Aggregating Evidence about the Positive and Negative Effects of Treatments using a Computational Model of Argument

Anthony Hunter[1] and Matthew Williams[2]

[1]Department of Computer Science, University College London.
[2]Department of Oncology, Charing Cross Hospital, London.

October 10, 2016

Content

- Computational models of argument
- Some applications of argumentation in healthcare
- Problem of evidence aggregation
- Aggregating evidence using argumentation
- Conclusions

Argumentation as a cognitive process

Arguments are normally based on imperfect information

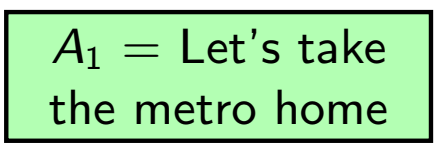
Arguments are normally constructed from information that is incomplete, inconsistent, uncertain and/or subjective, and from multiple sources.

Diverse examples of arguments

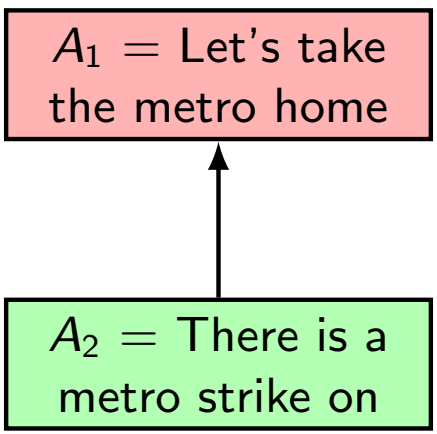
Mathematical	All squares have 4 corners. That is a square, and so it has 4 corners.
Epsitemic	If my sister was diagnosed with glaucoma, I would have known about it. As I don't, my sister hasn't been diagnosed with it.
Scientific	Mr Jones has angina. Aspirin has been shown to reduce risk of heart attack in angina patients. So prescribe daily aspirin.
Subjective	Prescribe nurofen because the patient prefers it, and the alternatives are not more effective clinically.

Abstract argumentation: Winning arguments

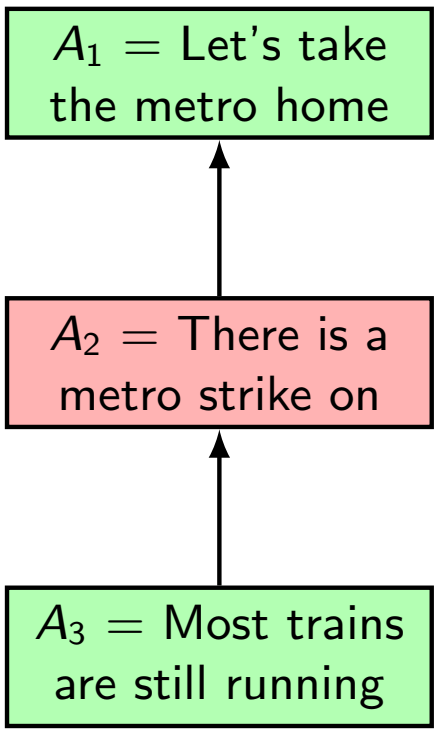
Green means the argument "wins" and red means the argument "loses".



Graph 1

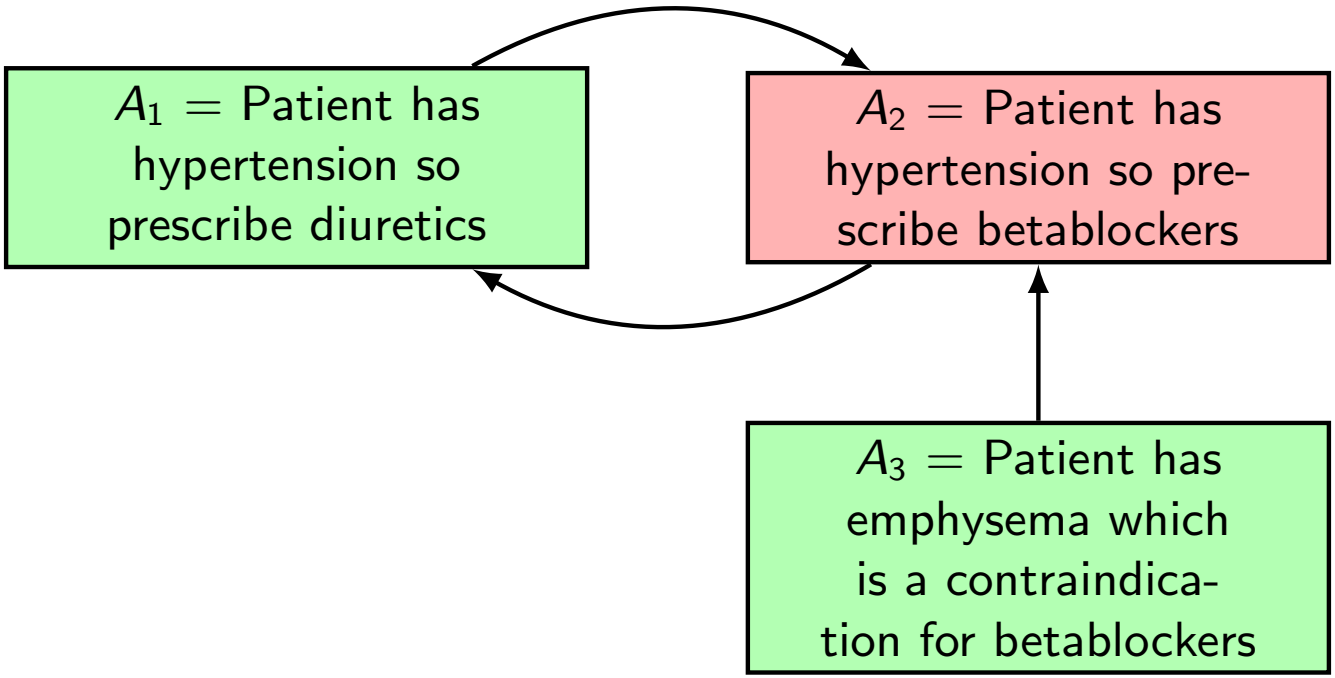


Graph 2



Graph 3

Abstract argumentation: Graphical representation



Towards applications in healthcare

Some examples of applications of argumentation in healthcare

- Computer decision support for GP prescribing (by John Fox et al.)
- Computer decision support for breast multi-disciplinary meetings (by Vivek Patkar, Dionisio Acosta, John Fox, et al.)
- Aggregating evidence about the positive and negative effects of treatments (by Anthony Hunter and Matthew Williams)
- Identifying clinical trials relevant for a specific patient (by Francesca Toni and Matthew Williams)
- Supporting patient decision making (by Anthony Hunter, Astrid Mayer and Kawsar Noor)

Some problems with primary evidence

- Evidence-based decision making relies on harnessing primary evidence (e.g. RCTs, observational studies, etc).
- But there is a lot of primary evidence to assimilate.
- Thousands of new results are published each year.
- The evidence is often
  - heterogeneous
  - uncertain
  - incomplete
  - inconsistent
- Published aggregates (e.g. systematic reviews, guidelines, etc) can help address the problem of dealing with primary evidence.

Some problems with aggregates

- However, published aggregates (e.g. systematic reviews, guidelines, etc) are
  1. expensive to produce
  2. take a long time to produce
  3. can become out of date quickly
  4. are for broad patient groups
  5. normally do not consider co-morbidities
  6. often use subjective or context-specific criteria to interpret the evidence, and these are not always made explicit
  7. decouple clinicians from the aggregation process, denying them the opportunity to use their own subjective or context-specific criteria
- Therefore there is a need for formal / computational tools to aggregate evidence.

Evidence-based decision making: Our aim

A *computational analysis framework for evidence* to help

Producers of aggregates (e.g. guideline development groups, systematic reviewers, etc) to analysis the available evidence to see what are justifiable ways to aggregate the evidence, and thereby make recommendations

Researchers to identify where the are important gaps in the current state of the literature and thereby identify new questions for clinical trials.

Clinicians to aggregate evidence using their subjective and contextual criteria for specific patients (perhaps with multiple issues)

Aggregating evidence concerning multiple outcomes

A simple example

Let CP denote contraceptive pill and NT denote no treatment.

ID	Left	Right	Indicator	Risk ratio	Outcome	p
e1	CP	NT	Pregnancy	0.05	superior	0.01
e2	CP	NT	Ovarian cancer	0.99	superior	0.07
e3	CP	NT	Breast cancer	1.04	inferior	0.01
e4	CP	NT	DVT	1.02	inferior	0.05

Our approach to aggregating evidence is based on argumentation.

Arguments based on evidence

Inductive arguments

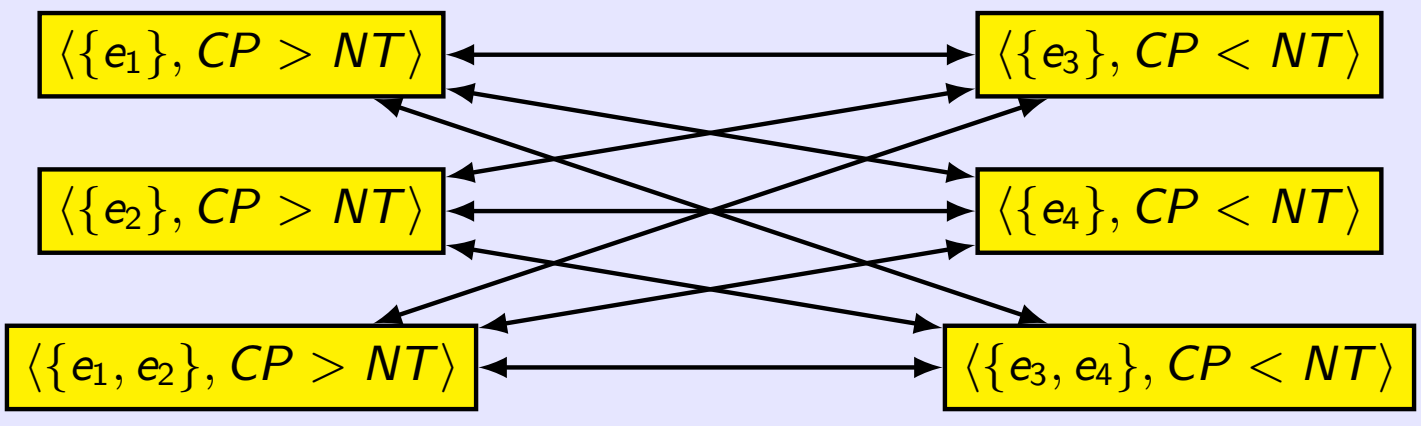
Given treatments τ_1 and τ_2, there are three kinds of inductive argument that can be formed.

1. $\langle X, \tau_1 > \tau_2 \rangle$, meaning the evidence in X supports the claim that treatment τ_1 is superior to τ_2.
2. $\langle X, \tau_1 \sim \tau_2 \rangle$, meaning the evidence in X supports the claim that treatment τ_1 is equivalent to τ_2
3. $\langle X, \tau_1 < \tau_2 \rangle$, meaning the evidence in X supports the claim that treatment τ_1 is inferior to τ_2.

Arguments with different claims can conflict

Example where CP is contraceptive pill and NT is no treatment

ID	Left	Right	Indicator	Risk ratio	Outcome	p
e1	CP	NT	Pregnancy	0.05	superior	0.01
e2	CP	NT	Ovarian cancer	0.99	superior	0.07
e3	CP	NT	Breast cancer	1.04	inferior	0.01
e4	CP	NT	DVT	1.02	inferior	0.05



Preferences over outcomes and their magnitude

To decide whether one choice is better than another, we need both the outcome type and its magnitude.

Example 1

- Benefit of choice 1 (CP): Relative risk of pregnancy is 0.01.
- Benefit of choice 2 (NT): Relative risk of breast cancer is 0.99.

Example 2

- Benefit of choice 1 (CP): Relative risk of pregnancy is 0.5.
- Benefit of choice 2 (NT): Relative risk of breast cancer is 0.5.

Preferences over inductive arguments

Example where CP is contraceptive pill and NT is no treatment

ID	Left	Right	Indicator	Risk ratio	Outcome	p
e1	CP	NT	Pregnancy	0.05	superior	0.01
e2	CP	NT	Ovarian cancer	0.99	superior	0.07
e3	CP	NT	Breast cancer	1.04	inferior	0.01
e4	CP	NT	DVT	1.02	inferior	0.05

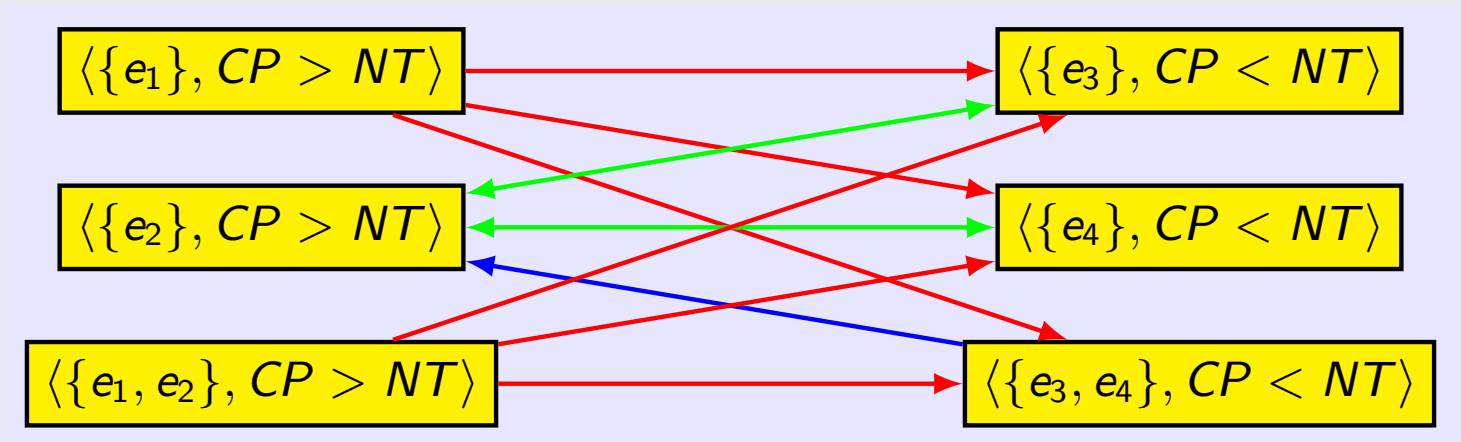
- Substantial reduction in pregnancy is *more preferred* to modest reduction in risk of either breast cancer or DVT.
- Modest reduction in risk of ovarian cancer is *equally preferred* to modest reduction in risk of either breast cancer or DVT.
- Modest reduction in risk of ovarian cancer is *less preferred* to modest reduction inower risk in both DVT and breast cancer.

Preferences over inductive arguments

The preferences over outcomes (and their magnitude) is used as the preference relation over arguments.

Example where CP is contraceptive pill and NT is no treatment

ID	Left	Right	Indicator	Risk ratio	Outcome	p
e1	CP	NT	Pregnancy	0.05	superior	0.01
e2	CP	NT	Ovarian cancer	0.99	superior	0.07
e3	CP	NT	Breast cancer	1.04	inferior	0.01
e4	CP	NT	DVT	1.02	inferior	0.05



Meta-arguments

Some types of meta-argument

- Meta-arguments are counterarguments to inductive arguments
- Meta-arguments are reasons based on weaknesses in the evidence in inductive arguments

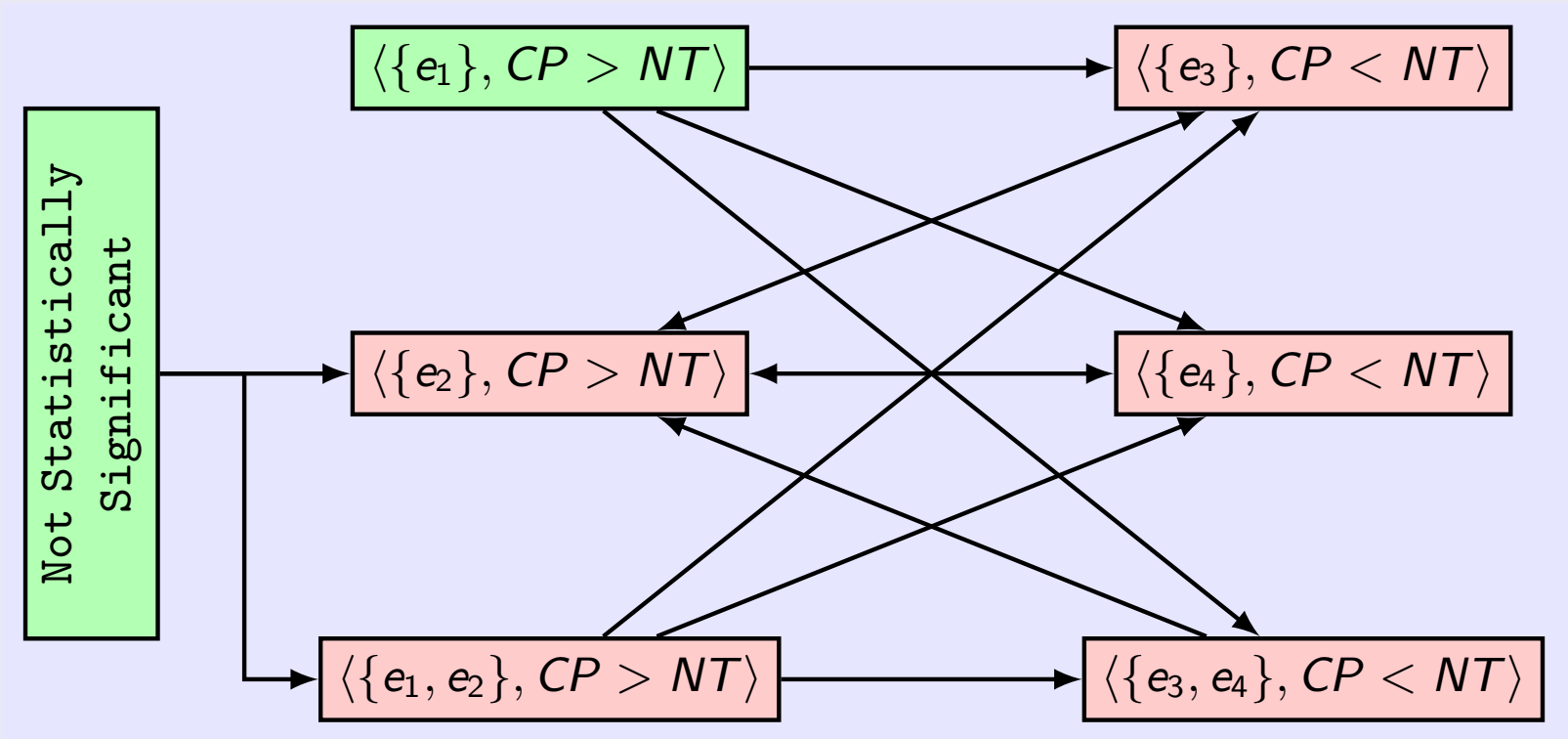
Some types of meta-argument

- The evidence contains flawed RCTs.
- The evidence is not statistically significant.
- The evidence is from trials with narrow patient class.
- The evidence has outcomes that are not consistent.

Argument graph with inductive and meta-arguments

Example where CP is contraceptive pill and NT is no treatment

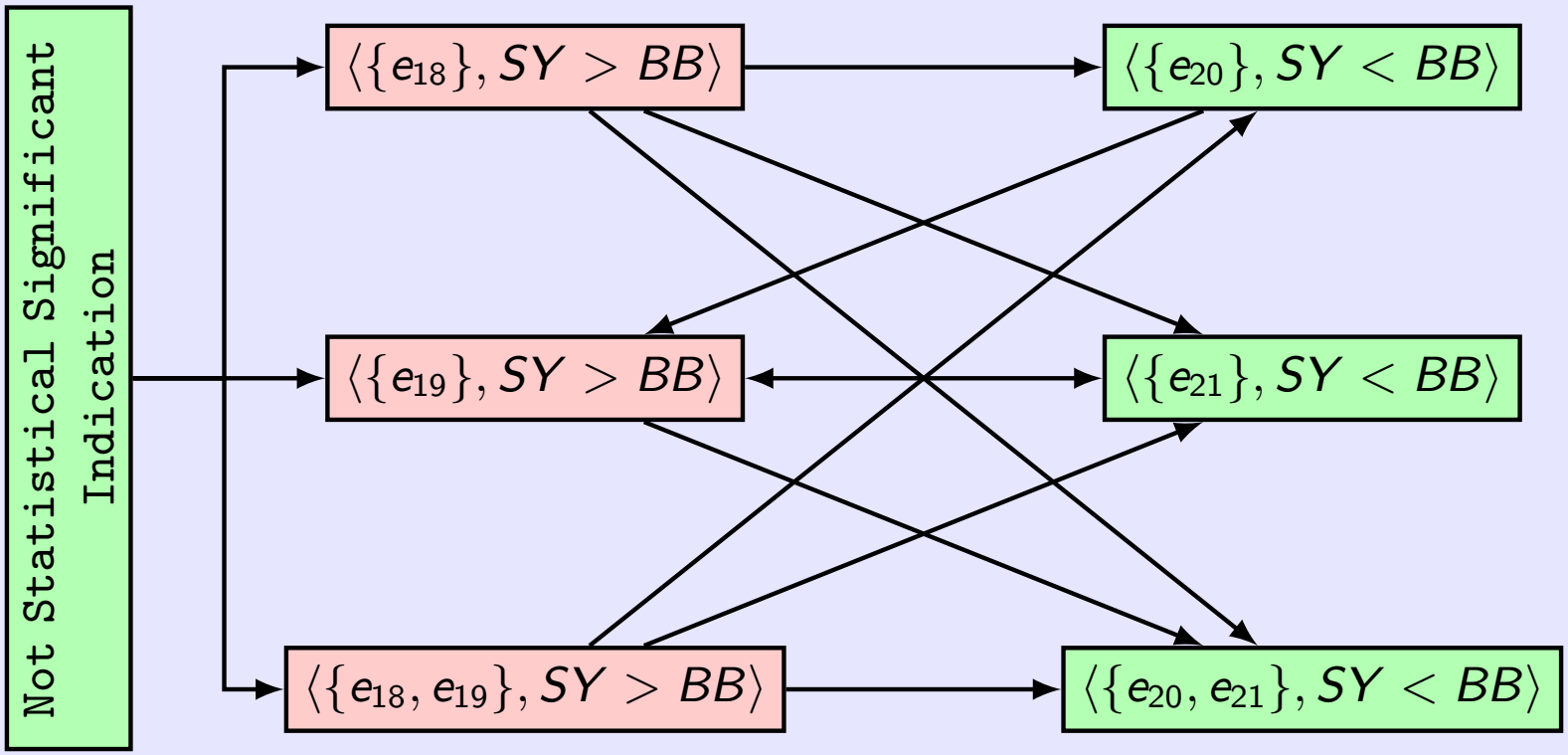
ID	Left	Right	Indicator	Risk ratio	Outcome	p
e1	CP	NT	Pregnancy	0.05	superior	0.01
e2	CP	NT	Ovarian cancer	0.99	superior	0.07
e3	CP	NT	Breast cancer	1.04	inferior	0.01
e4	CP	NT	DVT	1.02	inferior	0.05



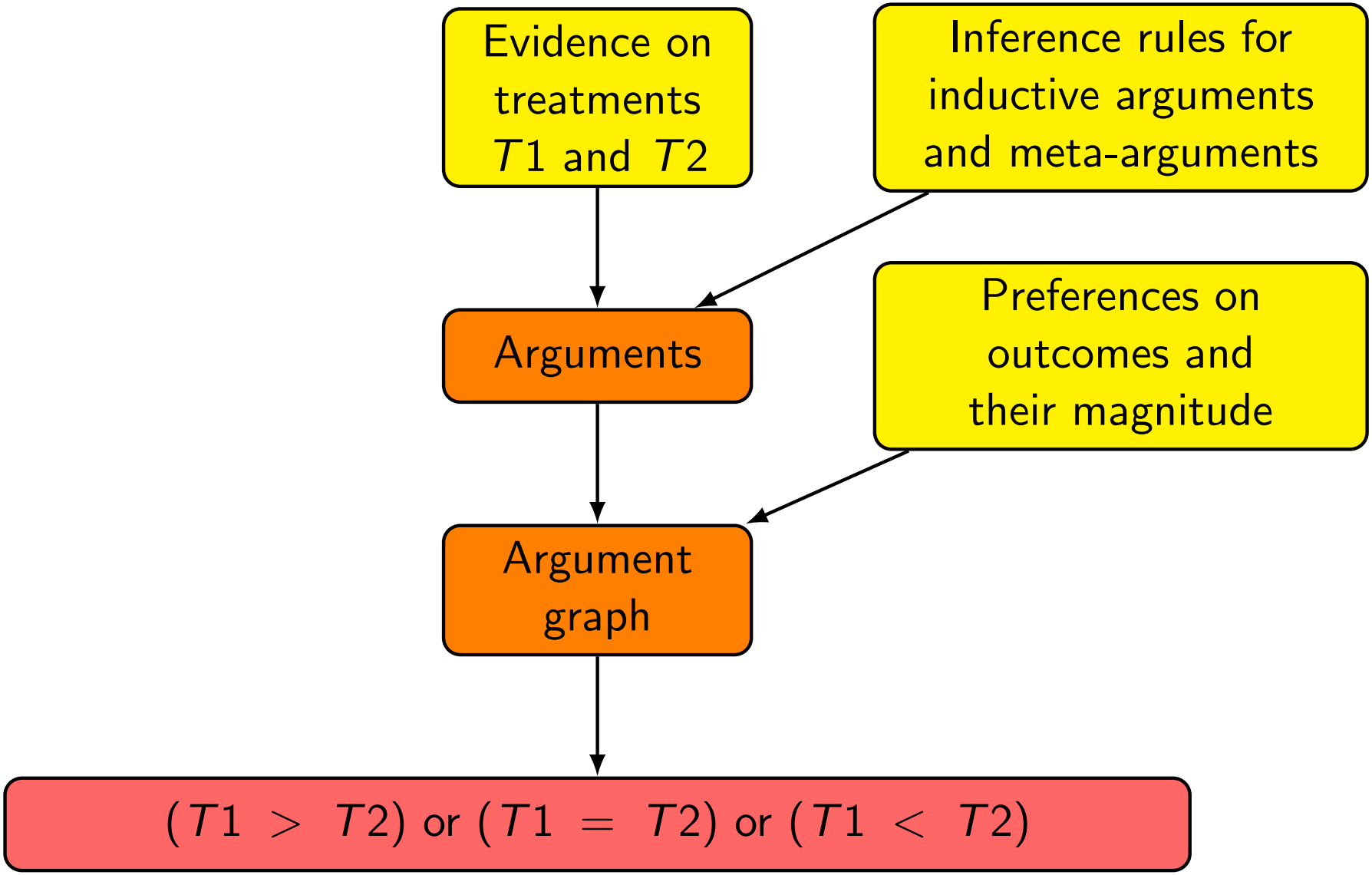
Argument graph with inductive and meta-arguments

Example with beta-blockers (BB) and sympathomimetics (SS)

	Left	Right	Outcome indicator	Value	Net	Sig	Type
e_{18}	SY	BB	visual field prog	0.92	$>$	no	MA
e_{19}	SY	BB	change in IOP	-0.25	$>$	no	MA
e_{20}	SY	BB	allergic prob	41.00	$<$	yes	MA
e_{21}	SY	BB	drowsiness	1.21	$<$	no	MA



Summary of our approach



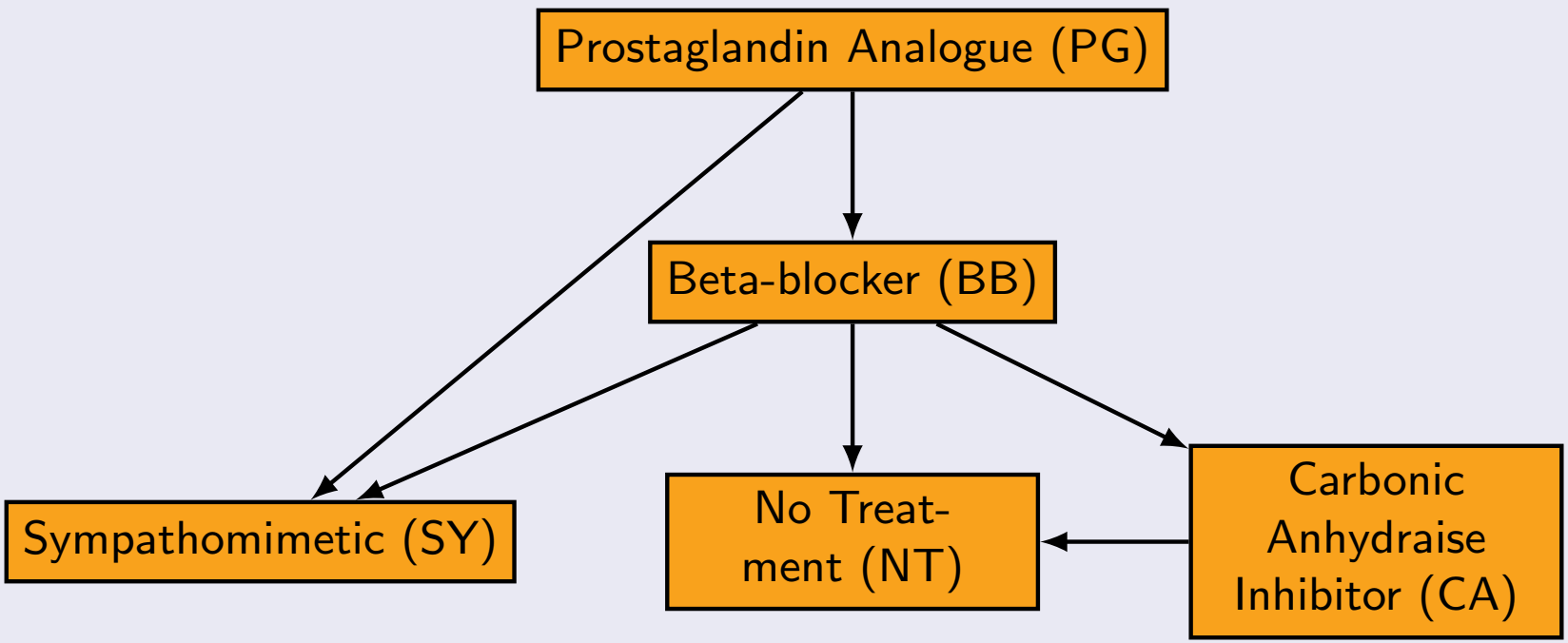
Case study: NICE Glaucoma Guideline

Evidence taken from the guideline

	Left	Right	Outcome indicator	Value	Net	Sig	Type
e_{01}	BB	NT	visual field prog	0.77	$>$	no	MA
e_{02}	BB	NT	change in IOP	-2.88	$>$	yes	MA
e_{03}	BB	NT	respiratory prob	3.06	$<$	no	MA
e_{04}	BB	NT	cardio prob	9.17	$<$	no	MA
e_{05}	PG	BB	change in IOP	-1.32	$>$	yes	MA
e_{06}	PG	BB	acceptable IOP	1.54	$>$	yes	MA
e_{07}	PG	BB	respiratory prob	0.59	$>$	yes	MA
e_{08}	PG	BB	cardio prob	0.87	$>$	no	MA
e_{09}	PG	BB	allergy prob	1.25	$<$	no	MA
e_{10}	PG	BB	hyperaemia	3.59	$<$	yes	MA
e_{11}	PG	SY	change in IOP	-2.21	$>$	yes	MA
e_{12}	PG	SY	allergic prob	0.03	$>$	yes	MA
e_{13}	PG	SY	hyperaemia	1.01	$<$	no	MA
e_{14}	CA	NT	convert to COAG	0.77	$>$	no	MA
e_{15}	CA	NT	visual field prog	0.69	$>$	no	MA
e_{16}	CA	NT	IOP $>$ 35mmHg	0.08	$>$	yes	MA
e_{17}	CA	BB	hyperaemia	6.42	$<$	no	MA
e_{18}	SY	BB	visual field prog	0.92	$>$	no	MA
e_{19}	SY	BB	change in IOP	-0.25	$>$	no	MA
e_{20}	SY	BB	allergic prob	41.00	$<$	yes	MA
e_{21}	SY	BB	drowsiness	1.21	$<$	no	MA

Case study: NICE Glaucoma Guideline

Superiority graph obtained via our argumentation approach



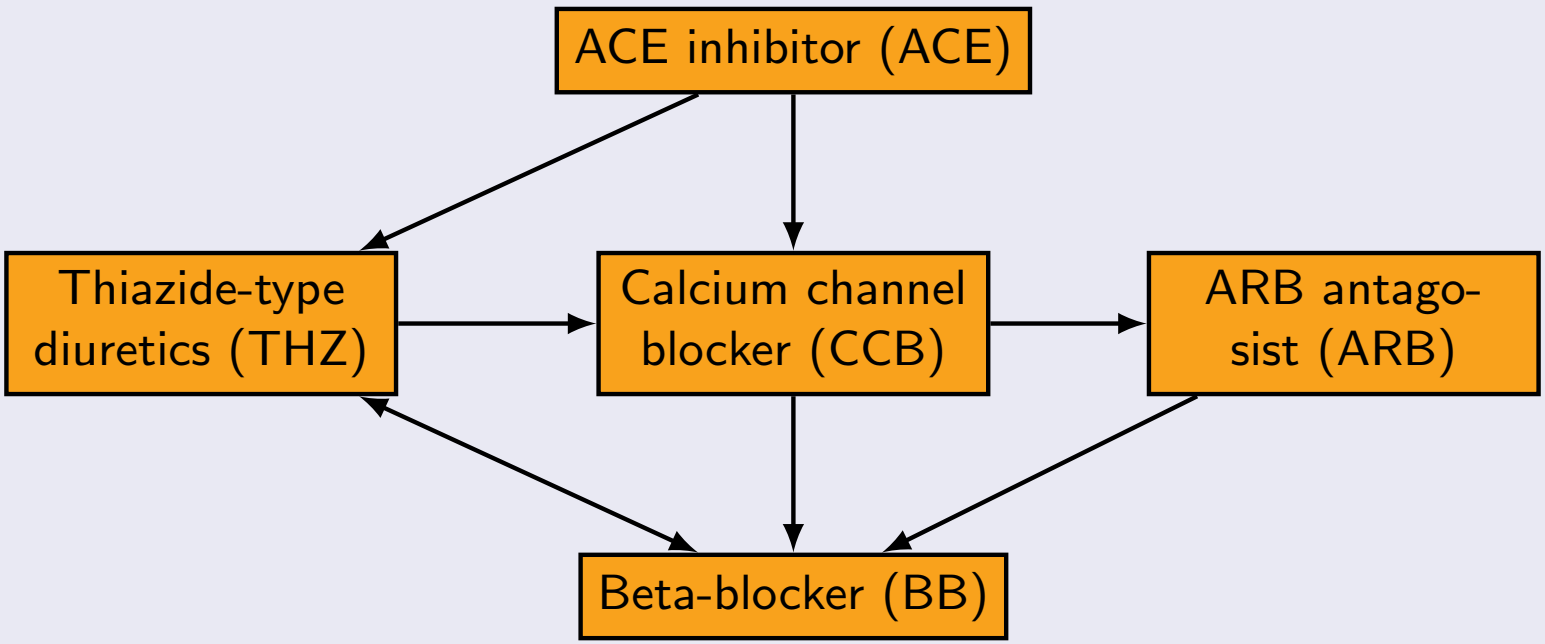
Case study: NICE Hypertension Guideline

Evidence taken from the guideline

	Left	Right	Outcome indicator	Value	Net	Sig	Type
e_{01}	BB	THZ	mortality	1.04	$<$	no	MA
e_{02}	ACE	CCB	mortality	1.04	$<$	no	MA
e_{03}	ACE	CCB	stroke	1.15	$<$	yes	MA
e_{04}	ACE	CCB	heart failure	0.85	$>$	yes	MA
e_{05}	ACE	CCB	diabetes	0.85	$>$	yes	MA
e_{06}	ARB	BB	mortality	0.89	$>$	no	MA
e_{07}	ARB	BB	myocardial infarction	1.05	$<$	no	MA
e_{08}	ARB	BB	stroke	0.95	$>$	no	MA
e_{09}	ARB	BB	heart failure	1.25	$<$	no	MA
e_{10}	ARB	BB	diabetes	0.75	$>$	yes	MA
e_{11}	ARB	CCB	mortality	1.02	$<$	no	MA
e_{12}	ARB	CCB	myocardial infarction	1.17	$<$	yes	MA
e_{13}	ARB	CCB	stroke	1.14	$<$	no	MA
e_{14}	ARB	CCB	heart failure	0.88	$>$	no	MA
e_{15}	ACE	THZ	mortality	1.00	$\sim$	no	MA
e_{16}	ACE	THZ	stroke	1.13	$>$	yes	MA
e_{17}	CCB	BB	mortality	0.94	$>$	no	MA
e_{18}	CCB	BB	myocardial infarction	0.93	$>$	no	MA
e_{19}	CCB	BB	stroke	0.77	$>$	yes	MA
e_{20}	CCB	BB	diabetes	0.71	$>$	yes	MA
e_{21}	CCB	THZ	mortality	0.97	$<$	no	MA
e_{22}	CCB	THZ	myocardial	1.02	$>$	no	MA
e_{23}	CCB	THZ	stroke	0.95	$<$	yes	MA
e_{24}	CCB	THZ	heart failure	1.38	$>$	yes	MA
e_{25}	CCB	THZ	diabetes	0.82	$<$	yes	MA

Case study: NICE Hypertension Guideline

Superiority graph obtained via our argumentation approach



Case study: NICE Pre-eclampsia Guideline

Evidence taken from the guideline

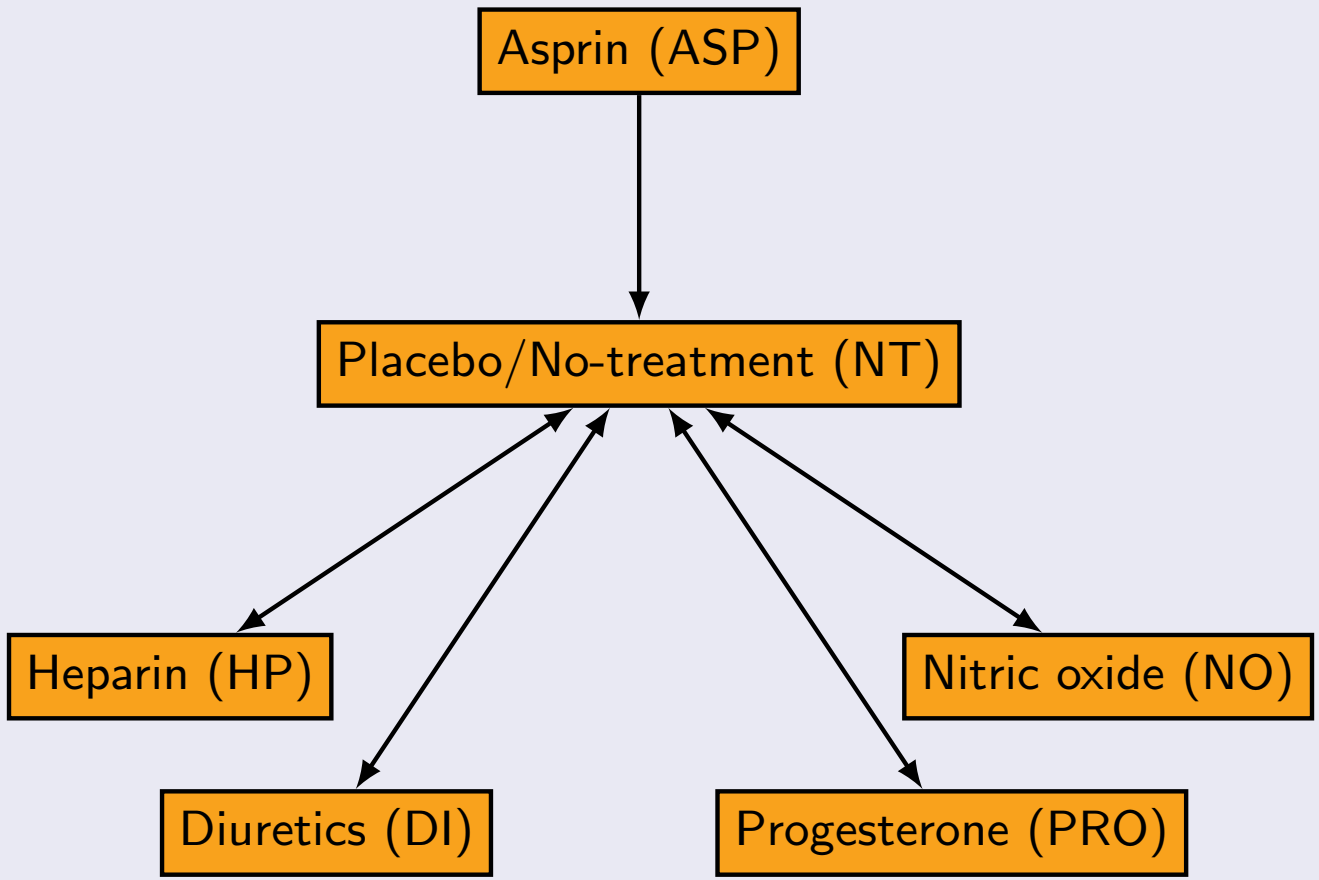
	Left	Right	Outcome indicator	Value	Net	Sig	Type	Note
e_{01}	HEP	NT	pre-eclampsia	0.26	$>$	yes	MA	1,2
e_{02}	HEP	NT	fetal growth restriction	0.22	$>$	yes	MA	1,2
e_{03}	HEP	NT	gestational diabetes	0.48	$>$	no	MA	1,2
e_{04}	DI	NT	pre-eclampsia	0.68	$>$	no	MA	
e_{05}	PRO	NT	pre-eclampsia	0.21	$>$	no	MA	
e_{06}	NO	NT	pre-eclampsia	0.83	$>$	no	MA	
e_{07}	ASP	NT	pre-eclampsia	0.83	$>$	yes	MA	
e_{08}	ASP	NT	preterm	0.92	$>$	yes	MA	
e_{09}	ASP	NT	fetal & neonatal death	0.86	$>$	yes	MA	
e_{10}	ASP	NT	small gestational age	0.90	$>$	yes	MA	

In the Notes column,

- denotes that the evidence is from non-randomized and non-blind trials,
- denotes that the trials are for very narrow patient classes.

Case study: NICE Pre-eclampsia Guideline

Superiority graph obtained via our argumentation approach



Using our method in practice

- Over 200 items of evidence concerning chemo-radiotherapy treatments for non-small cell lung cancer
- 27 treatment regimens considered
- 8 combinations of meta-arguments and preference criteria
- We obtained a finer-grained analysis of the literature than the Cochrane Review

M Williams, Z. Liu, A.Hunter and F. MacBeth (2015) An updated systematic review of lung chemo-radiotherapy using a new evidence aggregation method. Lung Cancer 87(3):290-5

Conclusions

Benefits of logical argumentation for evidence aggregation

- Each inductive argument is an aggregation of the evidence
- Preferences over outcomes and their magnitude are used to define preferences over inductive arguments.
- Each meta-argument provides a reason for rejecting an inductive argument based on weaknesses in the evidence used.
- Dialectical criteria used to determine which arguments are acceptable.

References

1. A Hunter and M Williams (2012) Aggregating evidence about the positive and negative effects of treatments, *AI in Medicine Journal*, 56:173-190.
2. A Hunter and M Williams (2015) Aggregation of Clinical Evidence using Argumentation: A Tutorial Introduction, *Foundations of Biomedical Knowledge Representation*, edited by Arjen Hommersom and Peter Lucas, LNCS volume 9521, Springer, pages 317–338.