

inductive arguments

abductive arguments ✓

analogies ✓

DEFEASIBLE ARGUMENTS

Special case: extrapolating/inductive arguments
Generalizing from a set of observations to a general rule

P1 The sun rose every morning so far.

C: Probably, the sun is going to rise tomorrow. ✗ | ✓

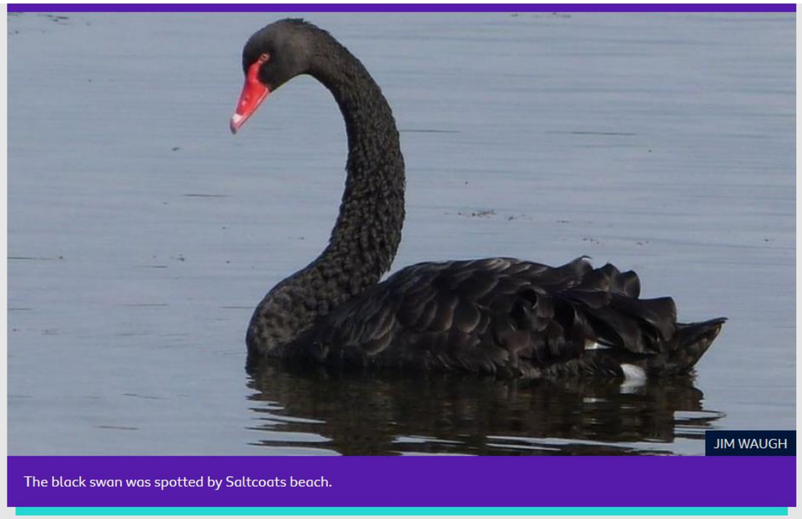
P1 All swans we saw so far where white.

P2 We saw a lot of swans.

C: Probably, all swans are white. ✗ | ✓

UK Wildlife: Australian black swan spotted in Scotland

21 May 2019 Last updated at 07:26



https://www.bbc.co.uk/newsround/48338856

We say this to be defeasible forceful if the sample is big enough to support the conclusion by ‘induction’.

COMMON KINDS OF DEFEASIBLE REASONING

Common kinds of defeasible reasoning

(general) defeasible arguments ✓

inductive arguments ✓

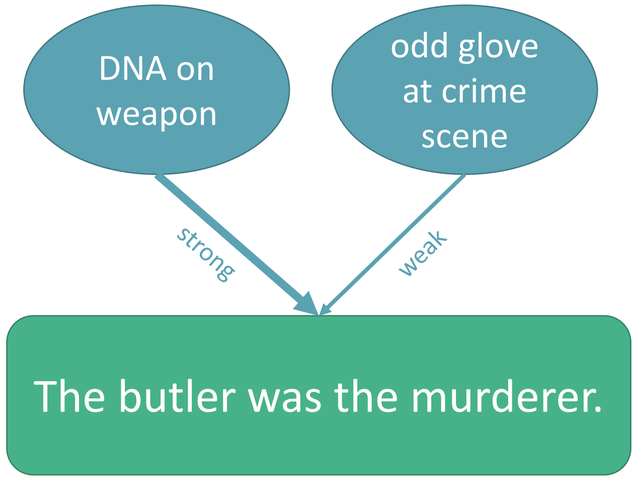
abductive arguments ✓

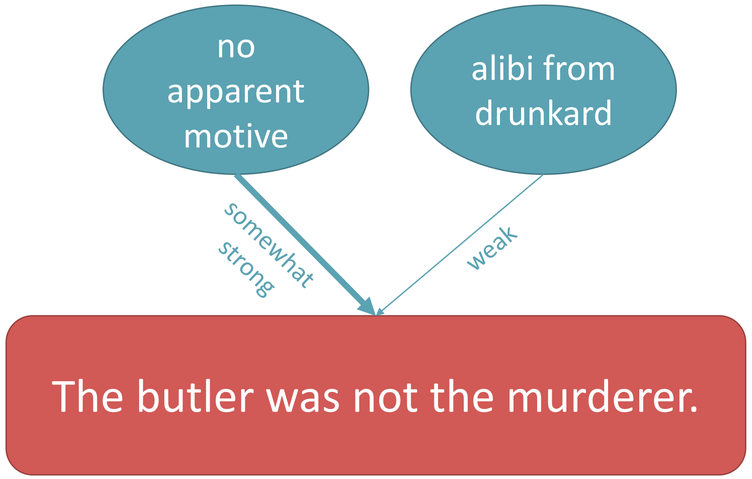
analogies ✓

Evidence

Very roughly: E is evidence for X iff $P(X|E) > P(X)$

There can be conflicting evidence and evidence of different strength



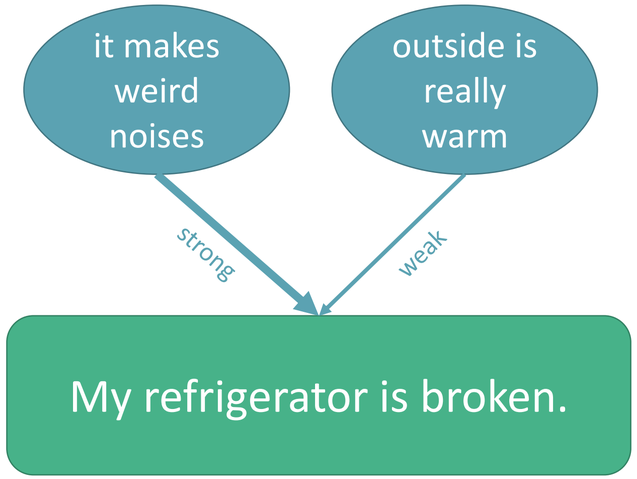


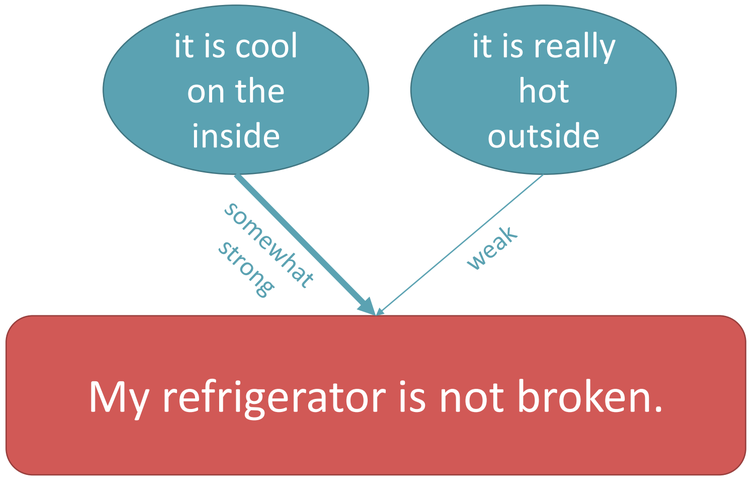
Often, the premises of defeasible arguments are evidence.

Evidence

Very roughly: E is evidence for X iff $P(X|E) > P(X)$

There can be conflicting evidence and evidence of different strength





Often, the premises of defeasible arguments are evidence.

does not have to be criminal evidence

Ethics for Nerds

An Advanced Course in Computer Science
Summer Semester 2020

Precise Thinking 3.3
Defeasible Arguments

A Complication and a Remark

Prof. Holger Hermanns,
Kevin Baum, Sarah Sterz

UNIVERSITÄT
DES
SAARLANDES

What's the indicator word?

Usually it is "probably", but…

P1: In lottery A you randomly draw exactly one lottery ticket from a box with 1,000,000 tickets of which exactly one is a winning ticket.

C: You won't draw the winning ticket in lottery A. ✗ | ✓

No keyword:
- not valid
- defeasibly cogent

P1: In lottery A you randomly draw exactly one lottery ticket from a box with 1,000,000 tickets of which exactly one is a winning ticket.

C: Probably, you won't draw the winning ticket in lottery A. ✓ | -

Keyword part of conclusion:
- valid
- not defeasibly cogent

P1: In lottery A you randomly draw exactly one lottery ticket from a box with 1,000,000 tickets of which exactly one is a winning ticket.

C: Probably, you won't draw the winning ticket in lottery A. ✗ | ✓

Keyword not part of conclusion:
- not valid
- defeasibly cogent

The "probably" can either be a keyword or part of the conclusion, and this can make a difference.

Inheritance of modifiers

If a premise contains a modifier that weakens a premise (like “probably”), then the conclusion will need to inherit it, if it is not caught by another premise.

Example:

P1: Probably, it is going to rain.

P2: If, probably, it is going to rain, then we should take an umbrella.

C: We should take an umbrella.

P2 catches the “probably”:

P1: a

P2: $a \rightarrow b$

C: b

P1: Probably, Jessica will be late.

P2: If Jessica will be late, then we will have to leave without her.

C: Probably, we will have to leave without Jessica.

The “probably” is not caught:

P1: Probably, a

P2: $a \rightarrow b$

C: Probably, b

https://www.reddit.com/r/ProgrammerHumor/comments/7uyafj/roses_are_red_violets_are_blue/