OEP: Ontology of Evolutionary Process

Practice & Method

Dr. Adam M. Goldstein

Darwin Manuscripts Project
http://darwin.amnh.org
American Museum of Natural History

Dr. Goldstein
http://shiftingbalance.org
z_californianus@shiftingbalance.org

New York Botanical Garden
Bronx, New York
7 November 2011

[slide #1]

Introduction

- Main topics today:
  - Ontology—a brief introduction
  - A quick scan of the kinds of phenomena that OEP is meant to illuminate
  - Overview account of the four classes of entities and their relationships according to OEP.
  - Assessment of the state of the discipline of formal ontology for biological sciences
  - Proposed alternative method to solve many problems posed by the dominant school of thought on ontological engineering
  -

[slide #2]

Where are we, and what's next?

[slide #3]

The fact of evolution

- In *The Origin of Species*, Darwin claimed that all organisms share a common ancestor or have an ancestor in a small group of organisms.
- This is the *fact of common descent*.
- Darwin often describes evolution as descent with modification.
- This claim was accepted almost immediately.
- Common descent is used to explain phenomena in every life science discipline.
- The research program: identify relationships of ancestry among the species.
- Establish *what begat what*.

[slide #4]

Mechanisms of evolution

- Darwin made a second claim: natural selection is the central mechanism of evolution.
- If no mechanism of evolution could be identified, species differences and complex adaptation would remain unexplained.
- Hardly anyone accepted Darwin's proposal about natural selection.
- The research program: evaluate Darwin's claim about natural selection and identify alternatives.
- These mechanisms of evolution are used to explain *how and why* dynamics of variation in the organic world occur.

[slide #5]

Where are we, and what's next?

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Fact vs. mechanism
How & why?
Ontology
Ontology of evolution
Method

[slide #6]

How and why does . . . occur? Phenomena of evolution

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Fact vs. mechanism
How & why?
Ontology
Ontology of evolution
Method

- Complex adaptation—displace argument from design
- Tempo and mode in macroevolution
- Patterns of molecular evolution
- Embryology, development, morphogenesis
- Differences in major structural elements, *Baupläne*
- Taxonomic relationships
- Phylogenesis
- Origin & nature of species as biological individuals
- Biogeography
- Many, many others . . .

[slide #7]

Where are we, and what's next?

Thomas Gruber's famous definition (1993)

Ontology "Explicit specification of a conceptualization"

Conceptualization "The objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them . . . . A conceptualization is an abstract, simplified view of the world that we wish to represent for some purpose."

Gruber's definition explained

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
What is ontology?
Ontology of evolution
Method

- Gruber's view is that ontologies are a means of information exchange between agents, machine or human.
- If the kind of content differs among data sources, there is a special need for a mechanism for exchange.
- Agents query one another in a standardized formal language.
- The formal language is a model of the general characteristics of the environment of the agents.

An Example

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
What is ontology?
Ontology of evolution
Method

- The Plant Ontology has the term "plant cell."
- A human researcher queries a robot search agent about specific measurements associated with specific kinds of plant cells.
- The robot searches several databases, each with different data sets about different plant species.
- The researchers curating the databases use the Plant Ontology to organize their data, with a common syntax to encode it.
- Because of this shared description of their specimens and the common encoding format, the robot identifies the correct information in each database.

A naturalistic account of formal ontology

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
What is ontology?
Ontology of evolution
Method

Ontology The ontology of domain D is a scheme for organizing information about things in D that reflects what those things are and what their relationships are to one another.

Question How do we know what the entities and relationships in D are?

Answer Our best source of information is (are) the science(s) of domain D.

Comment Terms in the ontology represent entities, properties; predicate-like representations of relationships between them.

Example Use Mx for "x has mass."

Example Use Cx for "x is a cell."

Example Use $R(x, y, z)$ for "x is between y and z."

Where are we, and what's next?

The central guiding design principle

OEP:
Ontology of Evolutionary Process

Dr. Adam M. Goldstein

Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

What explains an evolutionary process?

Formulate a general account of those entities and their relationships that explain an evolutionary process.

Desiderata for the ontology

OEP:
Ontology of Evolutionary Process

Dr. Adam M. Goldstein

Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

The ontology must be general enough to apply across:

- Levels of biological hierarchy
- Types of variation
- Types of evolving lineage
- Adaptive and non-adaptive processes
- Scales of time and space

Where are we, and what's next?

OEP:
Ontology of Evolutionary Process

Dr. Adam M. Goldstein

Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

Daniel Dennett and Richard Dawkins

- Dennett and Dawkins claim that the explanatory aim of evolutionary biology is to displace the argument from design.
- They say that Darwin's central innovation, which has turned out to be a lasting contribution, is the claim that natural selection explains adaptive complexity.
- While acknowledging Darwin's role in naturalizing organic phenomena, it must also be acknowledged that the Dennett-Dawkins view about evolutionary science is not likely to be true.
- The neutralist-selectionist dispute; isolation and subsequent reproductive separation in the origin of animal species; tempo and mode in large-scale evolution.

Sadly, a poverty of resources

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

- One strategy of response to Dennett and Dawkins would be to turn to the literature.
- How many papers report the occurrence of natural selection, and how many report the occurrence of (say) drift?
- Unfortunately, there is no way to search bibliographic records or the literature itself to answer this question in a reasonable amount of time.
- PubMed indexing fails miserably at capturing evolutionary topics
  - Not that NLM should be faulted for this—PubMed is, after all, intended for use in medical contexts.
- Free-text searching in abstracts, full article text, titles generates far too many false positives or negatives, depending on the search key.

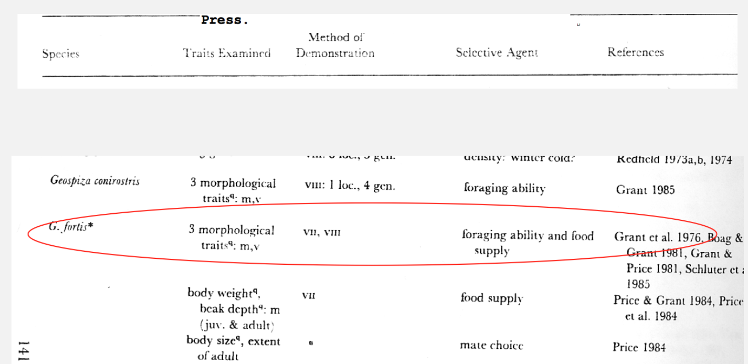
Darwin Manuscripts Project

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

- Darwin's marginalia, notes, and published works are continuous with one another—in many cases, his line of thought can be traced from initiation in notes or marginalia to a passage in a published work.
- Growing digital corpus at the Biodiversity Heritage Library presents the opportunity to link a user to a digital copy of a work mentioned by Darwin, even to a particular page or passage.
- Why not add a subject-oriented thread?
- Place Darwin's work in context of historically related works on topics he writes about?
- Example: the role of geographic isolation in the formation of new species.

From Endler's *Natural Selection in the Wild*

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method



Press.

Species	Traits Examined	Method of Demonstration	Selective Agent	References
			density? winter cold?	Redfield 1973a,b, 1974
Geospiza conirostris	3 morphological traits[q]: m,v	VIII: 1 loc., 4 gen.	foraging ability	Grant 1985
*G. fortis**	3 morphological traits[q]: m,v	VII, VIII	foraging ability and food supply	Grant et al. 1976, Boag & Grant 1981, Grant & Price 1981, Schluter et 1985
	body weight[q], beak depth[q]: m (juv. & adult)	VII	food supply	Price & Grant 1984, Price et al. 1984
	body size[q], extent of adult	*	mate choice	Price 1984

141

Where are we, and what's next?

The essence of an evolutionary process

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

Four elements for describing the history of a biological individual:

1. Some lineage or historical individual;
2. Some aspect in which that individual changes;
3. Some "environment," i.e., interaction of the individual as a whole or in parts with its surroundings; and
4. Some pattern or signature of the time-course of the changing aspect of the individual.

Progress?

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

Hypothesis: The four-element description can be generalized.

Initial assay: The *human population* changed by *natural selection* from *being primarily lactose intolerant to being tolerant* after *humans learned to domesticate diary animals*.

Assessment: The hypothesis is confirmed.

`EvolvingEntity`

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

- A historical individual.
- Is not continuously in the same state with regard to `EvolvableProperty`.

Example: Gene pool
Example: Gene locus
Example: Phylum
Example: Species

Context

- Analogous to environment, i.e., when considering the environment of organisms in their habitats.
- Contains or has an element of `EvolvingEntity` as a part.

Example Habitat of an organism (species or other population-like group of organisms)

Example Physiology of sex cell formation ("genes that violate Mendel's rules".

Example Habitat fragmentation (Influences rates of extinction, i.e., diversity of species in a phylum)

Example Geographic isolation (allopatry)

EvolvableProperty

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

- Quality of an element of *EvolvingEntity* that varies in character.
- Analogous to a morphological or physiological trait of an organism.

Example Alleles, interaction of organisms with some one of `Context`

Example Alleles, where `Context` includes genes that violate Mendel's rules

Example Quantitative trait e.g., height

Example Number of species in a phylum

Signature

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

- Pattern of change in `EvolvingProperty`; characteristic time-course of `EvolvingProperty`

Example Natural selection—better adapted organisms survive in greater proportion than less-well-adapted organisms

Example Species selection—species that have shorter life cycles have a higher probability of ramifying per unit time

Example Recurring mutations

Example Random drift, e.g., every element of `EvolvingProperty` has an equal probability of producing ancestors

The four classes and their relationships

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Architecture
Preshistory
The four classes
Method

Source of *x*	Source of *y*	*Pxy*
`Evolving entities`	`Contexts`	*x* is in context *y*
`Evolving entities`	`Evolvable properties`	*x* has character *y*
`Evolvable properties`	`Signature`	*x* evolves by *y*

Where are we, and what's next?

Kuhn's model of scientific change

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- In *The structure of scientific revolutions*, Thomas Kuhn describes what he claims to be a pattern in the history of science.
- Most of the time, an established science is "normal," that is, there are established theoretical principles, standards of evidence-giving, styles of reasoning.
- Established sciences will eventually reach a point of crisis: there will be some problem that the "normal" resources cannot be used to solve.
- The crisis is not resolved by solving the problem, but by the general acceptance by the scientific community of a new paradigm.

The pre-paradigm and interregnum periods

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- Kuhn identifies characteristics of the period before a science adopts its first paradigm and after a crisis has been reached but before a new paradigm has formed.
  - No discipline-wide agreement on basic issues of method, theory, observation, styles of reasoning
  - Scientists choose one working group over another based on loyalty, social networks, personal interests, happenstance
  - Partly as a result of domination of the field, a new paradigm takes shape around a particular approach
  - In retrospect, the others seem unreasonable, the work of quacks, pseudoscience.

The state of formal ontology in life sciences

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

Not looking up

The state of formal ontology in the life sciences most closely resembles the pre-paradigm stage.

Where are we, and what's next?

The Open Biological and Biomedical Ontology Group

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- the "OBO" has emerged as the most powerful influence on the community.
- They have become the de facto authority
- There are a set of principles that ontologies must adhere to if they are to be considered compliant
- Because they review ontologies for adherence to these principles, being approved by OBO is regarded similar to the way acceptance of a paper in a peer-reviewed journal is for an article.

Ontological Realism

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- The OBO group promotes a method and interpretation of formal ontology they term "realism"
  - Top-down development of ontologies
  - Ontologies about particular subjects must be formulated in an abstract vocabulary derived from Aristotle, purported to be a general description of reality
  - A central organization enforces compliance, establishes the correct interpretation of the Aristotelian principles.
- The OBO group's aim is to create a set of subject-area ontologies, unified by their adherence to the Aristotelian scheme.

Problems

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- The realist method is poorly suited for many tasks that ontology can be used for.

  For instance Comparing alternative theories; indexing literature; semantic web markup; describing experiments and hypotheses.
- Reformulating the description of one's domain using the Aristotelian vocabulary is difficult, requiring extensive consultation with OBO "gurus."
- Work groups sacrifice a degree of independence and control over how they want to structure their ontologies because the central authority may require changes.

Where are we, and what's next?

Ontological naturalism

OEP: Ontology of Evolutionary Process
Dr. Adam M. Goldstein
Introduction
Phenomena
Ontology
Ontology of evolution
Method
Confusion
Realism
Naturalism

- "Bottom-up" development
- Specialist sub-communities in a discipline develop their own ontologies
- Keeping to the vocabulary of the scientific theories of the domain is stressed
- A central organizing body provides resources and expertise to create and maintain links between the sub-community ontologies
- Absent the demand for adherence to Aristotelian principles, a wide range of uses for ontologies are permissible.