version 102422NY

The Argument for God's Existence

Selmer Bringsjord

Rensselaer AI & Reasoning (RAIR) Lab
Department of Cognitive Science
Department of Computer Science
Lally School of Management & Technology
Rensselaer Polytechnic Institute (RPI)
Troy, New York 12180 USA

IFLAI2
10/24/2022

version 102422NY

The Argument for God's Existence from AI

Selmer Bringsjord

Rensselaer AI & Reasoning (RAIR) Lab
Department of Cognitive Science
Department of Computer Science
Lally School of Management & Technology
Rensselaer Polytechnic Institute (RPI)
Troy, New York 12180 USA

IFLAI2
10/24/2022

The Ontological/Modal Argument Meets AI

The Ontological/Modal Argument Meets AI

Scott's version of Gödel verified!

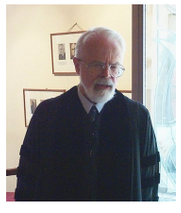
Gödel's version of Gödel falsified!

Anselm

Descartes

Leibniz

Gödel/Adams



ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-93

Automating Gödel's Ontological Proof of God's Existence with Higher-order Automated Theorem Provers

Christoph Benzmüller[1] and **Bruno Woltzenlogel Paleo**[2]

Abstract. Kurt Gödel's ontological argument for God's existence has been formalized and automated on a computer with higher-order automated theorem provers. From Gödel's premises, the computer proved: necessarily, there exists God. On the other hand, the theorem provers have also confirmed prominent criticism on Gödel's ontological argument, and they found some new results about it.

The background theory of the work presented here offers a novel perspective towards a *computational theoretical philosophy*.

1 INTRODUCTION

Kurt Gödel proposed an argumentation formalism to prove the existence of God [23, 30]. Attempts to prove the existence (or non-existence) of God by means of abstract, ontological arguments are an old tradition in western philosophy. Before Gödel, several prominent philosophers, including St. Anselm of Canterbury, Descartes and Leibniz, have presented similar arguments. Moreover, there is an impressive body of recent and ongoing work (cf. [31, 19, 18] and the references therein). Ontological arguments, for or against the existence of God, illustrate well an essential aspect of metaphysics: some (necessary) facts for our existing world are deduced by purely a priori, analytical means from some abstract definitions and axioms.

What motivated Gödel as a logician was the question, whether it

A1	Either a property or its negation is positive, but not both:	$\forall\phi[P(\neg\phi) \equiv \neg P(\phi)]$
A2	A property necessarily implied by a positive property is positive:	$\forall\phi\forall\psi[(P(\phi) \wedge \Box\forall x[\phi(x) \supset \psi(x)]) \supset P(\psi)]$
T1	Positive properties are possibly exemplified:	$\forall\phi[P(\phi) \supset \Diamond\exists x\phi(x)]$
D1	A *God-like* being possesses all positive properties:	$G(x) \equiv \forall\phi[P(\phi) \supset \phi(x)]$
A3	The property of being God-like is positive:	$P(G)$
C	Possibly, God exists:	$\Diamond\exists x G(x)$
A4	Positive properties are necessarily positive:	$\forall\phi[P(\phi) \supset \Box P(\phi)]$
D2	An *essence* of an individual is a property possessed by it and necessarily implying any of its properties:	$\phi\ ess.\ x \equiv \phi(x) \wedge \forall\psi(\psi(x) \supset \Box\forall y(\phi(y) \supset \psi(y)))$
T2	Being God-like is an essence of any God-like being:	$\forall x[G(x) \supset G\ ess.\ x]$
D3	*Necessary existence* of an individ. is the necessary exemplification of all its essences:	$NE(x) \equiv \forall\phi[\phi\ ess.\ x \supset \Box\exists y\phi(y)]$
A5	Necessary existence is a positive property:	$P(NE)$
T3	Necessarily, God exists:	$\Box\exists x G(x)$

Figure 1. Scott's version of Gödel's ontological argument [30].

Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)

The Inconsistency in Gödel's Ontological Argument: A Success Story for AI in Metaphysics

Christoph Benzmüller*
Freie Universität Berlin & Stanford University
c.benzmueller@gmail.com

Bruno Woltzenlogel Paleo
Australian National University
bruno.wp@gmail.com

Abstract

This paper discusses the discovery of the inconsistency in Gödel's ontological argument as a success story for artificial intelligence. Despite the popularity of the argument since the appearance of Gödel's manuscript in the early 1970's, the inconsistency of the axioms used in the argument remained unnoticed until 2013, when it was detected automatically by the higher-order theorem prover

istence of God, illustrate well an essential aspect of metaphysics: some (necessary) facts for our existing world are deduced by purely a priori, analytical means from some abstract definitions and axioms.

What motivated Gödel as a logician was the question, whether it

A5	Necessary existence is a positive property:	$P(NE)$
T3	Necessarily, God exists:	$\Box\exists x G(x)$

Figure 1. Scott's version of Gödel's ontological argument [30].

on the proof [Fuhrmann, 2016].

The in-depth analysis presented here substantially extends previous computer-assisted studies of Gödel's ontological argument. Similarly to the related work [Benzmüller and Woltzenlogel-Paleo, 2013a; 2014] the analysis has been conducted with automated theorem provers for classical higher-order logic (HOL; cf. [Andrews, 2014] and the references therein), even though Gödel's proof is actually formulated in higher-order *modal* logic (HOML; cf. [Muskens, 2006]

ECAI 2014
T. Schaub et al. (Eds.)

doi:10.3233/978-1-61499-419-0-93

Automating Gödel's Ontological Proof of God's Existence with Higher-order Automated Theorem Provers

Christoph Benzmüller[1] and **Bruno Woltzenlogel Paleo**[2]

Abstract. Kurt Gödel's ontological argument for God's existence has been formalized and automated on a computer with higher-order automated theorem provers. From Gödel's premises, the computer proved: necessarily, there exists God. On the other hand, the theorem provers have also confirmed prominent criticism on Gödel's ontological argument, and they found some new results about it.

The background theory of the work presented here offers a novel perspective towards a *computational theoretical philosophy*.

1 INTRODUCTION

Kurt Gödel proposed an argumentation formalism to prove the existence of God [23, 30]. Attempts to prove the existence (or non-existence) of God by means of abstract, ontological arguments are an old tradition in western philosophy. Before Gödel, several prominent philosophers, including St. Anselm of Canterbury, Descartes and Leibniz, have presented similar arguments. Moreover, there is an impressive body of recent and ongoing work (cf. [31, 19, 18] and the references therein). Ontological arguments, for or against the existence of God, illustrate well an essential aspect of metaphysics: some (necessary) facts for our existing world are deduced by purely a priori, analytical means from some abstract definitions and axioms.

What motivated Gödel as a logician was the question, whether it

A1	Either a property or its negation is positive, but not both:	$\forall\phi[P(\neg\phi) \equiv \neg P(\phi)]$
A2	A property necessarily implied by a positive property is positive:	$\forall\phi\forall\psi[(P(\phi) \wedge \Box\forall x[\phi(x) \supset \psi(x)]) \supset P(\psi)]$
T1	Positive properties are possibly exemplified:	$\forall\phi[P(\phi) \supset \Diamond\exists x\phi(x)]$
D1	A *God-like* being possesses all positive properties:	$G(x) \equiv \forall\phi[P(\phi) \supset \phi(x)]$
A3	The property of being God-like is positive:	$P(G)$
C	Possibly, God exists:	$\Diamond\exists x G(x)$
A4	Positive properties are necessarily positive:	$\forall\phi[P(\phi) \supset \Box P(\phi)]$
D2	An *essence* of an individual is a property possessed by it and necessarily implying any of its properties:	$\phi\ ess.\ x \equiv \phi(x) \wedge \forall\psi(\psi(x) \supset \Box\forall y(\phi(y) \supset \psi(y)))$
T2	Being God-like is an essence of any God-like being:	$\forall x[G(x) \supset G\ ess.\ x]$
D3	*Necessary existence* of an individ. is the necessary exemplification of all its essences:	$NE(x) \equiv \forall\phi[\phi\ ess.\ x \supset \Box\exists y\phi(y)]$
A5	Necessary existence is a positive property:	$P(NE)$
T3	Necessarily, God exists:	$\Box\exists x G(x)$

Figure 1. Scott's version of Gödel's ontological argument [30].

The Ontological/Modal Argument Meets AI

Scott's version of Gödel verified!

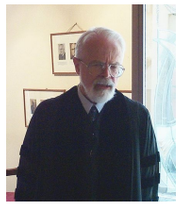
Gödel's version of Gödel falsified!

Anselm

Descartes

Leibniz

Gödel/Adams



?

The Ontological/Modal Argument Meets AI

Scott's version of Gödel verified!

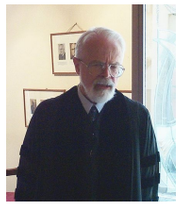
Gödel's version of Gödel falsified!

Anselm

Descartes

Leibniz

Gödel/Adams



The Ontological/Modal Argument Meets AI

Scott's version of Gödel verified!

Gödel's version of Gödel falsified!

?

Anselm

Descartes

Leibniz

Gödel/Adams

This is argumentation for God's existence that *uses* AI; it's not *based* on AI.

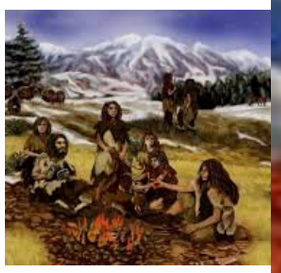
This is argumentation for God's existence that *uses* AI; it's not *based* on AI.

Today, The Argument for God's Existence *from* AI …

The Canyon of Discontinuity (or Darwin's Dread)

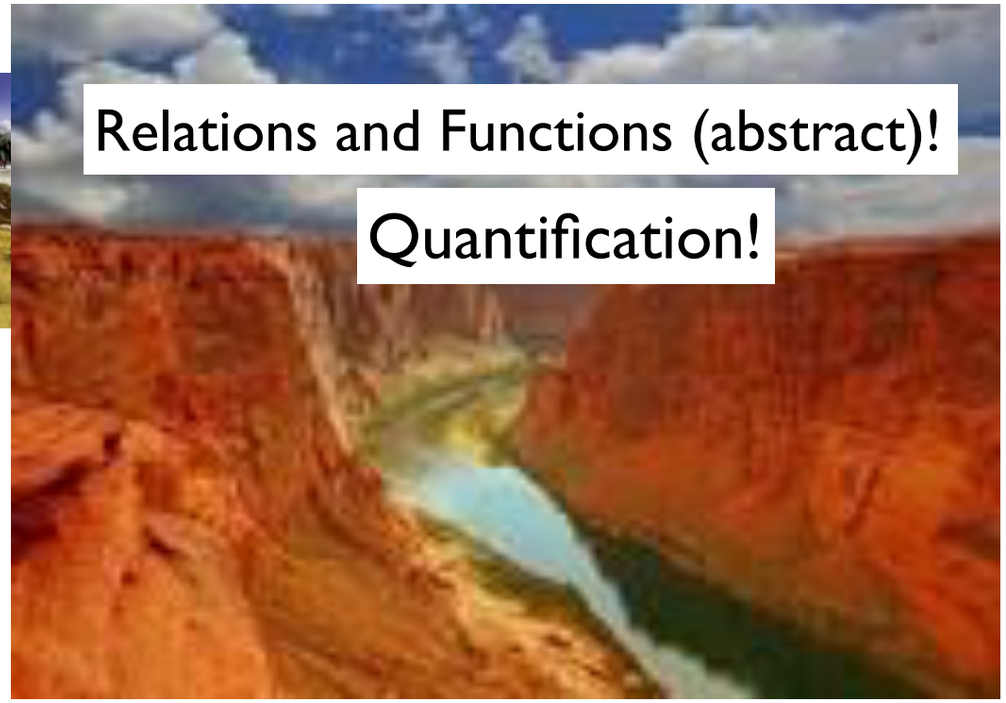
gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)



gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)

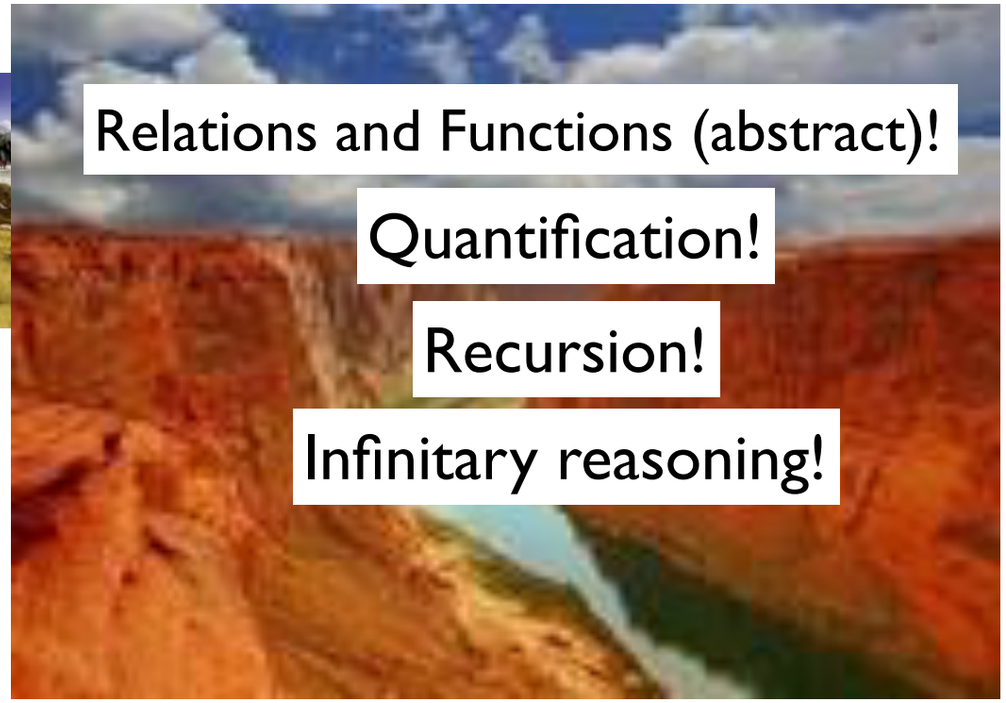


gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)

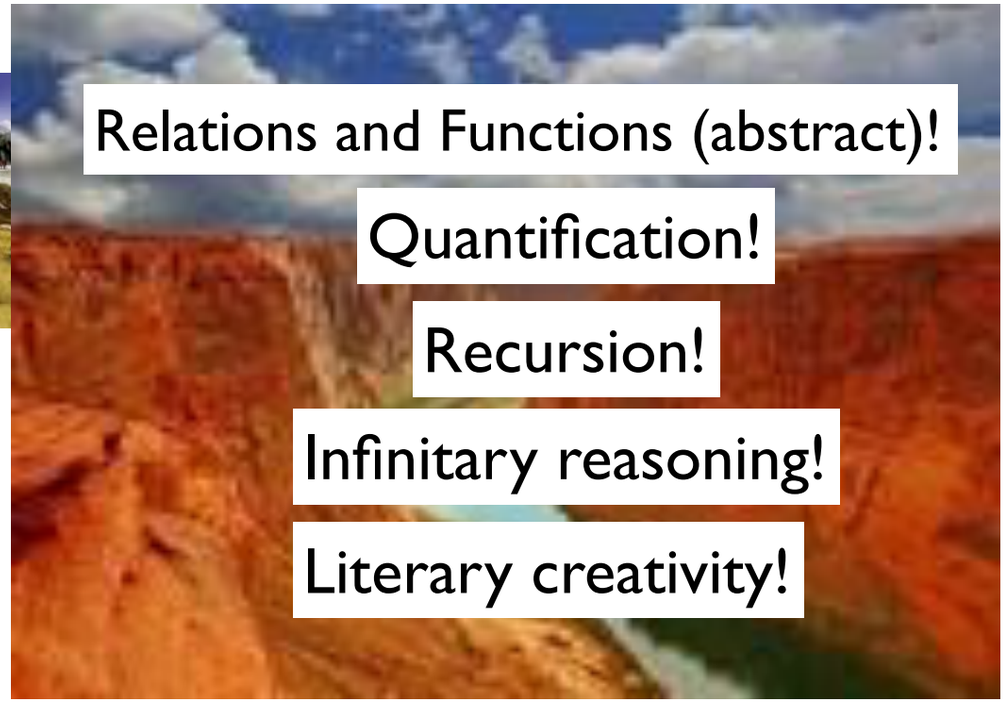
gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)



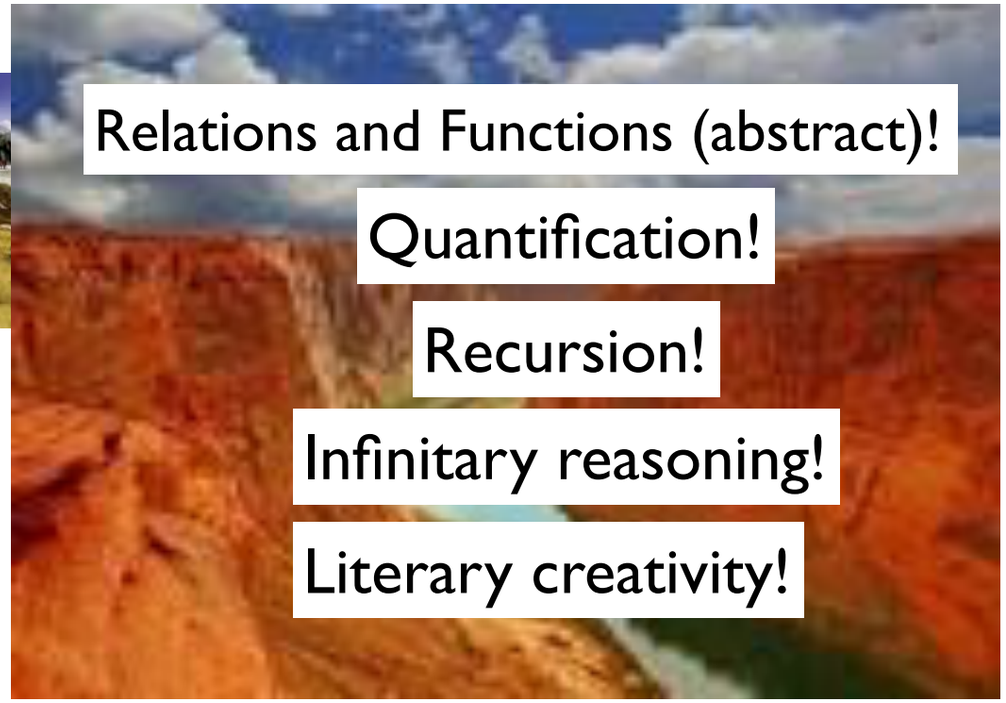
gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)



gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)

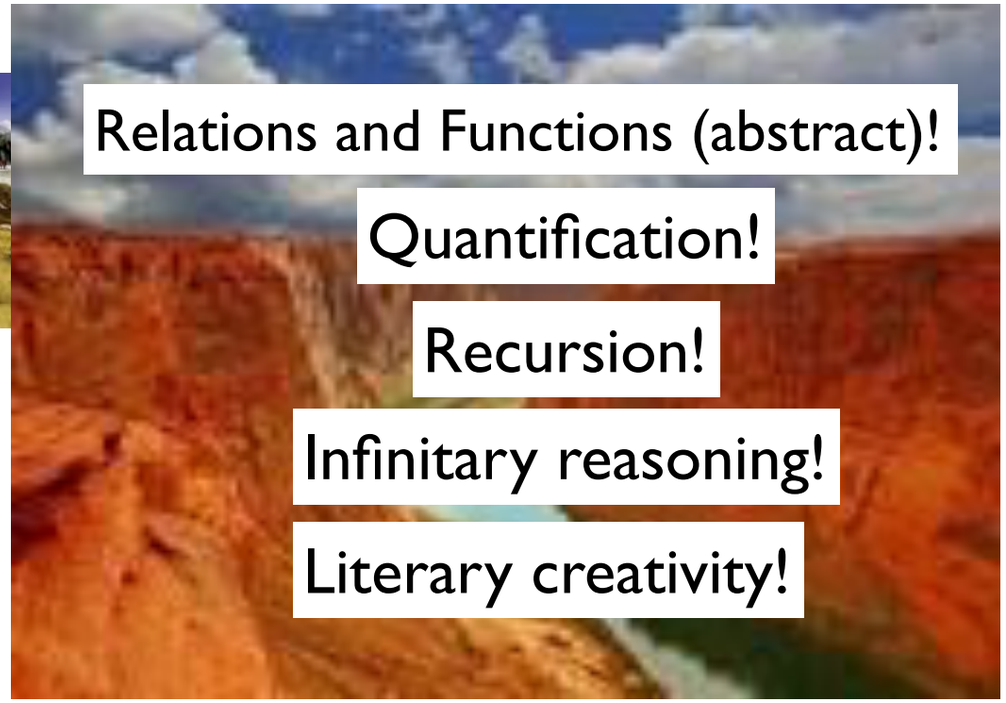


gradual increase in cognitive powers

P^{∞}

The Canyon of Discontinuity (or Darwin's Dread)

P^{HG}



P^{∞}

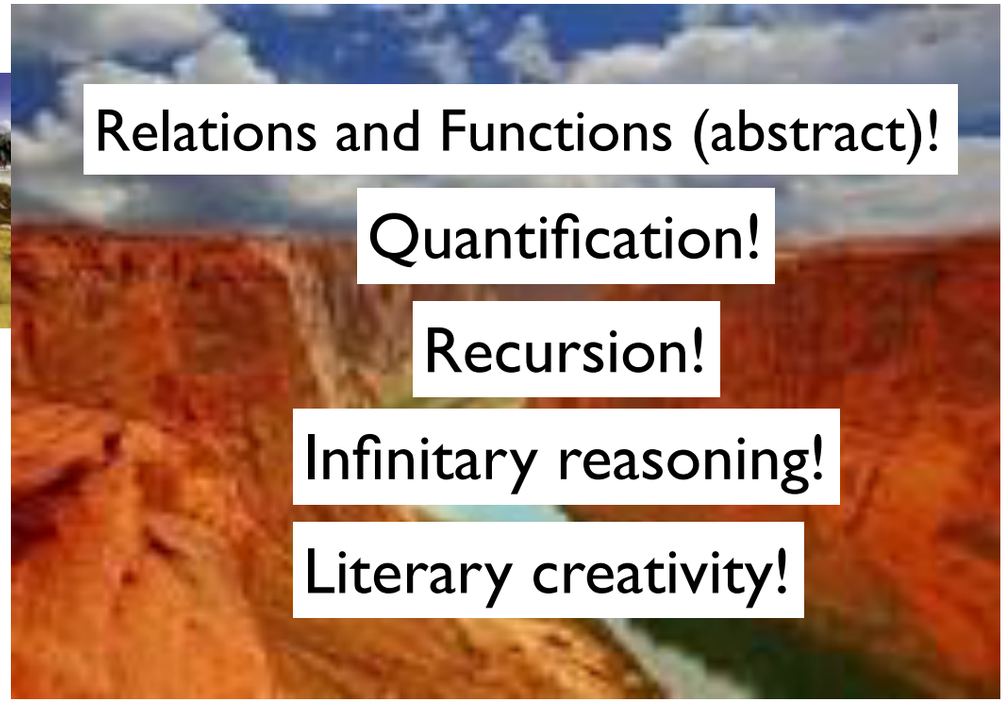
gradual increase in cognitive powers

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

gradual increase in cognitive powers



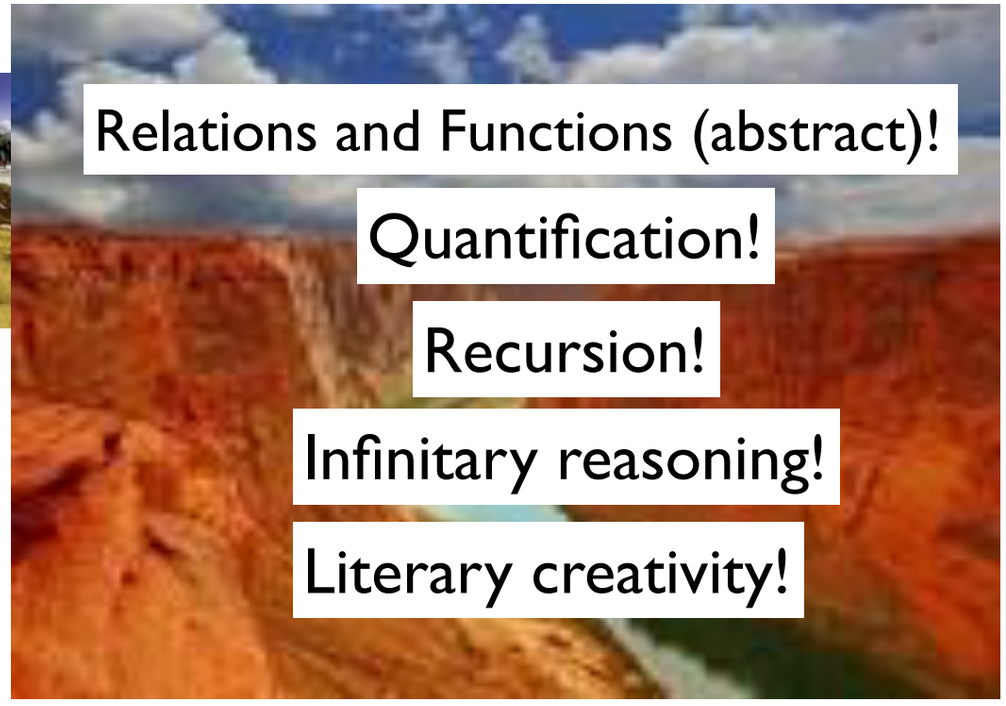
P^{∞}

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

gradual increase in cognitive powers



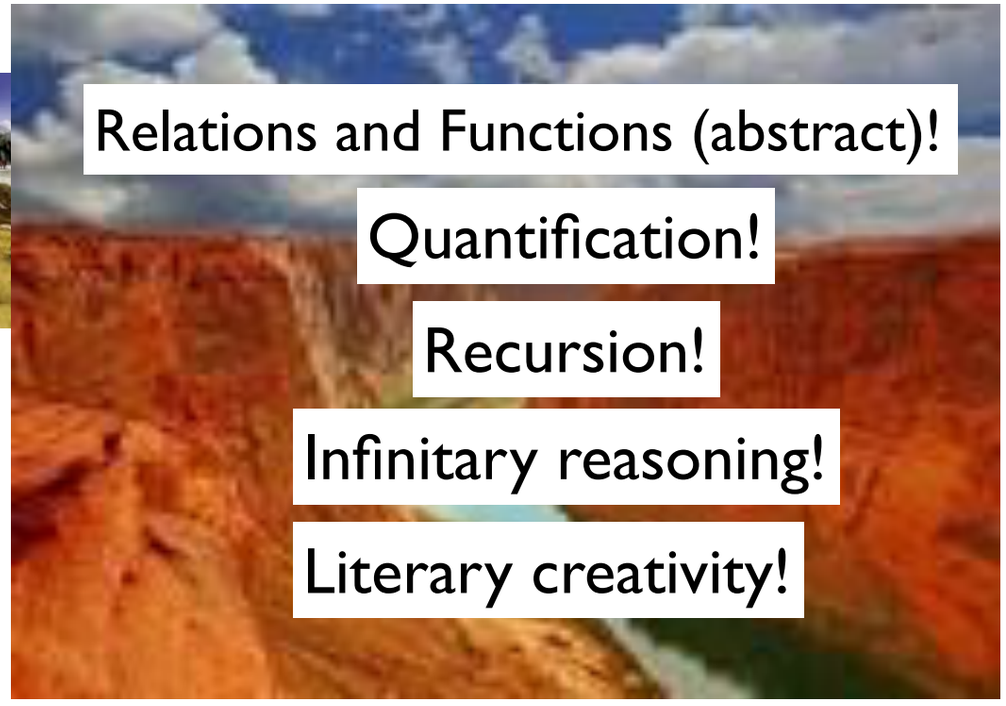
P^{∞}

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

gradual increase in cognitive powers



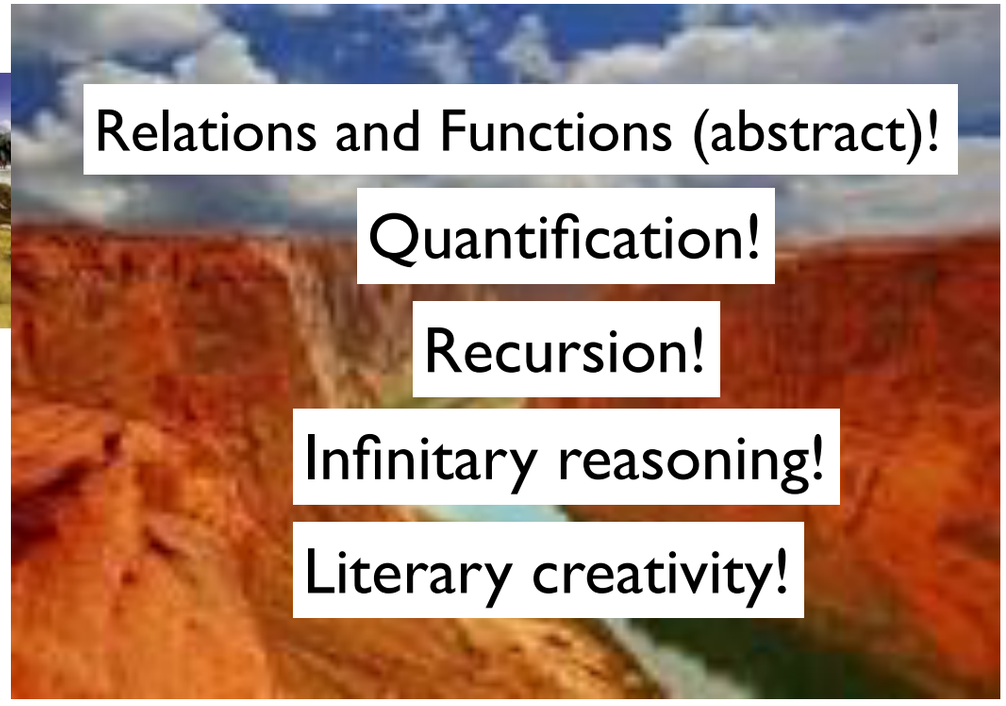
P^{∞}

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

gradual increase in cognitive powers



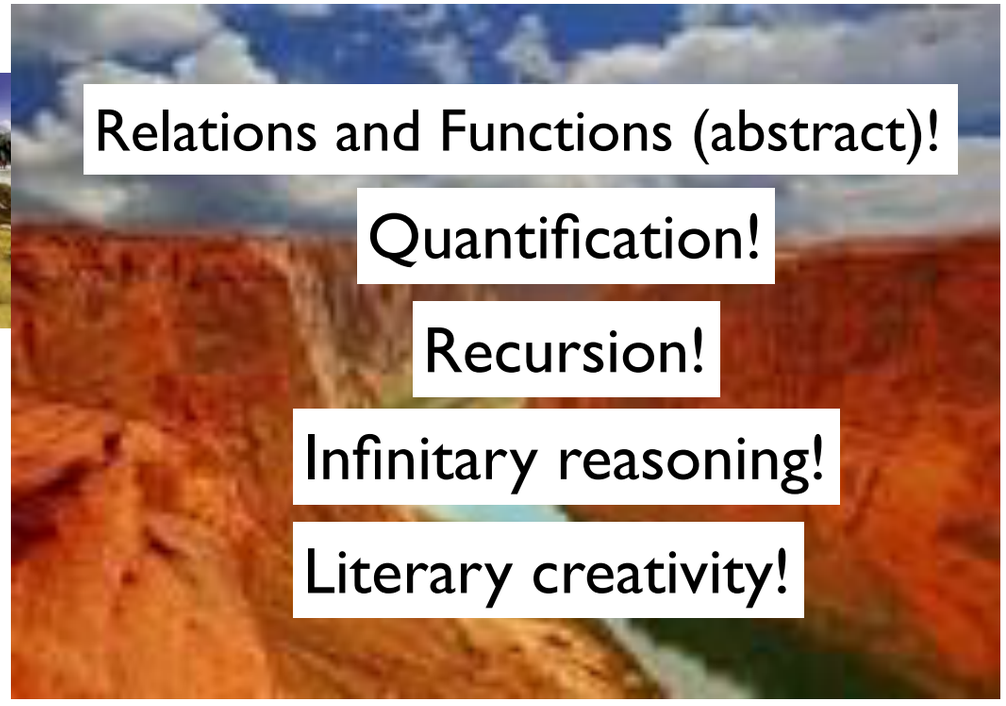
P^{∞}

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

gradual increase in cognitive powers



P^{∞}

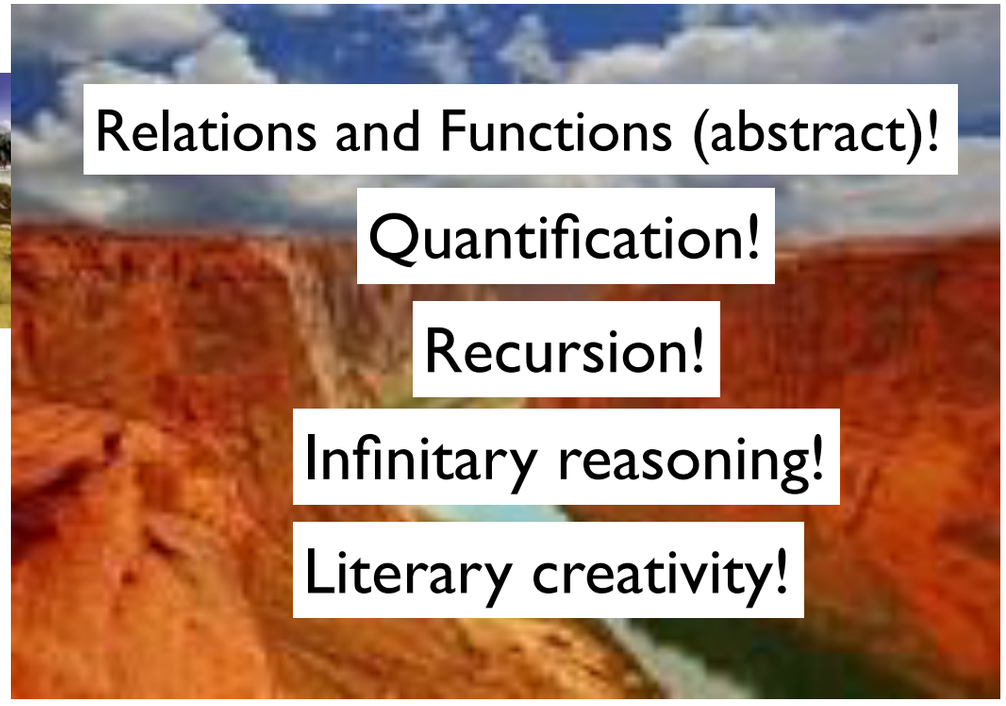
What about neanderthals?

The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}

P^{HG}

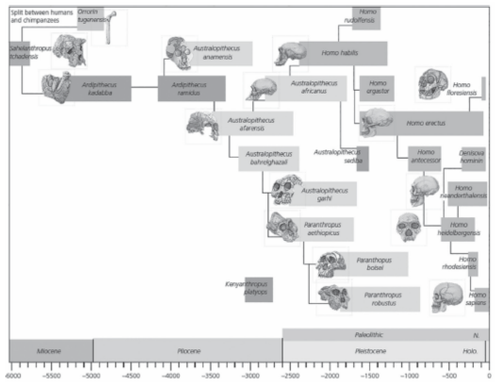
gradual increase in cognitive powers



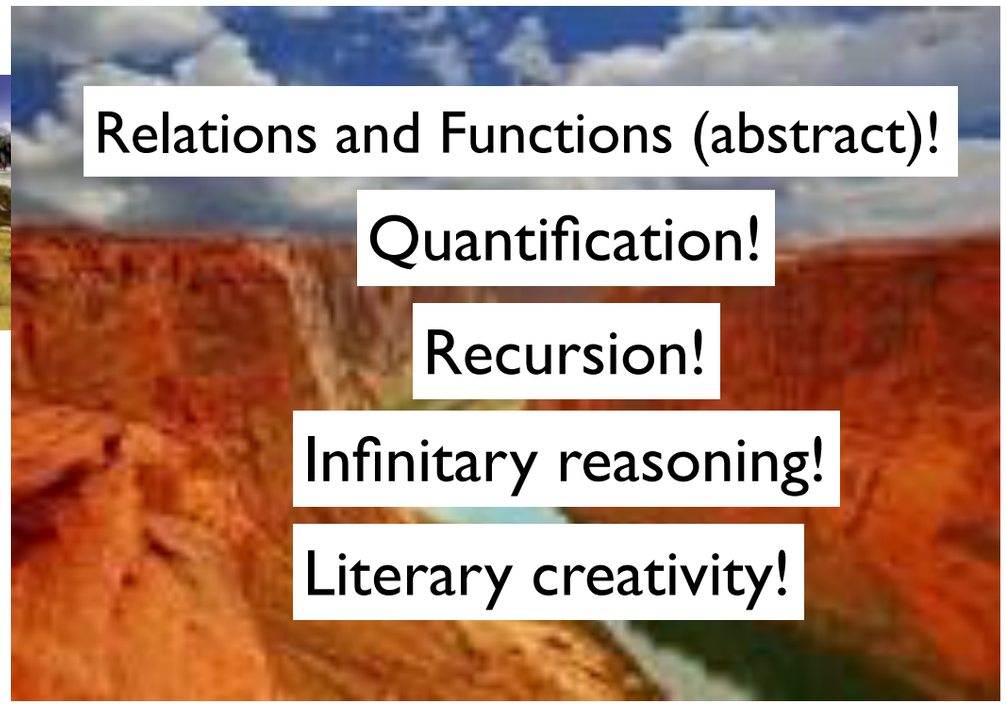
P^{∞}

What about neanderthals?

The Canyon of Discontinuity (or Darwin's Dread)



P^{AI}



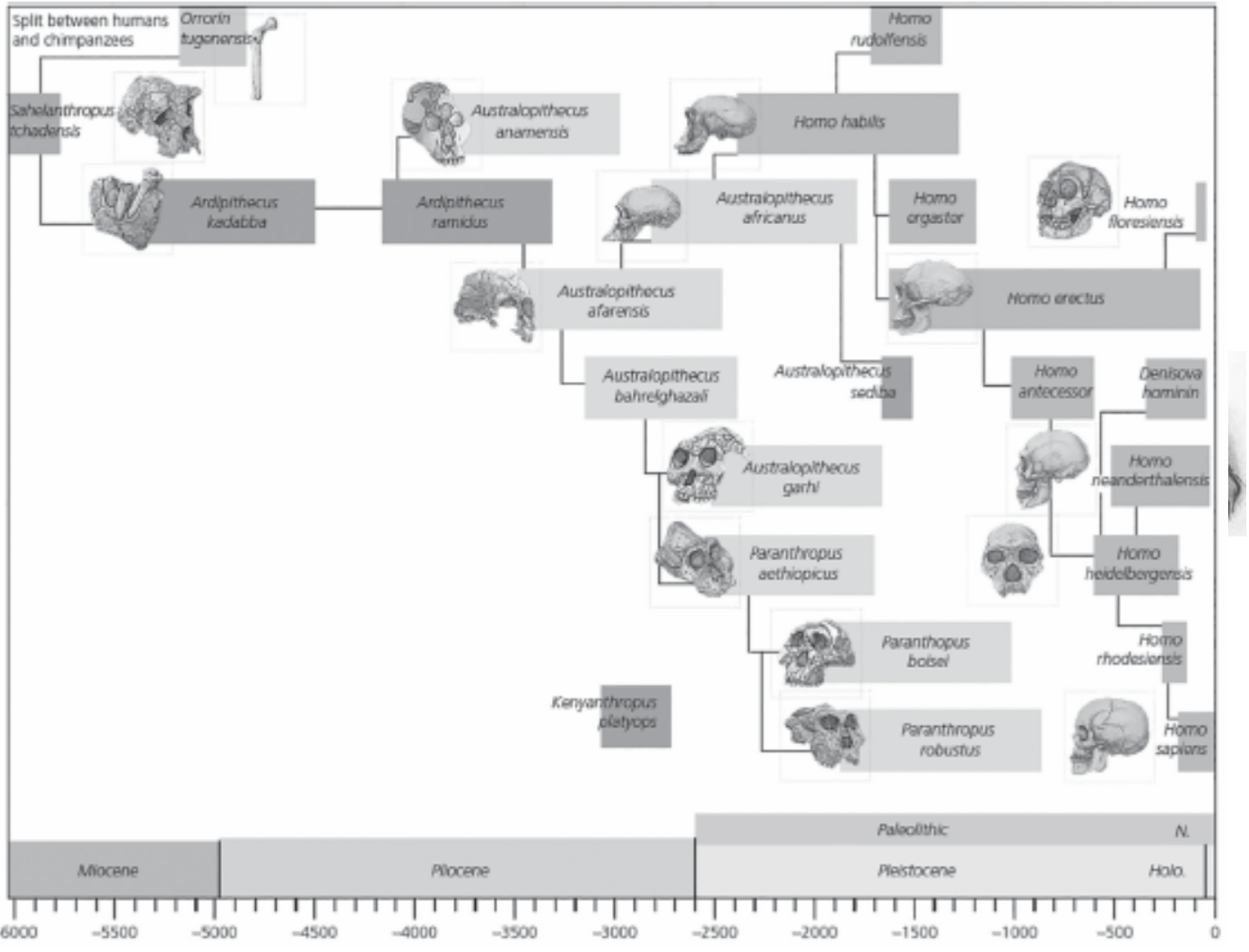
P^{HG}

gradual increase in cognitive powers

P^{∞}

What about neanderthals?

The Canyon of Discontinuity

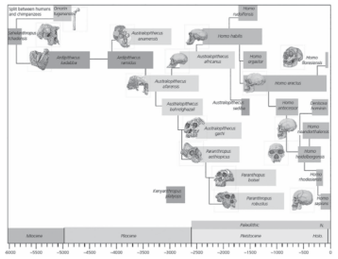


gra
cc

What about neanderthals?

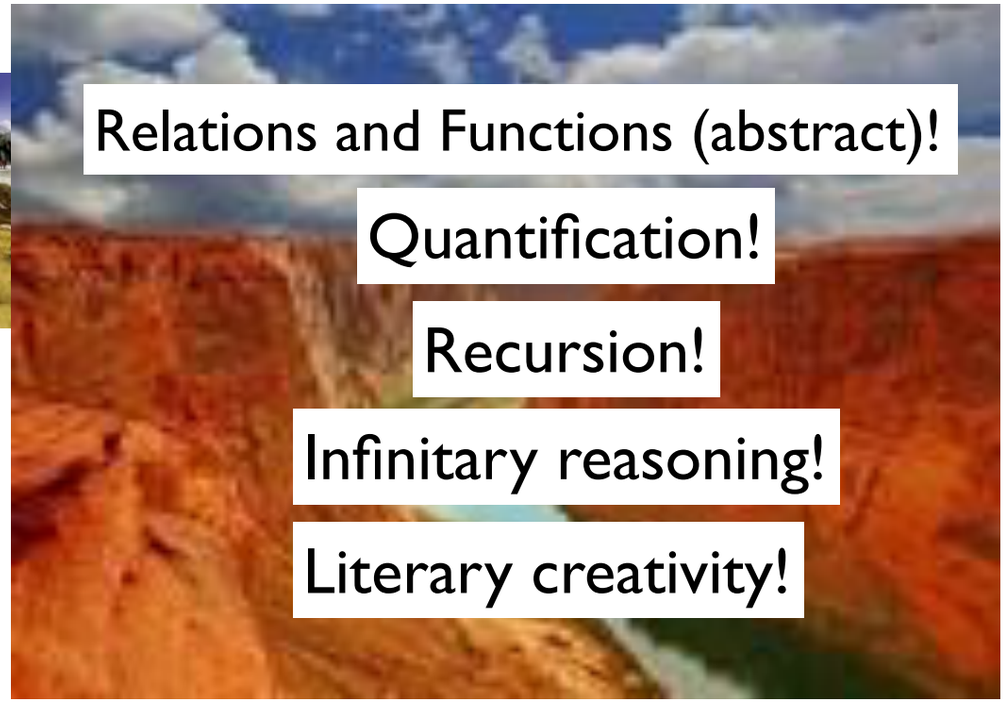
The Canyon of Discontinuity (or Darwin's Dread)

P^{AI}



P^{HG}

gradual increase in cognitive powers



P^{∞}

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/....	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ (where P^{AI} is a limit on the cognitive power of AI), and AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P^{\text{HG}})$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ (where P^{AI} is a limit on the cognitive power of AI), and AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P^{\text{HG}})$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ (where P^{AI} is a limit on the cognitive power of AI), and AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P^{\text{HG}})$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

Set "subtraction": The extreme cognitive powers, but none of the routine ones. So, that which we share with hunter-gatherer activity & the lower animals is irrelevant.

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

Songbirds

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

Songbirds

the birds sing differentially based on what song they hear

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$\Delta_a^B := (ab)^n$ and $a^n b^n$ are *both* starling-decidable

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$\Delta_a^B := (ab)^n$ and $a^n b^n$ are *both* starling-decidable

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } \textit{both} \text{ starling-decidable}$$

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } \textit{both} \text{ starling-decidable}$$

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } \textit{both} \text{ starling-decidable}$$

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } \textit{both} \text{ starling-decidable}$$

Some elements of some formalized nonhuman-animal behavior overlap some elements of some formalized human behavior.

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } both \text{ starling-decidable}$$

Some elements of some formalized nonhuman-animal behavior overlap some elements of some formalized human behavior. $\Delta_a^B \cap \Delta_h^B \neq \emptyset!$

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } both \text{ starling-decidable}$$

Some elements of some formalized nonhuman-animal behavior overlap some elements of some formalized human behavior. $\Delta_a^B \cap \Delta_h^B \neq \emptyset!$

After all, birds can see; humans can too; but but this is not relevant to my present purposes, which is to focus on:

Gentner, T., Fenn, K., Margoliash, D. & Nusbaum, H. (2006) "Recursive Syntactic Pattern Learning by Songbirds" *Nature* **440**: 1204–1207.

Humans

humans can easily enough decide these languages

$$\Delta_h^B := (ab)^n \text{ and } a^n b^n \text{ are (trivially) both human-decidable}$$

Songbirds

the birds sing differentially based on what song they hear

a "rattle" coded as *a*; a "warble" as *b*

$$\Delta_a^B := (ab)^n \text{ and } a^n b^n \text{ are } \textit{both} \text{ starling-decidable}$$

Some elements of some formalized nonhuman-animal behavior overlap some elements of some formalized human behavior. $\Delta_a^B \cap \Delta_h^B \neq \emptyset!$

After all, birds can see; humans can too; but but this is not relevant to my present purposes, which is to focus on:

$$P^\infty - P^{\mathrm{HG}} / P^\infty - P^{\mathrm{AI}}$$

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ and, where P^{AI} is a limit on the cognitive power of AI, AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P)$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ and, where P^{AI} is a limit on the cognitive power of AI, AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P)$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

See …

See …

doi: 10.1017/S0140525X08003543

Darwin's mistake: Explaining the discontinuity between human and nonhuman minds

Derek C. Penn
Department of Psychology, University of California–Los Angeles, Los Angeles, CA 90095; Cognitive Evolution Group, University of Louisiana, Lafayette, LA 70504
dcpenn@ucla.edu
http://reasoninglab.psych.ucla.edu/
http://www.cognitiveevolutiongroup.org/

Keith J. Holyoak
Department of Psychology, University of California–Los Angeles, Los Angeles, CA 90095
holyoak@lifesci.ucla.edu
http://reasoninglab.psych.ucla.edu/

Daniel J. Povinelli
Cognitive Evolution Group, University of Louisiana, Lafayette, LA 70504
ceg@louisiana.edu
http://www.cognitiveevolutiongroup.org/

Abstract: Over the last quarter century, the dominant tendency in comparative cognitive psychology has been to emphasize the similarities between human and nonhuman minds and to downplay the differences as "one of degree and not of kind" (Darwin 1871). In the present target article, we argue that Darwin was mistaken: the profound biological continuity between human and nonhuman animals masks an equally profound discontinuity between human and nonhuman minds. To wit, there is a significant discontinuity in the degree to which human and nonhuman animals are able to approximate the higher-order, systematic, relational capabilities of a physical symbol system (PSS) (Newell 1980). We show that this symbolic-relational discontinuity pervades nearly every domain of cognition and runs much deeper than even the spectacular scaffolding provided by language or culture alone can explain. We propose a representational-level specification as to where human and nonhuman animals' abilities to approximate a PSS are similar and where they differ. We conclude by suggesting that recent symbolic-connectionist models of cognition shed new light on the mechanisms that underlie the gap between human and nonhuman minds.

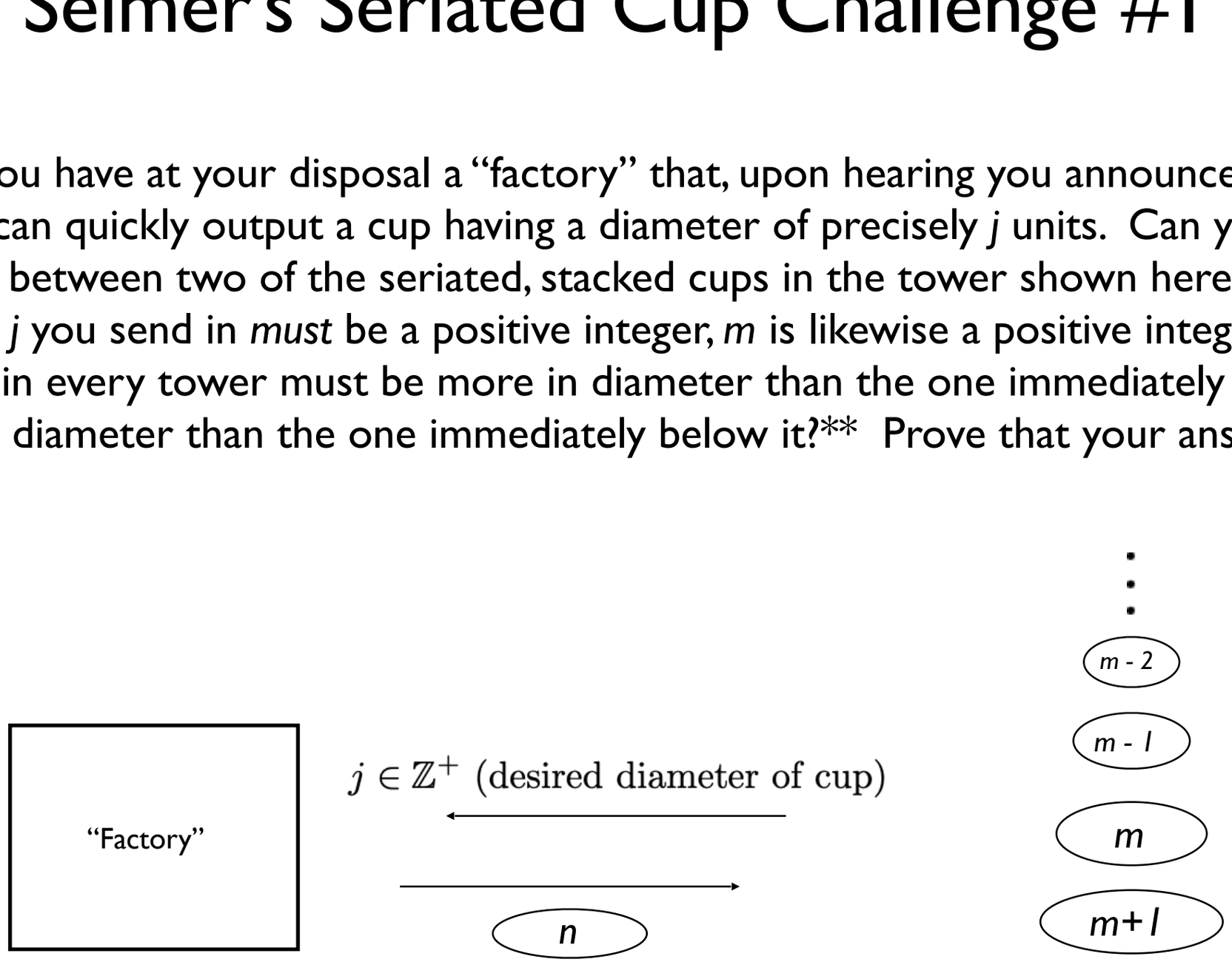
Keywords: analogy; animal cognition; causal learning; connectionism; Darwin; discontinuity; evolution; human mind; language; language of thought; physical symbol system; reasoning; same-different; theory of mind

1. Introduction

Human animals – and no other – build fires and wheels, diagnose each other's illnesses, communicate using symbols, navigate with maps, risk their lives for ideals, collaborate with each other, explain the world in terms of hypothetical causes, punish strangers for breaking rules, imagine impossible scenarios, and teach each other how to do all of the above. At first blush, it might appear obvious that human minds are qualitatively different from those of every other animal on the planet. Ever since Darwin, however, the dominant tendency in comparative cognitive psychology has been to emphasize the continuity between human and nonhuman minds and to downplay the differences as "one of degree and not of kind" (Darwin 1871). Particularly in the last quarter century, many prominent comparative researchers have claimed that the traditional hallmarks of human cognition – for example, complex tool use, grammatically structured language, causal-logical reasoning, mental state attribution, metacognition, analogical inferences, mental time travel, culture, and so on – are not nearly as unique as we once thought (see, e.g., Bekoff et al. 2002; Call 2006; Clayton et al. 2003; de Waal & Tyack 2003; Matsuzawa 2001; Pepperberg 2002; Rendell & Whitehead 2001; Savage-Rumbaugh et al. 1998; Smith et al. 2003; Tomasello et al. 2003a). Pepperberg (2005, p. 469) aptly sums up the comparative consensus as follows: "for over 35 years, researchers have been demonstrating through tests both in the field and in the laboratory that the capacities of nonhuman animals to solve complex problems form a continuum with those of

Selmer's Seriated Cup Challenge #1

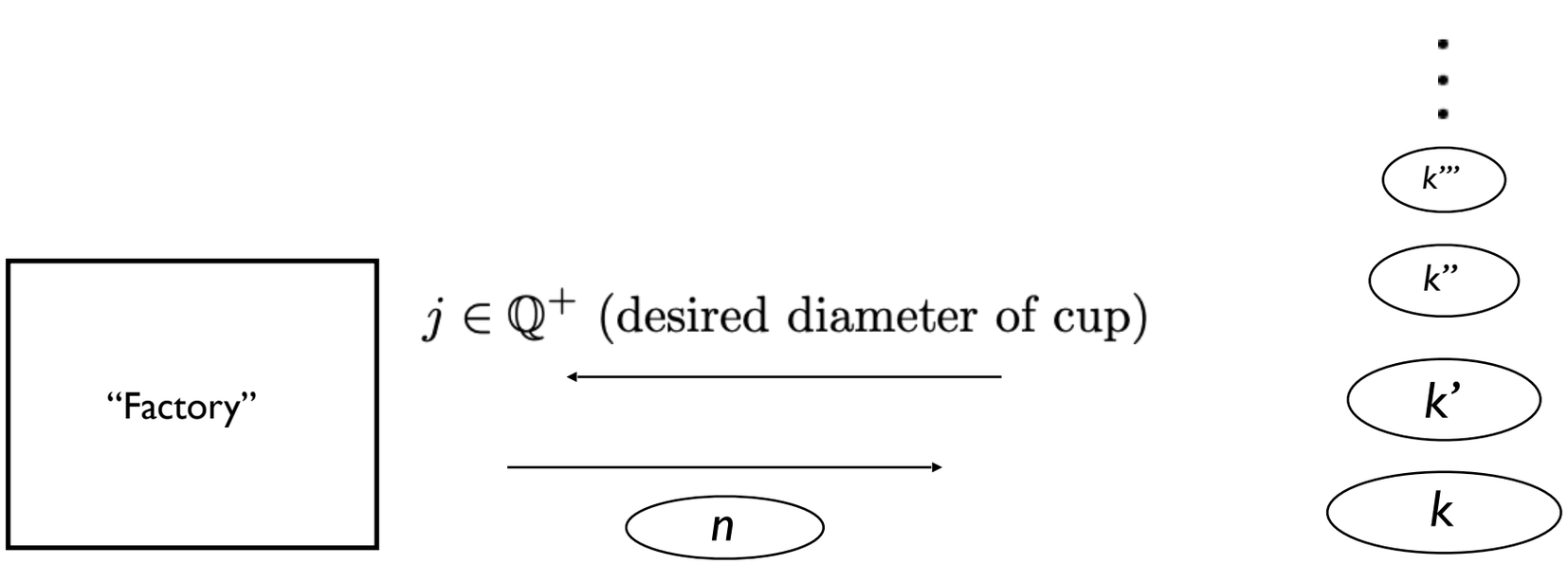
Suppose you have at your disposal a "factory" that, upon hearing you announce a number *j*, can quickly output a cup having a diameter of precisely *j* units. Can you insert a new cup between two of the seriated, stacked cups in the tower shown here? — where the *j* you send in *must* be a positive integer, *m* is likewise a positive integer, and every cup in every tower must be more in diameter than the one immediately above it, and less in diameter than the one immediately below it?** Prove that your answer is correct.



**E.g., if *m* = 3, the tower in that case will have a base cup 4 units in diameter, immediately above that a cup 3 units in diameter, then a cup 2 units in diameter, and then finally a top cup of 1 unit in diameter.

Selmer's Seriated Cup Challenge #2

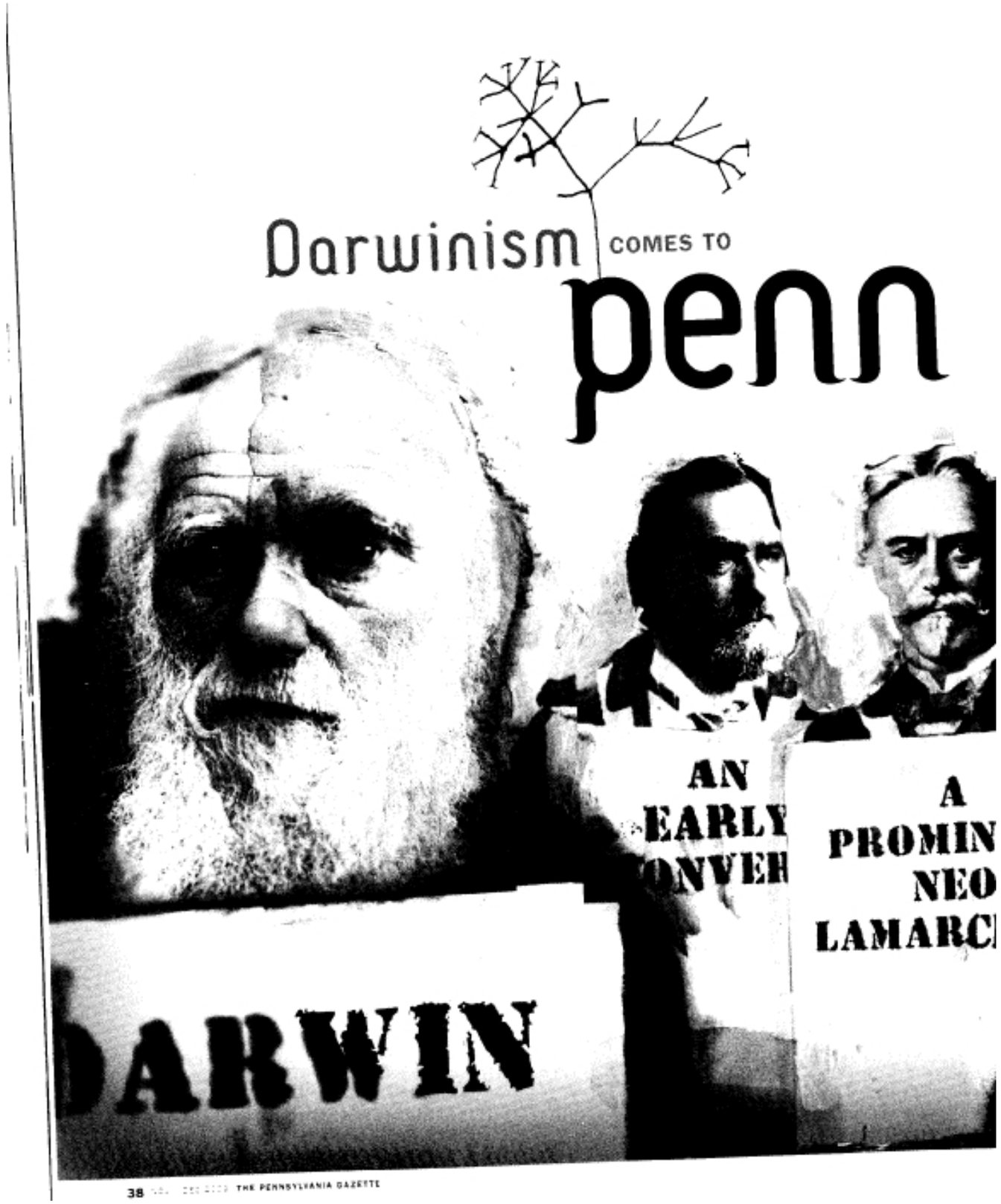
Suppose you have at your disposal a "factory" that, upon hearing you announce a number *j*, can quickly output a cup having a diameter of precisely *j* units. Can you insert a new cup between two of the seriated, stacked cups in the tower shown here? — where the *j* you send in *must* be a positive rational number; *k, k', k", k"'* … are likewise positive rational numbers, and every cup in every tower must be more in diameter than the one immediately above it, and less in diameter than the one immediately below it?** Prove that your answer is correct.



**E.g., if *k* = $\frac{1}{2}$, the tower in that case will have a base cup $\frac{1}{2}$ units in diameter, immediately above that there could be a cup $\frac{1}{3}$ units in diameter, then perhaps a cup $\frac{1}{4}$ units in diameter, and then perhaps finally a top cup of $\frac{1}{32}$ units in diameter.

Check your history books ...

From my alma mater: *Pennsylvania Gazette* Nov/Dec 2009



From my alma mater: *Pennsylvania Gazette* Nov/Dec 2009

From my alma mater: *Pennsylvania Gazette* Nov/Dec 2009

A century-and-a-half after the November 1859 publication of *On the Origin of Species*, a Penn microbiologist looks back at how Darwin's ideas were received by some of the University's leading thinkers. **BY HOWARD GOLDFINE**

ON June 18, 1858, Charles Darwin received a manuscript from Alfred Russel Wallace, which outlined a theory of evolution based on natural selection. Wallace's letter came from an island in the Malay Archipelago, where he was collecting field specimens and studying the distribution of species. Wallace, like Darwin, invoked the Malthusian concept that a struggle for existence within rapidly expanding populations would be the driving force for selection of natural variants within a species. Darwin's immediate reaction was one of dismay. He had been working on his "big book on species" since his five-year voyage on the Beagle (1831-36) and a relatively unknown naturalist had forestalled him. Darwin wrote to Charles Lyell, "If Wallace had my [manuscript] sketch written out in 1842, he could not have written out a better short abstract!"

Fortunately, Darwin had previously outlined his theory to his friends, the distinguished geologist Lyell and the botanist Joseph D. Hooker, and in a brief, unpublished draft to Asa Gray, a botanist at Harvard. Lyell and Hooker immediately arranged for Wallace's paper and a brief summary of Darwin's theory to be read simultaneously at the Linnaean Society in London on July 1, 1858. These were received with little comment. The president of the society later noted that nothing of great interest had happened that year.

ILLUSTRATION BY DAVID HOLLENBACH

A century-and-a-half after the November 1859 publication of *On the Origin of Species*, a Penn microbiologist looks back at how Darwin's ideas were received by some of the University's leading thinkers. **BY HOWARD GOLDFINE**

ON June 18, 1858, Charles Darwin received a manuscript from Alfred Russel Wallace, which outlined a theory of evolution based on natural selection. Wallace's letter came from an island in the Malay Archipelago, where he was collecting field specimens and studying the distribution of species. Wallace, like Darwin, invoked the Malthusian concept that a struggle for existence within rapidly expanding populations would be the driving force for selection of natural variants within a species. Darwin's immediate reaction was one of dismay. He had been working on his "big book on species" since his five-year voyage on the Beagle (1831-36) and a relatively unknown naturalist had forestalled him. Darwin wrote to Charles Lyell, "If Wallace had my [manuscript] sketch written out in 1842, he could not have written out a better short abstract!"

Fortunately, Darwin had previously outlined his theory to his friends, the distinguished geologist Lyell and the botanist Joseph D. Hooker, and in a brief, unpublished draft to Asa Gray, a botanist at Harvard. Lyell and Hooker immediately arranged for Wallace's paper and a brief summary of Darwin's theory to be read simultaneously at the Linnaean Society in London on July 1, 1858. These were received with little comment. The president of the society later noted that nothing of great interest had happened that year.

ILLUSTRATION BY DAVID HOLLENBACH

"On June 18, 1858, Charles Darwin received a manuscript from Alfred Russel Wallace, which outlined a theory of evolution based on natural selection. ... Darwin's immediate reaction was one of dismay. ... [That year] Wallace's paper, and a brief summary of Darwin's theory [were] read simultaneously (sic) at the Linnaean Society in London on July 1, 1858. ..."

Wallace rejected the claim that the human mind, with its capacity for abstract, rational thought, is the product of evolution by mutation and natural selection, on the basis of reasoned argument (Wallace's Paradox).

Wallace rejected the claim that the human mind, with its capacity for abstract, rational thought, is the product of evolution by mutation and natural selection, on the basis of reasoned argument (Wallace's Paradox).

Darwin did not.

And he defended his position in a book:
Descent of Man.

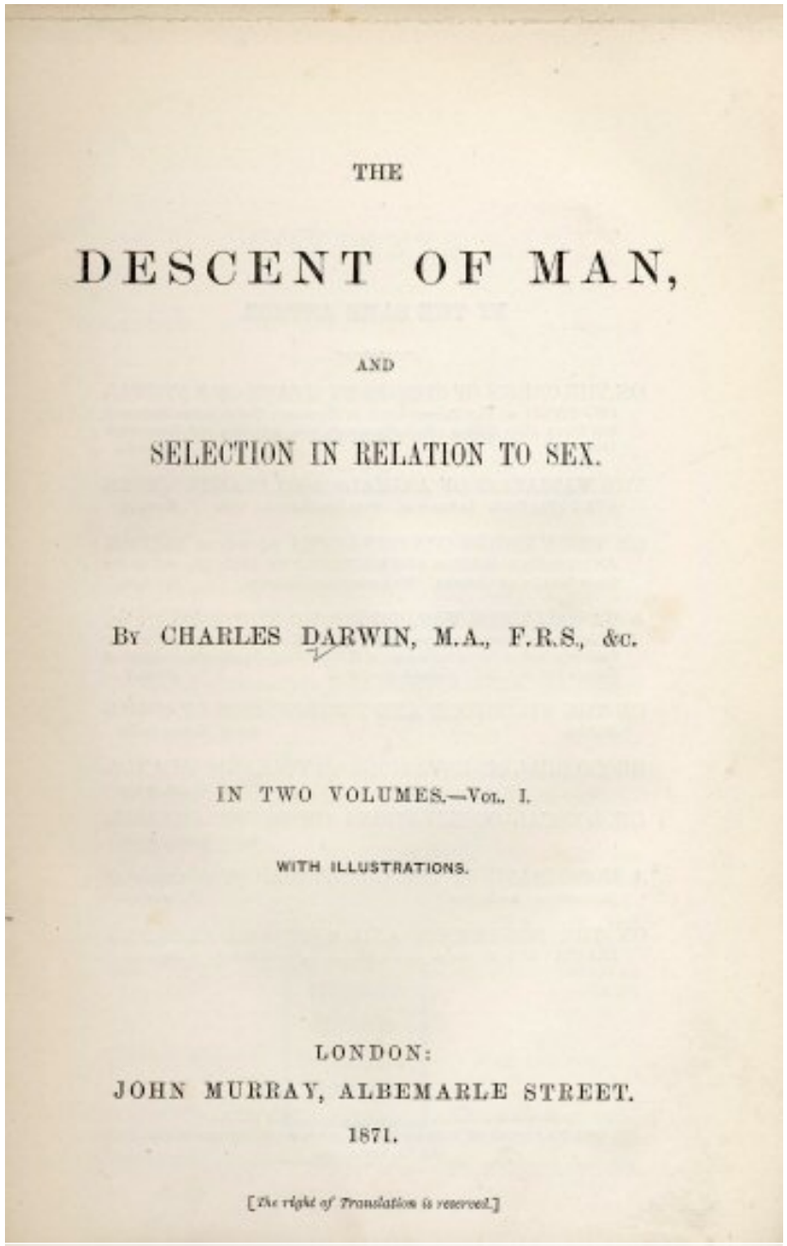
Wallace rejected the claim that the human mind, with its capacity for abstract, rational thought, is the product of evolution by mutation and natural selection, on the basis of reasoned argument (Wallace's Paradox).

Darwin did not.

And he defended his position in a book:
Descent of Man.

Wallace seems to me to be right; Darwin to be wrong…

The book that shook the world, and supposedly obliterated the stupid notion that human persons are made in (in Milton's unpacked version of the phrase) God's image.



THE

DESCENT OF MAN,

AND

SELECTION IN RELATION TO SEX.

By CHARLES DARWIN, M.A., F.R.S., &c.

IN TWO VOLUMES.—Vol. I.

WITH ILLUSTRATIONS.

LONDON:
JOHN MURRAY, ALBEMARLE STREET.
1871.

Praise for Darwin & *DoM*

Back cover of my Amazon.com version of *DoM*:
"Darwin's engaging literary style, charming modesty, brilliant argument, and discursive method of proof makes the book an exhilarating romp through Earth's natural history and Man's history ..."

Praise for Darwin & *DoM*

Back cover of my Amazon.com version of *DoM*:
"Darwin's engaging literary style, charming modesty, brilliant argument, and discursive method of proof makes the book an exhilarating romp through Earth's natural history and Man's history ..."

Really?

I found no brilliant arguments, and not a single proof.

A Key Proposition

A Key Proposition

$\bar{A}$

There is at least one mental power possessed by human persons, but not by any mere animal; and the mental powers of human persons are of a wholly different nature than those of mere animals.

Efficient Refutation of Darwin's *DoM*

	(1)	If human persons are the product of evolution, then it's not the case that $\bar{A}$ holds.	
	(2)	$\bar{A}$ does hold.	
$\therefore$	(3)	Human persons are not the product of evolution.	from (1), (2) by *modus tollens*

QED

Efficient Refutation of Darwin's *DoM*

	(1)	If human persons are the product of evolution, then it's not the case that $\bar{A}$ holds.	
	(2)	$\bar{A}$ does hold.	
$\therefore$	(3)	Human persons are not the product of evolution.	from (1), (2) by *modus tollens*

QED

Note: (3) doesn't deductively entail that *no* parts of human personhood are the product of evolution. In other words, (3) can be rephrased as: "Human persons are not solely and completely the product of evolution." As seen shortly, the power of human persons to carry out abstract, infinitary reasoning (as in the case of developing the tensor calculus) would be — according to Wallace & Bringsjord — something that evolution didn't produce.

Whence comes the first premise in this argument?

From Darwin Himself

From Darwin Himself

"If no organic being excepting man had possessed any mental power, or if his powers had been of a wholly different nature from those of the lower animals, then we should never have been able to convince ourselves that our high faculties had been gradually developed."

(*Descent of Man*, Part One, Chapter Two)

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ and, where P^{AI} is a limit on the cognitive power of AI, AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P)$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

E.g., from S.J. Gould: A big mutation happened but lay dormant for ~ 250,000 years.

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ and, where P^{AI} is a limit on the cognitive power of AI, AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P)$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

?

E.g., from S.J. Gould: A big mutation happened but lay dormant for ~ 250,000 years.

The Argument

	(1)	Hunter-gatherers possessed the cognitive power P^{∞} to e.g. invent the calculus and create literary art of the caliber of Blecher/Proust/Ibsen/. . . .	undisputed
	(2)	AI shows us that these early versions of us, to hunt and gather, needed only humble cognitive power P^{HG}, where $P^{\text{HG}} < P^{\infty}$, because $P^{\text{AI}} \approx P^{\text{HG}}$ and, where P^{AI} is a limit on the cognitive power of AI, AIs can hunt and gather.	see AI today
$\therefore$	(3)	We have $(P^{\infty} - P^{\text{HG}})$.	abstraction (1), (2)
	(4)	Our having $(P^{\infty} - P^{\text{HG}})$, *contra* Darwin, is inexplicable by gradual mutation and natural selection (i.e. P^{∞} is **discontinuous** from P^{HG}).	see critique of *DoM* see theorem/proof
	(5)	If our having $(P^{\infty} - P^{\text{HG}})$ is explicable, then $E_1 \vee E_2 \vee$ God exists.	sub-arg
	(6)	Our having $(P^{\infty} - P)$ is explicable.	undeniable
	(7)	$\neg E_1 \wedge \neg E_2$	sub-argument
$\therefore$	(8)	God exists.	*modus ponens* (5), (6), (7)

?