Neurocomputational Approaches to Decision Making

Parameterized complexity theory: An indispensable tool for the cognitive (neuro)scientist

Johan Kwisthout
(joint work with Iris van Rooij, Todd Wareham, & Mark Blokpoel)

This lecture

- This lecture focuses on the **use** of **computational models** of (**neuro**-)**cognitive phenomena** for **explanatory** purposes
- What you will learn here is a **modeling technique** for **analysing** particular computational models and **constraining** the search space of possible models
- Many such constraints are relevant: biological plausibility, realistic assumptions, falsifiability, ...
- We focus on **computational tractability** as a model constraint based on theoretical computer science

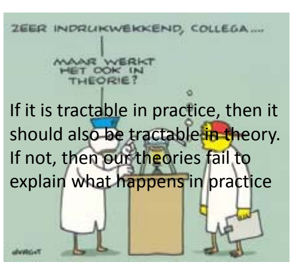
This lecture

- We will not focus on **particular** models of concrete neurocognitive phenomena related to decision making
- In contrast, I will discuss a technique that is relevant **for all** computational models of **all cognitive capacities**
- However, I will focus, without loss of generality, on **Bayesian cognitive models**, because:
  - Many models of decision-making sub-processes are in fact Bayesian models
  - We have mostly worked with such models and have the most interesting results of the technique I will describe

Bayesian inference to the best explanation

- Inference to the best explanation ("trying to make sense of the phenomena we observe") is a key concept in many computational models of cognitive capacities / domains
  - Baker et al.'s models of Theory of Mind and Action Understanding
  - Van Rooij et al.'s models of Intention Recognition and Recipient Design
  - Yuille & Kersten's model of Visual Perception
  - Chater & Manning 's models of Language Processing
- Yet, (Bayesian) inference to the best explanation is known to be a highly intractable problem in general
  - Bylander et al., 1991; Nordh & Zanuttini, 2005; Kwisthout, 2011
- Is this a problem for computational cognitive models that are based on Bayesian inference to the best explanation?

Fokke en Sukke



If it is tractable in practice, then it should also be tractable in theory. If not, then our theories fail to explain what happens in practice

"Very impressive, colleague, but does it also work in theory?"

The Tractability Constraint

"The computations postulated by a model of cognition need to be tractable in the real world in which people live, not only in the small world of an experiment ... This eliminates NP-hard models that lead to computational explosion." (Gigerenzer et al., 2008)

The Tractability Constraint

"The computations postulated by a **model of cognition** need to be tractable in the real world in which people live, not only in the small world of an experiment ... This eliminates NP-hard models that lead to computational explosion." (Gigerenzer et al., 2008)

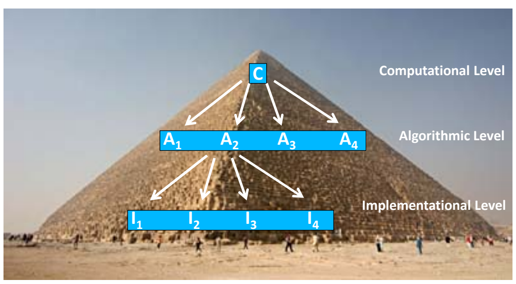
Marr's three levels of explanation

input i → Cognitive process → output o = f(i)

Level	Marr's levels	Question
1	Computational	What?
2	Algorithm	Method?
3	Implementation	Implementation?

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

A pyramid of explanations



http://en.wikipedia.org/wiki/Great_Pyramid_of_Giza

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

Computational-level Models of Cognition

input i → Cognitive process → output o = f(i)

 → Bayesian inference → 

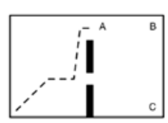
Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

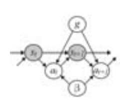
The Tractability Constraint

"The computations postulated by a model of cognition need to be tractable **in the real world in which people live**, **not only in the small world of an experiment** ... This eliminates NP-hard models that lead to computational explosion." (Gigerenzer et al., 2008)

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

Scalability





Baker, C.L., Tenenbaum J.B.,& Saxe, R.R. (2007) Goal Inference as Inverse Planning, CogSci'07

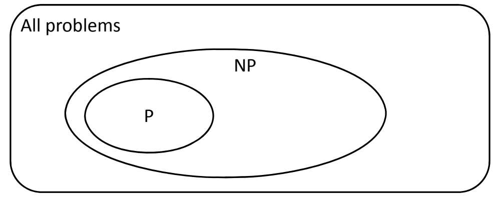
Does this scale to, e.g., recognizing intentions in a job interview?

Donders Institute for Brain, Cognition and Behaviour
Radboud University Nijmegen

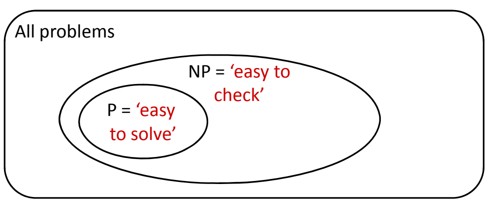
The Tractability Constraint

"The computations postulated by a model of cognition **need to be tractable** in the real world in which people live, not only in the small world of an experiment ... **This eliminates NP-hard models** that lead to computational explosion." (Gigerenzer et al., 2008)

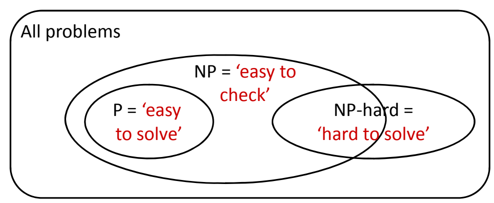
What are P, NP and NP-hard?



What are P, NP and NP-hard?



What are P, NP and NP-hard?



Intuitive example: hard to solve & easy to check

Sudoku

			7	1	2	9		3
5		2		3	8			
			9					
		8					9	4
	1		5		9		8	
9	5					7		
					3			
			1	9		3		8
6		1	2	8	5			

8	4	6	7	1	2	9	5	3
5	9	2	6	3	8	1	4	7
1	7	3	9	5	4	8	2	6
2	6	8	3	7	1	5	9	4
3	1	7	5	4	9	6	8	2
9	5	4	8	2	6	7	3	1
7	8	9	4	6	3	2	1	5
4	2	5	1	9	7	3	6	8
6	3	1	2	8	5	4	7	9

Why NP-hard is considered intractable

NP-hard functions cannot be computed in polynomial time (assuming $P \neq NP$). Instead they require exponential time (or worse) for their computation, which is why they are considered intractable (in other words, unrealistic to compute for all but small inputs).

n	n^2
5	25
20	400
50	2500
100	10000
1000	10^6

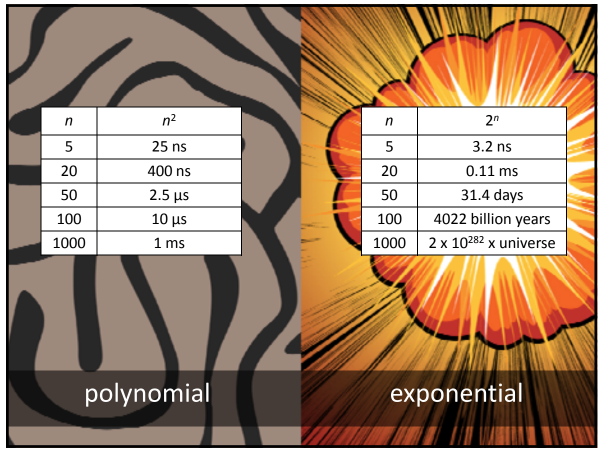
n	2^n
5	32
20	1.05^6
50	1.13×10^{15}
100	1.27×10^{30}
1000	1.07×10^{301}

Why NP-hard is considered intractable

Assuming a computer with 10 Gigaflops computing speed (10^{10} basic instructions per second):

n	n^2
5	2.5 ns
20	40 ns
50	0.25 µs
100	1 µs
1000	0.1 ms

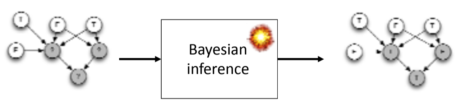
n	2^n
5	3.2 ns
20	0.11 ms
50	31.4 days
100	4022 billion years
1000	2×10^{282} x universe



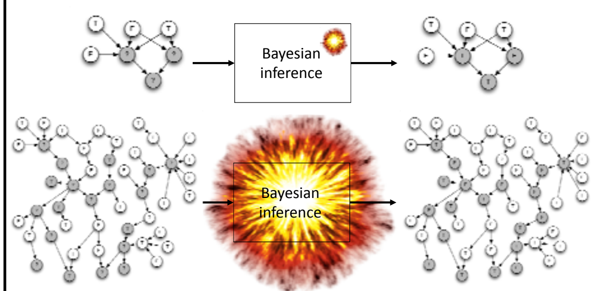
n	n^2
5	25 ns
20	400 ns
50	2.5 µs
100	10 µs
1000	1 ms

n	2^n
5	3.2 ns
20	0.11 ms
50	31.4 days
100	4022 billion years
1000	2×10^{282} x universe

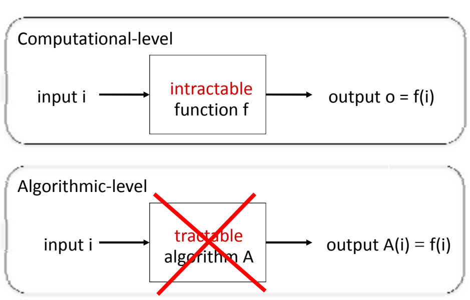
Bayesian Inference is Intractable (NP-hard)



Bayesian Inference is Intractable (NP-hard)

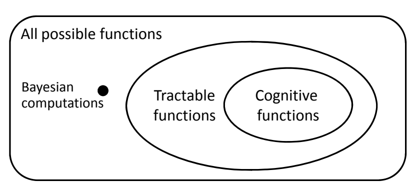


Intractable function, tractable algorithm?



Intractability as "good news"
tractability is a model constraint

Tractability as a theoretical constraint helps cognitive modelers by ruling out unrealistic models



Can approximation help us?

Computational-level

input i → intractable function f → output o = f(i)

consistent

Algorithmic-level

input i → tractable algorithm A → output A(i) ≈ f(i)

For Bayesian computations: in general, NO

Computational-level

input i → intractable function f → output o = f(i)

Algorithmic-level

input i → ~~tractable algorithm A~~ → output A(i) ≈ f(i)

Classical Bayesian approximation results

It is **NP-hard** to:

- *compute posteriors* approximately (Dagum & Luby, 1994)
- *fixate beliefs* approximately (Abdelbar & Hedetniemi, 1998) or with a *non-zero* probability (Kwisthout, 2011)
- fixate a belief that *resembles* the most probable belief (Kwisthout, 2012)
- fixate a belief that is *likely the most probable* belief (Kwisthout & van Rooij, 2012)
- fixate a belief that *ranks within the k most probable* beliefs (Kwisthout, 2014)

So, what now?

- Bayesian computations are NP-hard, that is, intractable in general; both to compute exactly and to approximate for any sensible meaning of 'approximation'
- "Now what?" – three ways of dealing with intractability
  - The **doomsday** approach ("the pessimists")
  - The **hand-waving** approach ("the optimists")
  - The **analytical or rational** approach ("the realists")

The doomsday approach

Bayesian inferences are NP-hard

Bayesian models are no good models of the brain

- The **pessimists** throw away the baby with the bath-water: because Bayesian computations are NP-hard, that doesn't rule out that *many* instances of Bayesian models problems can be solved tractably

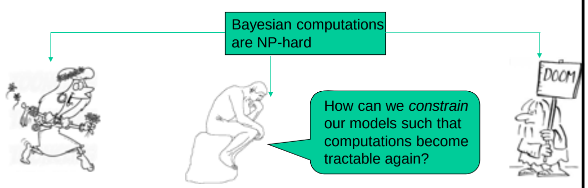
The hand-waving approach

Bayesian inferences are NP-hard

OK, fine; we'll just assume that the mind *approximates* Bayesian computations then…

- The **optimists** try to solve the problem by asserting that approximation, satisficing, and using heuristics will be sufficient to overcome intractability. However, we already saw that approximate Bayesian computations are just as hard as exact Bayesian computations…

The analytical approach

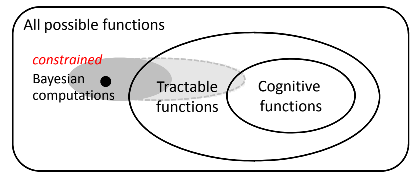


- The **realists** see the strength of Bayesian models but acknowledge that they are too broad and need to be constrained in order to overcome intractability. They will look for ***problem parameters*** that – when bounded – render the problem tractable

Intractability as "good news"

tractability is a model constraint

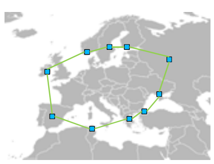
Constraining computational models may help to buy tractability – and even cognitive plausibility

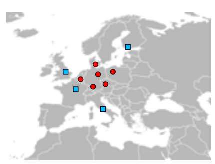


How to constrain Bayesian inferences?

Step 1. Identify **parameters** of the model that can be proven to be sources of intractability

- In general, NP-hard problems take exponential time **in the worst case** to solve → some instances are easy, some are hard
- Identify what makes these instances hard (or easy)

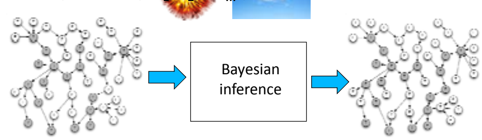




How to constrain Bayesian inferences?

Step 1. Identify **parameters** of the model that can be proven to be sources of intractability

$\exp(n) \Rightarrow \exp(k_1, k_2, \ldots k_m)\mathrm{poly}(n)$



For example, k_1 : max out degree
k_2 : max # unknowns
k_3 : etc.

How to constrain Bayesian inferences?

Step 1. Identify **parameters** of the model that can be proven to be sources of intractability

$\exp(n) \Rightarrow \exp(k_1, k_2, \ldots k_m)\mathrm{poly}(n)$

Step 2. Constrain the model to **small** values for the parameters k_1, k_2, …k_m. (Note: n can still be large!)

Step 3. Verify that the constraints hold for humans in **real-life** situations, and **test in the lab** if performance breaks down when parameters are large

What makes Bayesian inferences tractable?

Exact inferences

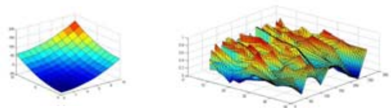
- ✗ degree of network?
- ✗ cardinality of variables?
- ✗ length of paths/chains?
- ✓ structure of dependences?
- ✓/✗ posterior probability?

Approximate inferences

- ✗ degree of network?
- ✗ cardinality of variables?
- ✗ length of paths/chains?
- ✓ structure of dependences?
- ✓/✗ posterior probability?
- ✓ characteristics of the probability distribution?

CPDs and approximate inference

- Local search techniques, MC sampling, etc., are dependent on the landscape of the probability distribution



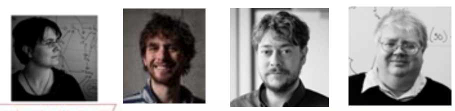
- For some Bayesian inference problems, this landscape can be parameterized – we can prove bounds on the success of the approximation algorithm relative to this parameter
- Kwisthout & Van Rooij (2013), Bridging the gap between theory and practice of approximate Bayesian inference. *Cognitive Systems Research, 24, 2–8.*
- Kwisthout (2015), Tree-Width and the Computational Complexity of MAP Approximations. *Journal of AI Research, in press*.

Our version of the Tractability Constraint

"The computations postulated by a model of cognition need to be tractable in the real world in which people live, not only in the small world of an experiment ...

This eliminates NP-hard models that lead to computational explosion." (Gigerenzer et al., 2008)

This poses the need for a **thorough analysis** of the sources of complexity underlying NP-hard models, and **eliminates NP-hard models** expect those that can be proven to be **fixed-parameter tractable** for parameters that may safely be assumed to be small in the real world.



Tractable-design cycle