AQ1-AQ4

Rational dynamics in efficient inquiry

David L. Barack

Abstract

Which premisses should we use to start our inquiries? Which transitions during inquiry should we take next? When should we switch lines of inquiry? In this paper, I address these open questions about inquiry, formulating novel norms for such decisions during deductive reasoning. I use the first-order predicate calculus, in combination with Carnap's state description framework, to state such norms. Using that framework, I first demonstrate some properties of sets of sentences used in deduction. I then state some norms for decisions made during deductive reasoning, establishing initial benchmarks for efficient deduction by ideal reasoners. When deciding which transition to make next, reasoners should choose the most informative transition, the one that maximally reduces uncertainty in the sense of ruling out the largest number of state descriptions relevant to their inquiry. Finally, inspired by optimal foraging theory, I show that, under certain assumptions of ignorance, reasoners should change premiss sets when their information intake drops below the global average information intake across premiss sets.

Keywords: reasoning, inquiry, foraging, efficient, bounded rationality.

1. Introduction

When inquiring, reasoners often must decide which steps to take next, such as deciding which conclusion to draw, premiss to select or line of inquiry to abandon. Much extant work (e.g. Alchourrón et al. 1985, Spohn 2012) focuses on how to formally model changes in belief, with less on how to choose beliefs, premisses, or lines of inquiry. The present work takes steps to address these questions from a normative perspective. A basic guiding principle is to seek premisses or transitions that help resolve inquiries in the smallest number of steps. Here I will outline a formalization of efficient deductive inquiry using the first-order predicate calculus and a state description framework, after Carnap. I have elected to use the first-order predicate calculus to illustrate the simple formalizability of norms for efficient inquiry.[1]

1 My motivations for choosing the first-order predicate calculus are pragmatic, to make my account more accessible to the philosophical community. While at times inelegant, basic logic is still adequate for illustrating novel norms for rational inquiry.

Analysis Vol. XX | Number XX | XX 2024 | 1–10 doi: https://doi.org/10.1093/analys/anae074

The account is not purely logical, referring to sentences accepted by a reasoner and descriptions that are still live possibilities for the reasoner. The key intuition is that deduction involves deriving and accepting sentences between the starting point of inquiry and the conclusion, which provides a measure of progress by ruling out live possibilities. I discuss and prove several theorems about sets of sentences that help state principles for efficient inquiry. I will also discuss a general directive for such inquiry in the sense of maximizing the average reduction in uncertainty regarding a resolution, an approach motivated by foraging theory. These principles and directives are not usable by actual reasoners. The aim is to provide an initial account of how well reasoners could do were they unbounded and knowledgeable or ignorant in certain ways, laying the foundation for future work.

2. *State descriptions and sets of sentences*

Let there be a name for every element in the domain $\mathfrak{D}$ of individuals and let there be a set of predicates for all properties and relations in which the individuals stand. Let Σ be the set of atomic sentences formed from the Cartesian product of the predicates and relations and constants for members of $\mathfrak{D}$. So, for example, if Rxy is a two-place predicate denoting some binary relation, then Σ will contain all atomic sentences formed from Rxy using all constants for members of $\mathfrak{D}$ (i.e. {Raa, Rab, Rba, …} $\in \Sigma$).

Intuitively a state description is a specification of who has which properties and who stands in which relations for some world. This can be formalized after Carnap 1950, letting a state description $\mathscr{S}$ be a set (herein, for a finite number of individuals, properties and relations) consisting of either the affirmation or the denial (exclusive) of every atomic sentence ϕ in Σ (e.g. for two-place predicate R and individuals named by a, b, c, …, {Raa, ~Rab, ~Rba, Rac, …} = $\mathscr{S}$). As Carnap notes, the construction of $\mathscr{S}$ requires that all the atomic sentences be logically independent of one another, and that no combination of atomic sentences entails another or its negation. Otherwise some $\mathscr{S}$ will be self-contradictory.

Call Δ the set of sentences accepted by the reasoner at time t discretized into steps.[2] Assume that Δ is consistent and can contain atomic or non-atomic sentences. As we reason, we use sentences in Δ to draw inferences and help close our inquiry. Assume further that, at the start of inquiry, not every sentence ϕ such that $\Delta \vdash \phi$ is in Δ, where the turnstile '$\vdash$' is interpreted as (say) derivable within a system of natural deduction (e.g. from Mates 1965).

At the start of reasoning, some subset of all state descriptions are ways the world might be that are relevant to the reasoner's inquiry.[3] Call this set $\mathfrak{S}$. For

2 Reasoners perhaps do not accept sentences but rather propositions; for logical convenience, I will speak of accepting sentences.

3 Determining relevance is an outstanding problem that space prevents me from discussing.

example, for predicate F and names a and b, at the start of inquiry $\mathfrak{S}$= {{Fa, Fb}, {~Fa, Fb}, {Fa, ~Fb}, {~Fa, ~Fb}}. Let $\mathfrak{S}_t(\Delta)$ be the set of state descriptions explicitly consistent with Δ at time t (i.e. there is no $\delta \in \Delta$ such that $\sim\delta \in \mathcal{S}$ for any $\mathcal{S} \in \mathfrak{S}(\Delta)$). Call this the set of live state descriptions. Using the example, if Δ = {Fa, Fa → ~Fb}, then $\mathfrak{S}_t(\Delta)$ = {{Fa, Fb}, {Fa, ~Fb}}. Below I will show that, as reasoning proceeds, state descriptions are often ruled out and the cardinality of $\mathfrak{S}_t(\Delta)$ will be less than or equal to the cardinality of $\mathfrak{S}_{t-1}(\Delta)$ at the previous step. I will not be addressing cases where the cardinality of $\mathfrak{S}_t(\Delta)$ grows in size, such as after retracting a rejected sentence (Harman 1986).

For every sentence ϕ such that $\phi \notin \Delta$ and $\Delta \vdash \phi$, the change in the size of $\mathfrak{S}_t(\Delta)$ if the sentence is accepted and Δ becomes $\Delta^* = \Delta \cup \phi$ can be assigned a number. These calculated changes impose a preorder on the sentences:

Theorem 1 (Preordering theorem). The set Φ of all sentences ϕ such that $\Delta \vdash \phi$ can be preordered.

Lemma. For any sentence ϕ such that $\Delta \vdash \phi$, the acceptance of ϕ would yield a change in the size of the set $\mathfrak{S}(\Delta)$ of state descriptions ≥ 0.

Proof of the lemma. Either ϕ is atomic or not. Suppose ϕ is atomic. Then accepting ϕ will rule out any state descriptions that contain $\sim\phi$. If no state descriptions contain $\sim\phi$, then the set $\mathfrak{S}$ stays the same size. Suppose ϕ is not atomic. The proof then proceeds by induction over the types of molecular sentences. Accepting the negation of some atomic sentence $\psi \in \Phi$ will rule out any state descriptions that contain ψ; otherwise the set $\mathfrak{S}$ stays the same size. Accepting the conjunction of two atomic sentences $\psi \in \Phi$ and $\chi \in \Phi$ will rule out any state descriptions that contain $\sim\psi$ or $\sim\chi$; otherwise the set $\mathfrak{S}$ stays the same size. Accepting the disjunction of two atomic sentences $\psi \in \Phi$ and $\chi \in \Phi$ will rule out any state descriptions that contain $\sim\psi$ and $\sim\chi$; otherwise the set $\mathfrak{S}$ stays the same size. The material conditional can be interchanged with disjunction, and the biconditional interchanged with two material conditionals. Let the universal quantifier sentence $\forall x\ \phi(x)$ be understood as a sentence ϕ such that ϕ contains the bound variable x. Replace every occurrence of x in ϕ with an arbitrary name for the members of $\mathfrak{D}$ and delete the quantifier (i.e. universal instantiation). Then ϕ is a quantifier-free atomic or molecular sentence and the proof proceeds as above. Repeat for every $d \in \mathfrak{D}$. The other quantified sentences can be interchanged with the universal quantifier in the usual way (i.e. $\sim\exists x \sim\phi(x) \equiv \forall x\ \phi(x)$ etc.). More complex sentences can be treated by recursion over the connectives.□

Proof of Theorem 1. A set X is preordered iff for all $x_i, x_j, x_k \in X$, $x_i \leq x_i$ and if $x_i \leq x_j$ and $x_j \leq x_k$ then $x_i \leq x_k$. Now the acceptance of any sentence AQ6
ϕ will yield a change in the size of the set of state descriptions ≥ 0 (from

the lemma). But then every sentence can be assigned a natural number equal to the change in the size of the set of state descriptions. Since any set of natural numbers can be preordered, Φ can be preordered by such an assignment. □

Theorem 1 states that, for all $\phi \in \Phi$, those ϕ can be placed in an order corresponding to the calculated change in the size of the set of state descriptions, with ties permitted. Using the above example, a reasoner could derive (among others) Fa, Fa → ~Fb or ~Fb from Δ; Fa and Fa → ~Fb are both assigned 0 and ~Fb is assigned 1.

Ruling out live state descriptions requires there to be some sentence ϕ derivable from Δ such that ϕ is explicitly inconsistent with some of those live state descriptions. This does not trivialize the above proof, however. The sentences in Δ are both explicitly consistent with the elements of $\mathfrak{S}(\Delta)$ and can imply a sentence ϕ that is inconsistent with them. Because only atomic sentences are contained in any $\mathcal{S} \in \mathfrak{S}(\Delta)$, only atomic sentences in Δ are available to rule out state descriptions; atomic sentences derivable from but not yet in Δ can then rule out elements of $\mathfrak{S}(\Delta)$.

The following theorem states that there exists some sentence such that accepting it would maximize the change in the size of the set of live state descriptions.

Theorem 2 (Maximum change theorem). There is a sentence ϕ such that $\Delta \vdash \phi$ and the cardinality of the live state description set before the inference ($|\mathfrak{S}_{t\text{-}1}(\Delta)|$) minus the cardinality of the live state description set after the inference ($|\mathfrak{S}_t(\Delta \cup \{\phi\})|$) is maximal, i.e. $\forall\psi\, \exists\phi(\max_{\psi}(|\mathfrak{S}_{t\text{-}1}(\Delta)| - |\mathfrak{S}_t(\Delta\cup\{\psi\})|) \leq \max_{\phi}(|\mathfrak{S}_{t\text{-}1}(\Delta)| - |\mathfrak{S}_t(\Delta\cup\{\phi\})|))$, where $\max_{\eta}(\cdot)$ means that there is some sentence η that maximizes the formula in $(\cdot)$.

Proof. The set of sentences Φ such that $\Delta \vdash \phi \in \Phi$ can be preordered (Theorem 1). But then there is some $\phi \in \Phi$ that is at the end of this preordering, the sentence that changes the size of the set of state descriptions more than or equal to any other $\phi \in \Phi$.□

Granted some set of live state descriptions, there is at each step during reasoning a best next sentence to accept (or many, if there are ties), the one(s) that maximally reduce the size of the set of possible state descriptions.

If the best next sentence is an atomic sentence and is not in Δ, half the remaining state descriptions are ruled out: either ϕ or $\sim\phi$ is derivable from Δ, which rules out all state descriptions that contain $\sim\phi$ or ϕ respectively. Sometimes this sentence will be derivable in one step, but at other times an inference chain, a finite sequence of derivations from the set of accepted sentences, will be required (such as moving from ϕ & ψ to ϕ and ψ). These inference chains can themselves be preordered by the number of state descriptions ruled out. Let Γ_1 be an initial set of sentences and Γ_i for step i is $\Gamma_{i-1} \cup \{\phi_{i-1}\}$ where $\Gamma_{i-1} \vdash \phi_{i-1}$.

Theorem 3 (Chain ordering theorem). The set G of all non-total inference chains can be preordered.

Proof. Let G be the set of non-total ordered n-tuples of sentences such that, for all $g \in G$, $\forall \phi_i \in g\ \Gamma_i \vdash \phi_i$ where Γ_i is constructed by starting with Γ_1, taking some ϕ_j such that $\Gamma_1 \vdash \phi_j$, adding ϕ_j to Γ_1 to yield Γ_2, and so on for a finite number of steps until some $\Gamma_n \vdash \phi_i$. Each n-tuple $g \in G$ is composed of sentences in the order in which they are entailed by the premisses and the earlier sentences in the sequence. The members of Γ_1 are not in the g. Non-total ordered n-tuples are such that $\forall g \in G$, g does not contain every sentence of every other $g \in G$. In virtue of their construction, each g is an inference chain and G is the set of all non-total inference chains.

Now, every set of sentences has a best next sentence (proved above). This property is also possessed by every ordered n-tuple of sentences since (for our purposes) an ordered n-tuple is a type of set whose elements are ordered. So too then does every g, since g is an ordered n-tuple of sentences formed from Γ_i. Consequently, $\forall g \in G$ can be pre-ordered by their best sentences. Let the maximal set be the proper subset of G each member of which is an ordered n-tuple that has the same, largest number assigned to it. Call this g*. By definition, some $\Gamma_i \vdash \phi_i^*$ where ϕ_i^* is the sentence that gives the number assigned to g*. Call Γ_i^* this set of sentences of Γ_i.

Note that there may be many such Γ_i^* and they may imply ϕ_i^* after different numbers of steps. Each Γ_i^* has some size, the length of the ordered n-tuple, which is a natural number. Every Γ_i^* then maps on to a rational number, the ratio of its assigned number to the number corresponding to the length of the n-tuple. But the rational numbers can be preordered. Hence, the members $g \in G$ can be preordered.□

The chain ordering theorem applies the notion of a best next sentence to inference chains. All chains will score some value $(m-n)/s$ for $\mathfrak{S}$ of size m before AQ7 the inference chain, size n after, and s steps in g. Let the maximal sentence be a sentence that initiates one of the chains at the end of the ordering. The chain ordering theorem will be used below to formulate a norm for reasoning.

3. *Norms of selection*

The preceding theorems demonstrate that sentences the reasoner could accept can be ordered by the number of state descriptions they rule out were the sentence to be accepted. These theorems can be used to state norms for selecting sentences.

One norm in reasoning is always to select a maximal sentence, a sentence that initiates some maximal chain. The maximal chain principle can be stated:

(MCP: Maximal chain principle) Of all the possible inference chains deducible from Δ, choose a chain at the end of the preorder established over G. Call this chain $g^\star$. Since $g^\star$ is an ordered n-tuple of sentences, there will be some sentence ϕ that is the first sentence in the n-tuple corresponding to $g^\star$. Select that ϕ.

Proof sketch. The proof follows from Theorem 3.

MCP says of the sentences in Δ, select (one of) the best sentence(s) derivable from Δ, namely a sentence ϕ such that ϕ is derivable from Δ in a single step and is the first step in a chain of inferences leading to ϕ^* at the end of $g^\star$. Intuitively, MCP says to choose the sentence that follows from the premisses and leads to the greatest overall reduction in uncertainty in the sense of a decrease in the size of the set of live state descriptions.

A more complex version of our working example above can illustrate MCP. Instead of two individuals, suppose there are four, a, b, c and d. At the start of inquiry, suppose Δ = {Fa, Fa v Fb v Fd → Fc, ~FavFc}, and so AQ8 $\mathfrak{S}(\Delta)$ = {{Fa, Fb, Fc, Fd}, {Fa, Fb, ~Fc, Fd}, {Fa, ~Fb, Fc, Fd}, {Fa, ~Fb, ~Fc, ~Fd}, {Fa, Fb, Fc, ~Fd}, {Fa, Fb, ~Fc, ~Fd}, {Fa, ~Fb, Fc, ~Fd}, {Fa, ~Fb, ~Fc, ~Fd}}. There are multiple sub-total inference chains from Δ, including those that insert unhelpful moves that reduce their scores; consider for simplicity just these two:

AQ9 Chain 1: Fa => Fa v Fb; Fa v Fb => Fa v Fb v Fd; Fa v Fb v Fd, Fa v Fb v Fd → Fc => Fc.

Chain 2: Fa, Fa → (~Fb & Fc) => (~Fb & Fc); (~Fb & Fc) => ~Fb; (~Fb & Fc) => Fc.

Chain 1 scores $(8 - 4)/3 = 1.333$ whereas chain 2 scores $(8 - 2)/3 = 2$. Of these two chains, chain 2 should be selected as it has the higher score.

The existence of a maximal sentence and the use of MCP is an idealization. Clearly, to prove the truth of some theorem, the reasoner should choose the maximal sentence, the step that resolves the uncertainty in the fewest number of steps. However, that maximal sentence may not be available because of the reasoner's ignorance or lack of confidence, cognitively bounds (e.g. finite attention or memory) or resource constraints (e.g. finite time). Consequently they may not be able to select the maximal sentence, leaving it unclear which step to take next.

4. *Norms of search*

Granted that the maximal sentence is often unavailable, how can reasoners inquire efficiently? The problems include which inference to make next, which chain to select next, when to switch chains of inference and when to switch premisses. Here I will present a principle to guide decisions about when to change premiss sets, providing an answer to the question of how to inquire efficiently using the above formalization of reducing uncertainty about the state of the world.

Consider the following marginal ranked information principle for changing premiss sets (after Charnov 1976, Stephens and Krebs 1986):

> (MRIP: Marginal ranked information principle) Given some premiss set P, reasoners should select inference chains in the rank order determined by the reverse preordering from P until the information intake rate drops below the expected information intake rate across all chains, at which time they should choose new premisses.

A reverse preorder is the preorder from Theorem 3 flipped from greatest to least. MRIP states that when the information intake from the current inference chain drops below the expected across all chains, the reasoner should change their premiss set to generate new inference chains. By a change in premiss set, I refer to either using other unused sentences in Δ or adding new sentences to Δ drawn from some other source (such as perception, background knowledge, testimony or others) and using those. The expectation across all chains should reflect the central tendency of the information intake; I will assume that the average is the relevant central tendency, but different contexts of inquiry may commend other measures.

What is the information intake? The preordering over inference chains is determined by the expected change in the size of the set of live state descriptions. As inference chains are followed, the size of the set of live state descriptions decreases, reducing uncertainty about the state description that resolves the current inquiry. This reduction in uncertainty is what is meant by 'information'. Information intake, then, refers to the change in the size of the set of live state descriptions resulting from deriving sentences, and the information intake rate refers to the information intake divided by the number of steps required to gather that information.

There are questions about the scope of the average information intake. Should every chain from all possible inquiries go into this average, and how should we compute such averages? In reply, unlike the next best sentence, which requires a global search across all next steps, these averages are computed from past progress, which reasoners have access to. Keeping track of past performance is also a topic of long interest in foraging, and numerous models have been developed for understanding how organisms might track these quantities (see e.g. Ollason 1980, McNamara and Houston 1985). These averages are also a measure of progress in an inquiry, and keeping track of how well an inquiry is proceeding is a reasonable thing for reasoners to do. Even when restricting the scope of the average to the current inquiry, should the average apply to just the current premisses or to all relevant premiss sets? Different averages will be relevant to different questions. For changing premisses, the average should be computed over the relevant premiss sets; for changing inference chains, the restriction to the current premiss set is justified. So different restrictions will apply depending on the question.

Granted some assumptions about reasoners and their inquiry, a proof can be sketched that MRIP is the most efficient norm for changing premiss sets. Suppose reasoners cannot gather information and search for premiss sets simultaneously and that outcomes of choices occur with unity probability. Suppose premiss sets can be categorized into types i, perhaps based on how much information they provide, their subject matter or some other feature. Let λ_i be the proportion of premiss sets of type i found by the reasoner. Let t_i be the time spent gathering information in time steps and let $g_i(t_i)$ denote the information gain function for premiss sets of type i. Assume $g_i(t_i)$ has three properties: (i) $g_i(0) = 0$; (ii) $g'i(0) > 0$; (iii) there exists some $\hat{t}$ such that $g''i(t) < 0$ for all $t \geq \hat{t}$. The first property is that at the start of information gathering there is no information gain, the second property is that the first step provides some information and the third property is true of any finite set of propositions that satisfy the chain ordering theorem in their reverse preordering and that stand in at least one strict inequality *inter alia* in their reverse preorder. Hence MRIP is evaluated over inference chains, not individual steps. Finally, let u_i be the average time in time steps between premiss sets of type i and $s_i(u_i)$, the information cost of searching per time step. Then the fully general long-run information intake rate R is given by

$$R = \frac{\sum_{i=1}^{n} \lambda_i g_i(t_i) - \sum_{i=1}^{n} \lambda_i w_i s_i(u_i)}{\sum_{i=1}^{n} \lambda_i u_i + \sum_{i=1}^{n} \lambda_i t_i},$$ AQ10

that is, over all patch types, the information gain minus the information cost due to time spent searching for premiss sets, divided by the sum of the time spent searching for premiss sets and the time spent gathering information.

We wish to find the optimal time at which to cease information gathering from a premiss set. I will make several idealizations to simplify this problem. First, suppose there is only one premiss set type. Then there is no summation over premiss set types and $\lambda = 1$:

$$R = \frac{g(t) - us(u)}{u + t}.$$

Note that this supposition implies that the rate at which premiss sets provide information, g(t), is the same across all premiss sets. Though clearly false, this implication can be conceptualized as some global average intake function g(t). Second, suppose that searching costs only time and not information. Then

$$R = \frac{g(t)}{u + t}.\dagger$$

We wish to maximize this quantity. To do so, we take the derivative with respect to time, which yields

$$R'(t) = \frac{g'(t)(u + t) - g(t)}{(u + t)^2}.$$

Now set the resulting equation to 0, which, given the properties of g(t) above, implies (see Stephens and Krebs 1986: 29 for proof)

$$g'(\hat{t})(u+\hat{t})-g(\hat{t})=0$$

for optimal leave time $\hat{t}$. By rearrangement of terms, this further implies

$$g'(\hat{t})=g(\hat{t})/(u+\hat{t}).$$

Note that this equality has the instantaneous information intake rate on the left-hand side and the average long-run information intake rate on the right (compare to † above). Consequently, the maximum information intake rate occurs when the instantaneous intake rate equals the long-run average. So reasoners cannot do better when reasoning from the current set of premisses, in the sense of improving their information intake rate, than when the current information intake equals the average. When the information intake rate from the current premisses equals that average, reasoners should decide to change premisses and search for new ones, proving MRIP.

A few comments on MRIP. MRIP states that after each inference chain using Δ, reasoners evaluate their information intake rate and, if that rate falls below the average, they should shift to new premisses, even if the current premiss set remains promising. However, MRIP assumes the goal is to maximize the long-run information intake, whereas the real goal is to resolve our inquiries. So there are contexts in which MRIP should be ignored. Suppose I am attempting to deduce the truth of some theorem. As I near the end of the proof, there may only be smaller amounts of information left to gather. In that case, instead of switching when my information intake rate drops low, I should ignore MRIP and persevere to see out the end of my inquiry. Similar conclusions follow regarding inductive and other types of inquiry.

5. *Conclusion*

In this article, I have presented a normative analysis of efficient inquiry. Granted general conditions of ignorance, normative principles can be stated for guiding inquiry. The statement of these novel principles, such as the maximal chain principle for choosing chains of inferences or the marginal ranked information principle for when to select new premisses, demonstrates that guidance can be provided for how to efficiently conduct one's inquiry, advancing our understanding of reasoning. The principles stated are generally outside the cognitive capacities of actual reasoners. Nonetheless, they provide a first-pass benchmark for understanding how (perhaps ideal) reasoners should make such decisions during inquiry. The discussion of these principles also highlighted important constraints on when to change premiss sets, as nearing the end of inquiry demands sticking with our premisses instead of changing them. Next steps include adapting the account to measures for continuous time or for open-ended state descriptions, a shift to a more powerful logical formalisms such as Hintikka's model sets (Hintikka 1973) or dynamic epistemic logic (van Benthem 2011) and formulating feasible norms for actual reasoners. More broadly, outside of consistency, validity and related norms, how one should guide one's reasoning is under-addressed in philosophy. By characterizing reasoning as a process of gathering information and drawing formal analogies to foraging, my account offers opportunities to better understand how to guide our reasoning, including which inferences to make, which premisses to choose and when to switch between the different activities involved in reasoning. Future work will draw connections to work on bounded rationality, investigate the nature of the search space in reasoning and explore creativity in reasoning such as uncovering next steps or new premisses.[4]

University of Pennsylvania
USA
University of Antwerp
Belgium

AQ11

References

Alchourrón, C.E., P. Gärdenfors and D. Makinson. 1985. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic* 50: 510–30.

Carnap, R. 1950. *Logical Foundations of Probability*. Chicag, IL: University of Chicago Press.

Charnov, E.L. 1976. Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9: 129–36.

Harman, G. 1986. *Change in View: Principles of Reasoning*. Cambridge, MA: MIT Press.

Hintikka, J. 1973. *Logic, Language-Games and Information: Kantian Themes in the Philosophy of Logic*. Oxford: Clarendon Press.

Mates, B. 1965. *Elementary Logic*, Oxford University Press. AQ12

McNamara, J.M. and A.I. Houston. 1985. Optimal foraging and learning. *Journal of Theoretical Biology* 117: 231–49.

Ollason, J. 1980. Learning to forage – optimally? *Theoretical Population Biology* 18: 44–56.

Spohn, W. 2012. *The Laws of Belief: Ranking Theory and Its Philosophical Applications*. AQ13 Oxford University Press.

Stephens, D.W. and J.R. Krebs. 1986. *Foraging Theory*. Princeton, NJ: Princeton University Press.

van Benthem, J. 2011. *Logical Dynamics of Information and Interaction*. Cambridge AQ14 University Press.

4 Many thanks to Chris Peacocke for numerous discussions over the years, to Corey Maley for helpful comments and to Kyle Landrum for sanity checks.