Slide 1

Modeling Minds Workshop

Making sense of "making sense":
A complexity-theoretic perspective on approximate Bayesian inferences

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

Donders Institute
for Brain, Cognition and Behaviour

Radboud University Nijmegen

Slide 2

"Making sense"

- 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
  - Wolpert & Ghrahamani's models of Motor Planning

- Yet, (Bayesian) abduction in known to be a highly intractable problem in general
  - Bylander et al., 1991; Nordh & Zanuttini, 2005; Kwisthout, 2011

- Is this a problem for **explanatory** computational models?

Slide 3

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?"

Slide 4

Bayesian intractability…

- …is a problem with our models of the cognitive capacity

- …undermines the plausibility of computational models of cognition that are faced with this intractability
(Van Rooij, 2008)

- …is *inherent* to the computational function itself; not how it is realized; it cannot be solved by assuming a realization that acts "as if" it computes that function
(Van Rooij et al., in press)

- …is not an "ailment approximation can cure"
(Kwisthout & van Rooij, 2012)

Slide 5

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)

Slide 6
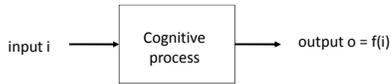
The Tractability Constraint

"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."
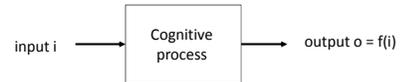(Gigerenzer et al., 2008)

Slide 1: Marr's three levels of explanation

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?

Slide 2: Computational-level Models of Cognition

Computational-level Models of Cognition

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

[network diagram] → [Bayesian inference] → [network diagram]

Slide 3: The Tractability Constraint

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)

Slide 4: Scalability

Scalability

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

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

Slide 5: The Tractability Constraint

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)

Slide 6: Why NP-hard is considered intractable

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

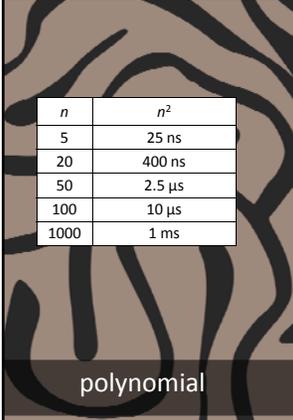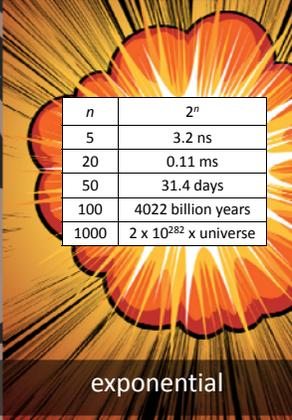| n	2^n
5	32
20	1.05^6
50	1.13×10^{15}
100	1.27×10^{30}
1000	1.07×10^{301}

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

polynomial   exponential
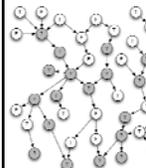

Bayesian Inference is Intractable (NP-hard)

Bayesian inference

Bayesian Inference is Intractable (NP-hard)

Bayesian inference

Bayesian inference

Intractable function, tractable algorithm?

Computational-level

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

Algorithmic-level

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

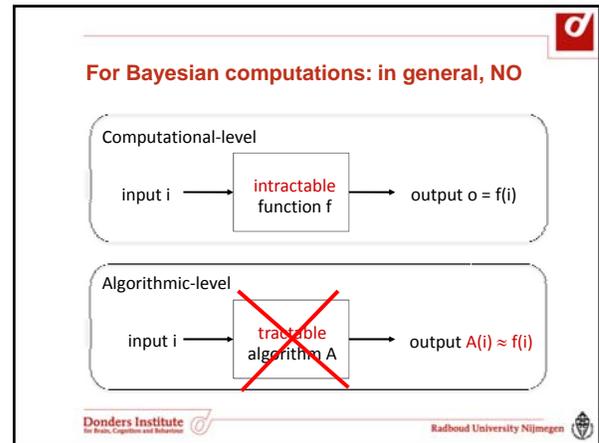## Intractability as "good news"
tractability is a model constraint

f

All possible functions

Tractable functions   Cognitive functions

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)

Intractability as "good news"
tractability is a model constraint

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

All possible functions

constrained Inference ●

constrained Belief fixation ●

Tractable functions

Cognitive functions
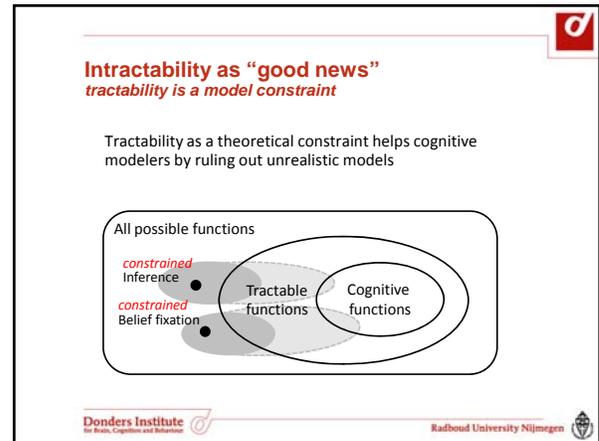
How to constrain Bayesian inferences?

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

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

→ | Bayesian inference | →

For example, k_1 : max out degree
k_2 : max # unknowns
k_3 : etc.
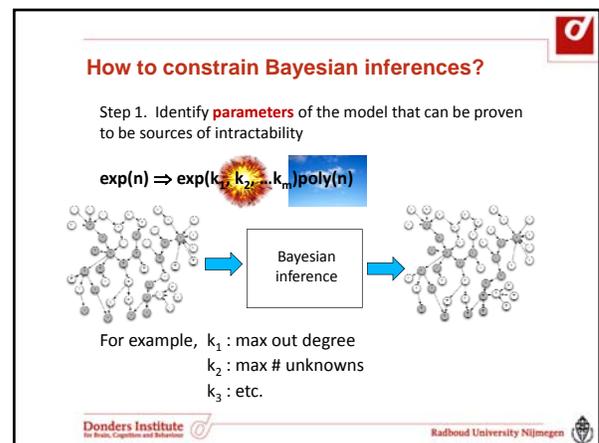
How to constrain Bayesian inferences?

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

$$\exp(n) \Rightarrow \exp(k_1, k_2, ...k_m)\text{poly}(n)$$

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	Approximate inferences
❌ degree of network?	❌ degree of network?
❌ cardinality of variables?	❌ cardinality of variables?
❌ length of paths/chains?	❌ length of paths/chains?
✅ structure of dependences?	✅ structure of dependences?
✅/❌ posterior probability?	✅/❌ 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
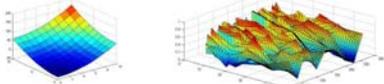
- For some Bayesian inference problems, this landscape can be parameterized – we can prove bounds on the success of the approximation algorithm relative to this parameter

- Kwisthout & Van Rooij (2013), Bridging the gap between theory and practice of approximate Bayesian inference. *Cognitive Systems Research, 24, 2–8.*

- Kwisthout (2014, under review), Treewidth and the Computational Complexity of MAP Approximations.

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

Computational level theory

Informal

Formal

Empirical testing

Tractability testing