NEWSLETTER ON PHILOSOPHY AND COMPUTERS

FROM THE EDITOR, JON DORBOLO

FROM THE CHAIR, MARVIN CROY

ARTICLE

ERIC STEINHART
“Generating & Interpreting Metaphors with NETMET”

BOOK REVIEWS

L. Floridi, ed.: The Blackwell Guide to the Philosophy of Computing and Information

L. Floridi, ed.: The Blackwell Guide to the Philosophy of Computing and Information
Review of “Mind and AI”
Reviewed by Susan A. J. Stuart

L. Floridi, ed.: The Blackwell Guide to the Philosophy of Computing and Information
Review of “Science and Technology”
Reviewed by Marvin Croy

L. Floridi, ed.: The Blackwell Guide to the Philosophy of Computing and Information
Review of “Cybernetics” and “Artificial Life”
Reviewed by Ron Barnette
FROM THE EDITOR

Editorial Board
Jon Dorbolo, Editor
4140 The Valley Library
Oregon State University
Corvallis, OR 97331-4502
Jon.Dorbolo@orst.edu
dorboloj@ucs.orst.edu
<ucs.orst.edu/~dorboloj>
Phone: 541.737.3811

Bill Uzgalis, Associate Editor
Oregon State University
Wuzgalis@orst.edu

Ron Barnette, Teaching in Cyberspace Editor
Valdosta State University
rbarnett@grits.valdosta.peachnet.edu

Douglas Birsch, Computing Ethics Editor
Shippensburg University
dbrbs@ark.ship.edu

Larry Hinman, Internet Resources Editor
San Diego State University
hinman@acusd.edu

FROM THE CHAIR

Marvin Croy
University of North Carolina at Charlotte
mjcroy@email.uncc.edu

The Philosophy and Computers (PAC) Committee continues to be immersed in a variety of challenging and rewarding activities. First, the Barwise prize was awarded to Dan Dennett at the recent Eastern meeting in Boston. I opened the session with a few remarks about Jon Barwise and the nature of the prize, and Dennett was introduced by Ron Barnette, who was once his student and who is a previous member of our committee. Dennett gave an intriguing presentation, and we had plenty of discussion. Here are two photos of the event, one of Dan with the prize, and one of Dennett with current committee members who were able to attend (Peter Boltuc and Bruce Umbaugh), plus Ron and myself. Overall, it was a joyful event, and I’m looking forward to the next awarding of the Barwise prize to Deborah Johnson in Chicago at the 2005 Central Division meeting.

In October 2004, I traveled to the University of Delaware (home of the APA) for a meeting of APA committee chairs. It reinforced my belief that the reason I belong to the APA is because of what the committees stand and strive for. However, many committees are frustrated at not having the resources to attain their goals. During that meeting, I, along with others, supported two proposals to be considered by the Board at their November meeting. The first was to give committee chairs more direct input into who is appointed to their committees, and the second was to change the composition of the Board to include more committee representation. (Currently, only standing committees have a seat on the Board.) Neither of these proposals made any headway at the Board meeting. I do not know their exact fate since the Board has as yet published no minutes of that meeting. While there is some psychological comfort in knowing that frustrations over under-supported projects are widely shared among committees, the PAC committee will be undaunted in pursuit of its goals.

One central project of our committee is the development of a system for collecting and disseminating reliable information concerning the use of computers throughout the profession. I’ve sent the first component of our on-line survey to Chris Caputo, the APA webmaster. Michael Kelly has been supportive of this effort, even though some committees prefer to attempt
data collection via mail-out surveys. This is an expensive proposition, and I’m hoping that our success with an electronic survey will lead the way for other committees to obtain reliable data. Our main problem will be the number of departments who won’t take the time to supply data, but we’ll formulate some strategies and cross that bridge when we come to it. Once we obtain data on the first component (computers in teaching), we’ll generate a report and pass the findings on to the Committee on Teaching. Then, we’ll focus on computers in research and, finally, on computers in professional cooperation.

Andrew Light (president of the Society for Philosophy and Technology) and Michael Kelly have received an $81,000 grant from the National Science Foundation (NSF) to increase and improve philosophical reflection upon technology, mostly via APA activities. As chair of the PAC Committee, I sit on an advisory board that will serve to guide this project. (Other board members include Peter French, Deborah Johnson, and Hugh LaFollet.) As a consequence, some PAC APA sessions will be directed toward this NSF project’s goals. However, there will still be plenty of opportunity for PAC Committee members to plan and contribute to these sessions. So, I encourage our committee members (and interested others) to continue their efforts in this direction. Looking ahead, there’s been a proposal by Bruce Umbaugh and Chris Grau to plan a session for the 2005 Eastern meeting, and we will probably be awarding the 2005 Barwise prize at the 2006 Pacific session. We’ll start our deliberations on the Barwise prize following the 2005 Central meeting.

The APA’s Committee on Committees has recently appointed two new members to our committee. They are Susan Stuart of the University of Glasgow and Brandon Fitelson of the University of California—Berkeley. As of July 1, they’ll be on board, and Patrick Grim and Noam Cook will be cycling off. (Both Patrick and Noam have been active in organizing APA sessions and their contributions will be missed.) You may have read in the most recent Proceedings that, starting in 2006, the PAC Committee will have an associate chair. At the October Chairs meeting, there was much concern expressed about the lack of continuity on the committees. The opportunity for committees to have an associate chair who serves for one year prior to actually becoming the chair is one solution to this problem, and this procedure will be adopted starting in 2006 for our PAC Committee. I will continue as chair until 2007, and, during my last year, I will basically familiarize the associate chair with our projects and will delegate (hopefully many) duties to that person.

I will be traveling to the Pacific 2005 meeting and taking part in a session stimulated by Peter Boltuc, who will also be participating, along with Patrick Suppes. The general topic will be “Perspectives on Philosophy On-Line.” Noam Cook has been active in arranging a session on “Computers and the Mediation of Human Experience.” His session is jointly sponsored with the Society for Philosophy and Technology. So, if you’re traveling to the Pacific Division meeting, please join us for a stimulating session and perhaps a lunch or dinner afterwards.

That’s all for now. Let me know if you have questions, concerns, or suggestions related to PAC Committee activities.
NETMET takes a list of propositions as its input. They are encoded in NETMET’s input file and are written in a canonical form similar to that of the predicate calculus. NETMET comes with thirteen input files. These include codings of (1) the SOCRATES IS A MIDWIFE analogy from Plato’s Theaetetus; (2) the MEMORY IS A WAX TABLET and MEMORY IS AN AVIARY analogies from Plato’s Theaetetus; (3) the SOUND IS A WAVE analogy; (4) the ATOM IS A SOLAR SYSTEM analogy; (5) the COULOMB’S LAW IS NEWTON’S LAW analogy; and (6) the argument by analogy for other minds.

Given an input file, NETMET is able to perform several basic functions: (1) display the words and propositions in the file; (2) generate lexical subnets in which a word sits; (3) find subnets analogous to a given subnet; (4) compute analogical mappings across subnets; (5) generate metaphors involving distinct but analogous subnets; (6) transfer information from one subnet to another based on an analogy; and (7) produce partial literal interpretations of metaphors generated by analogical transference. NETMET also allows the user to explicitly group propositions into subnets by specifying such groupings in the input file. An explicit subnet has the form FIELD \( \{ P_1, P_2, \ldots, P_n \} \) where \( P \) is a proposition (the term “FIELD” indicates NETMET’s origin in Kittay’s semantic field theory of metaphor). The propositions enclosed in curly braces are all put into the same subnet. Table 1 shows two subnets taken from Plato’s (1984) SOCRATES IS A MIDWIFE metaphor.

<table>
<thead>
<tr>
<th>FIELD {</th>
<th>FIELD {</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:contains(mother, womb)</td>
<td>T1:contains(Theaetetus, mind)</td>
</tr>
<tr>
<td>2:goes(womb, thru:cycle)</td>
<td></td>
</tr>
<tr>
<td>3:ovulates(womb)</td>
<td></td>
</tr>
<tr>
<td>4:menstruates(womb)</td>
<td></td>
</tr>
<tr>
<td>5:produces(womb, egg)</td>
<td></td>
</tr>
<tr>
<td>6:discards(womb, egg)</td>
<td></td>
</tr>
<tr>
<td>7:if S2 then {S3, S4}</td>
<td></td>
</tr>
<tr>
<td>8:if S3 then {S5}</td>
<td></td>
</tr>
<tr>
<td>9:if S4 then {S6}</td>
<td></td>
</tr>
<tr>
<td>10:contains(womb, baby)</td>
<td>T6:contains(mind, idea)</td>
</tr>
<tr>
<td>11:produce(mother, baby)</td>
<td>T7:produce(Theaetetus, idea)</td>
</tr>
<tr>
<td>12:give-birth(mother, baby)</td>
<td>T8:express(Theaetetus, idea)</td>
</tr>
<tr>
<td>13:help(midwife, mother)</td>
<td>T9:help(Socrates, Theaetetus)</td>
</tr>
<tr>
<td>14:pass(liveborn(baby), physical test)</td>
<td>T10:pass(true(idea), cognitivetest)</td>
</tr>
<tr>
<td>15:opposites(liveborn, stillborn)</td>
<td>T11:opposites(true, false)</td>
</tr>
</tbody>
</table>

Since it claims that metaphors are generated from analogies, STM requires a theory of analogy. NETMET uses theories of analogy developed by Gentner (1983), Falkenhainer, Forbus & Gentner (1989), and Holyoak & Thagard (1989, 1990). An analogy is a triple \((S, T, f)\) where \(S\) and \(T\) are subnets and the counterpart function \(f\) is a structure-preserving function from \(S\) to \(T\) (i.e., an approximate isomorphism). The generation of an analogy is traditionally divided into two major subphases: (1) the access phase and (2) the mapping phase (Hall, 1989). The access phase generates the source and target subnets \(S\) and \(T\); the mapping phase generates the counterpart function \(f\).

To generate the target and source subnets, NETMET begins with a cue. The cue is a single word, which the user selects from the words in the input file. NETMET generates the subnets of which the cue is a member. Once the subnets of the cue have been generated, the user must select one of these as the target. NETMET then generates subnets that are analogous to the target. These are the candidate source nets. Once the candidate source nets have been generated, the user must select one as the source. For instance, suppose NETMET is given an input file containing the fields in Table 1. Suppose the user picks “idea” as the cue. NETMET displays the subnet of “idea” as the target. NETMET then indicates that the field of “childbirth” is (the only) source subnet.

After analogical access has produced target and source subnets, the analogical mapping phase generates the counterpart function \(f\) that maps the source \(S\) onto the target \(T\). Several matching principles are used to generate ordered pairs that preserve the local relational structure of the source. The counterpart function emerges through the interaction of these local matches. NETMET uses the connectionist technique of constraint satisfaction to find the counterpart map (Holyoak & Thagard, 1989, 1990). A map from the subnet of “childbirth” to the subnet of “idea” is shown in Table 2.

STM holds that a metaphor is a statement that combines terms from a target subnet with terms from an analogous but distinct source subnet. Terms from the target are subscripted with \(T\), while those from the source are subscripted with \(S\).

| Table 1. Subnets from Plato’s SOCRATES IS A MIDWIFE metaphor. |
|---|---|
| FIELD \{ | FIELD \{ |
| 1:contains(mother, womb) | T1:contains(Theaetetus, mind) |
| 2:goes(womb, thru:cycle) | |
| 3:ovulates(womb) | |
| 4:menstruates(womb) | |
| 5:produces(womb, egg) | |
| 6:discards(womb, egg) | |
| 7:if S2 then \{S3, S4\} | |
| 8:if S3 then \{S5\} | |
| 9:if S4 then \{S6\} | |
| 10:contains(womb, baby) | T6:contains(mind, idea) |
| 11:produce(mother, baby) | T7:produce(Theaetetus, idea) |
| 12:give-birth(mother, baby) | T8:express(Theaetetus, idea) |
| 13:help(midwife, mother) | T9:help(Socrates, Theaetetus) |
| 14:pass(liveborn(baby), physical test) | T10:pass(true(idea), cognitivetest) |
| 15:opposites(liveborn, stillborn) | T11:opposites(true, false) |

| Table 2. Counterpart correspondence \(f\) for the subnets in Table 1. |
|---|---|---|---|
| mother \(\rightarrow\) Theaetetus | T5 | S5 \(\rightarrow\) T2 |
| womb \(\rightarrow\) mind | T3 | S6 \(\rightarrow\) T1 |
| baby \(\rightarrow\) idea | T7 | S11 \(\rightarrow\) T9 |
| egg \(\rightarrow\) idea | T9 | S13 \(\rightarrow\) T10 |
| midwife \(\rightarrow\) Socrates | T11 | S14 \(\rightarrow\) T10 |
| liveborn \(\rightarrow\) true | |
| stillborn \(\rightarrow\) false | |
Consider the sentence, “Socrates is a midwife.” From Table 1, we see that: “Socrates” comes from the target; “midwife” comes from the source. The copula links phrases from distinct but analogous subnets. The copula and the sentence containing it are thus metaphorical. We thus get the syntactically decorated sentence: [:Socrates]_T [is]_{MET} [:a midwife]_S.

Metaphors of the form [:x]_T [is]_{MET} [:y]_S are trivial. A great advantage of STM, indeed, one that separates it from all other theories of metaphor so far, is that it incorporates metaphors into standard generative grammar. By decorating standard syntactic rules (in Backus-Naur form) with parameters for the situations S and T, and marking phrases as metaphorical when S is not equal to T, STM can handle metaphors of any grammatical complexity. NETMET implements only about a dozen grammatical forms. I list a few of these forms in Table 3.

Table 3. A few grammatical forms handled by NETMET.

| [DET NOUN1]_T [BE]_{MET} [DET NOUN2]_S | Socrates is a midwife. |
| [DET NOUN1]_T [BE]_{MET} [DET ADJ]_S [NOUN2]_S | The mind is an intellectual womb. |
| [DET NOUN1]_T [BE DET NOUN2]_S of [DET NOUN3]_S | An idea is a child of the mind. |
| [DET NOUN1]_T [BE]_{MET} [DET ADJS]_S | Some ideas are stillborn. |
| [DET NOUN1]_T [VERBS]_S [PREP NOUN2]_S | Students give birth to ideas. |
| [DET NOUN1]_T [VERBS]_S [PREP DET NOUN2]_S of [DET NOUN3]_T | A student gives birth to a child of his mind. |

In these forms, brackets indicate phrase groups; subscripts S and T indicate situations; “DET” is a determiner; “NOUN” is a noun; “BE” is a form of the verb “to be”; “ADJ” is an adjective; and so on.

Besides generating metaphors, NETMET also performs analogical transference. Analogical transference is particularly important for philosophy of science because it creates a new theory of the target. Using the counterpart function, NETMET moves propositions from the source to the target. NETMET examines every proposition P in the source. If there is no analogous proposition for P in the target, then NETMET (1) makes a copy of P; (2) replaces every noun or adjective in the copy of P with its analogue, and (3) adds the modified copy of P to the target. Analogical transference produces a new target subnet in which many propositions are metaphorical predications (verb-predication metaphors, metaphorical rules, etc.). An example of analogical transference is shown in Table 4. Analogically transferred propositions are in boldface; words (or phrases) used metaphorically are enclosed in square brackets.

I deliberately chose an unusual metaphor generated from the SOCRA T ES IS A MIDWIFE analogy: [:The mind]_T [goes through a menstrual cycle]_T]. Reasoning by analogy is a powerful tool for making hypotheses, and NETMET can generate surprisingly novel metaphors. The metaphor is by no means merely poetic, obscene, or surreal. It is true if the mind goes through some kind of creativity cycle. Whether such a cognitive creativity cycle exists is a genuine scientific question. In extreme forms, such a cycle may take the form of manic-depressive syndrome; the relation between manic-depressive syndrome and creativity is an active research topic (Jamison, 1993).

NETMET can construct partial literal interpretations of the verb-predication metaphors generated by analogical transference. These interpretations are not truth-conditions, but they are used in the truth-conditional analyses of STM. According to Black (1962, 1979), metaphors are interpreted by generating implication complexes. An implication complex consists of all the entailments of a proposition. Implication complexes are generated by the repeated application of rules. An implication complex has the form of a tree. Those propositions that have no entailments are the leaves of an implication complex. For example, the implication complex of the literal proposition, “S2:goes(womb, thru:cycle)” is shown in Table 5. The leaves are S5 and S6.

Given a verb-predication metaphor produced by analogical transference, NETMET partially interprets that metaphor by generating its implication complex. For example, the implication complex of T12[:goes(mind, thru:cycle)]_T is shown in Table 6. The leaves of an implication complex are of particular interest. The conjunction of the literal leaves of the implication complex constitutes a partial literal interpretation or partial literal paraphrase of the metaphorical predication.

Although NETMET deals only with the syntactic aspects of STM, it is important to see that these syntactic aspects are grounded in a genuinely semantic analysis of metaphor. STM depicts reality as a logical space carved up into situations. A situation is a Tarskian model. There is a part-whole lattice of situations. Some situations are sufficiently large and complete to be called possible worlds (in the cosmic sense). But possible worlds have parts (submodels). These are like Hintikka’s “small worlds,” or “scenarios.” Some situations are analogous to one another. An analogy is a triple (S, T, f) where S and T are situations and f is a structure-preserving map. Situation S is the source and T is the target. We say x in T is the f-counterpart of y in S iff (S, T, f) is an analogy and f maps x in T onto y in S. The notion of a counterpart is familiar from David Lewis. Truth-conditions for metaphors are based on counterpart correspondences.

Consider, for instance, this short text from Shakespeare’s Romeo and Juliet, Act II Scene II: “[Juliet appears above at a window] ROMEO: But, soft! what light through yonder window breaks? It is the east and Juliet is the sun!” The text specifies an analogy: Juliet is to the window as the sun is to the east; just as the sun appears in the east, so also Juliet appears in the window. Table 7 lays out the analogy as a triple (S, T, f).

Metaphors combine terms from distinct but analogous situations. Consider the metaphor, “Juliet is the sun.” Since “Juliet” denotes an individual in S, I write it as [Juliet], and “the sun” denotes an individual in S, I write it as [the sun]. STM assigns both literal and metaphorical truth-conditions to all sentences. Let [x]_T and [y] denote individuals in the subscripted situations.
Table 4. Subnets after analogical transference.

<table>
<thead>
<tr>
<th>FIELD</th>
<th>FIELD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1:contains(mother, womb)</td>
<td>T1:contains(Theaetetus, mind)</td>
</tr>
<tr>
<td>S2:goes(womb, thru:cycle)</td>
<td>T12:[goes(mind, thru:[cycle])]</td>
</tr>
<tr>
<td>S3:ovulates(womb)</td>
<td>T13:[ovulates(mind)]</td>
</tr>
<tr>
<td>S4:menstruates(womb)</td>
<td>T14:[menstruates(mind)]</td>
</tr>
<tr>
<td>S5:produces(womb, egg)</td>
<td>T2:produces(mind, idea)</td>
</tr>
<tr>
<td>S6:discards(womb, egg)</td>
<td>T3:discards(mind, idea)</td>
</tr>
<tr>
<td>S7:if S2 then {S3, S4}</td>
<td>T4:forgets(mind, idea)</td>
</tr>
<tr>
<td>S8:if S3 then {S5}</td>
<td>T5:if T3 then {T4}</td>
</tr>
<tr>
<td>S9:if S4 then {S6}</td>
<td>T15:[if T12 then {T13, T14}]</td>
</tr>
<tr>
<td>S10:contains(womb, baby)</td>
<td>T16:[if T13 then {T2}]</td>
</tr>
<tr>
<td>S11:produce(mother, baby)</td>
<td>T17:[if T14 then {T3}]</td>
</tr>
<tr>
<td>S12:give-birth(mother, baby)</td>
<td>T6:contains(mind, idea)</td>
</tr>
<tr>
<td>S13:help(midwife, mother)</td>
<td>T7:produce(Theaetetus, idea)</td>
</tr>
<tr>
<td>S14:pass(liveborn(baby), physicaltest)</td>
<td>T18:[give-birth(Theaetetus, idea)]</td>
</tr>
<tr>
<td>S15:opposites(liveborn, stillborn)</td>
<td>T8:express(Theaetetus, idea)</td>
</tr>
<tr>
<td></td>
<td>T9:help(Socrates, Theaetetus)</td>
</tr>
<tr>
<td></td>
<td>T10:pass(true(idea), cognitivetest)</td>
</tr>
<tr>
<td></td>
<td>T11:opposites(true, false)</td>
</tr>
</tbody>
</table>

Literal truth-conditions are as expected from Tarski: \([x]_T \text{is LIT}[y]_S \text{LIT}\) is true iff \(x = y\). Hence \([\text{Juliet}]_T \text{is LIT}[\text{the sun}]_S \text{LIT}\) is true iff \(\text{Juliet} = \text{the sun}\); but that is false. The metaphorical truth-conditions are based on counterpart correspondence: \([x]_T \text{is the } f\text{-counterpart of } y \text{ in } S\) \(\Leftrightarrow (S, T, f) \text{ is an analogy } (x \text{ in } T \text{ is the } f\text{-counterpart of } y \text{ in } S)\). Hence \([\text{Juliet}]_T \text{is the } f\text{-counterpart of } [\text{the sun}]_S \text{ in } S\). Table 7 lays out the very analogy that makes \([\text{Juliet}]_T \text{is the } f\text{-counterpart of } [\text{the sun}]_S \text{ in } S\). STM can assign truth-conditions to metaphors of many grammatical forms using standard semantic techniques (popularized by Montague, Davidson, and so on).

The analysis of metaphors in terms of counterpart correspondences among analogous parts of possible worlds links metaphors to fascinating (and problematic) issues elsewhere in philosophy. Metaphors are linked to the issues of identity and indiscernibility in modal logics. They are linked to Quinean notions of ontological relativity and ontological reduction. Since analogies are based on isomorphisms, metaphors are made true by certain kinds of symmetries. Thus the use of analogy and metaphor is linked to deep questions about symmetry and symmetry breaking in science.

Table 5. An implication complex for a literal proposition.

<table>
<thead>
<tr>
<th>Source Situation S</th>
<th>Target Situation T</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2:goes(womb, thru:cycle)</td>
<td>T13:[ovulates(mind)]</td>
</tr>
<tr>
<td>S5:produces(womb, egg)</td>
<td>T14:[menstruates(mind)]</td>
</tr>
<tr>
<td>S6:discards(womb, egg)</td>
<td>T4:forgets(mind, idea)</td>
</tr>
</tbody>
</table>

Table 6. An implication complex for a metaphorical proposition.

<table>
<thead>
<tr>
<th>Target Situation T</th>
<th>Analogy f</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1: woman( Juliet);</td>
<td>theSun → Juliet</td>
</tr>
<tr>
<td>T2: window( aWindow);</td>
<td>theEast → aWindow</td>
</tr>
<tr>
<td>T3: appears( Juliet, aWindow)</td>
<td></td>
</tr>
<tr>
<td>T4:forgets(mind, idea)</td>
<td></td>
</tr>
</tbody>
</table>
philosophical theories can benefit from computational modeling and (2) as a classroom tool for computer-assisted instruction in the philosophy of language and science. It illustrates the power of what Jim Moor has called “the computational turn” in philosophy.

References

BOOK REVIEWS

The Blackwell Guide to the Philosophy of Computing and Information
L. Floridi, editor, Philosophy of Computing and Information. (Blackwell Publishing, 2004). Paperback $34.95, Hardback $69.95
This book is an invaluable resource for all who are interested in the complex range of issues that make up the philosophy of computing and information movement. It is especially valuable to teachers who want to acquaint students with the philosophical concepts and problems of computing and information. The guide consists of twenty-six chapters, each written by an authority in the relevant topic area. The chapters are grouped under seven topic areas. All of the articles are clearly written with a minimum of assumed technical language, such that they are likely accessible to any informed reader.

The Guide is accompanied by a website (http://www.blackwellpublishing.com/pci/), which provides a thirty-six-page glossary of technical and philosophical terms, along with a twenty-one-page bibliography of major resources related to the areas covered in the book.

To encourage critical discourse based on the Guide, we solicited reviews of its parts. Four of those reviews appear below, as well as an interview with Blackwell Acquisitions Editor Nick Bellarini.

Contents of The Blackwell Guide to the Philosophy of Computing and Information
1. Introduction
What is the Philosophy of Information? Luciano Floridi
2. Four Concepts
Computation, Jack B. Copeland
Complexity, Alasdair Urquhart
System, Klaus Mainzer
Information, Luciano Floridi
3. Computers in Society
Computer Ethics, Deborah G. Johnson
Communication and Interaction, Charles Ess
Cyberspace, Wesley Cooper
Digital Art, Dominic McIver Lopes
4. Mind and AI
The Philosophy of AI and Its Critique, J. James H. Fetzer
Computationalism, Connectionism and the Philosophy of Mind, Brian P. McLaughlin
5. Real and Virtual Worlds
Ontology, Barry Smith
Virtual Reality, Derek Stanovsky
The Physics of Information, Eric Steinhart
Cybernetics, Roberto Cordeschi
Artificial Life, Mark A. Bedau
6. Language and Knowledge
Information and Content, Jonathan Cohen
Knowledge, Fred Adams
Formal Languages, Graham White
Hypertext Theory, Thierry Bardini
7. Logic and Probability
Logic, Aldo Antonelli
Probabilistic Reasoning in AI, Donald Gillies
Decision and Game Theory, Cristina Bicchieri
8. Science and Technology
Computing in the Philosophy of Science, Paul Thagard
Methodology of Computer Science, Timothy Colburn
Philosophy of IT, Carl Mitcham
Computational Modeling as Philosophical Methodology, Patrick Grim

Interview with Nick Bellorini
Commissioning Editor, Philosophy
Blackwell Publishing
Nick.Bellorini@oxon.blackwellpublishing.com
Please give some background on yourself and your work as Blackwell Philosophy editor.
Nick Bellorini: I’m one of two Philosophy books acquisitions editors at Blackwell, responsible for one of the oldest and largest publishing programs in the discipline (around 40-50 new titles each year, 30+ journals) with a backlist dating to the Philosophical Investigations in 1953. My own background is in philosophy, and I received my doctorate from Oxford University in 2000, where I worked in the philosophies of mind and language.
Nick Bellorini: Our early publishing in this area—The Digital Phoenix and Cyberphilosophy—was primarily the product of our relationship with the journal Metaphilosophy. The decision to publish the Guide was the result of our responding positively to the energy and enthusiasm of Luciano Floridi and our sense that the project was a serious attempt by a collection of first rate philosophers to map out a nascent field of study and research. As to whether the publishing is commercially viable, the honest answer is that the jury is out. We hope the work being done but also, importantly, upon the degree to which those who are more and those less decision-theoretic in outlook.

Though certain areas of the subject get ever more technical, I certainly get no sense of a methodological or conceptual transforming of the discipline.
Outside philosophy, one gets a sense that the concept of information has increased prominence in social science, especially in areas like economics and geography. Blackwell is also due to publish a reference work on the “Digital Humanities” through our cultural studies list.

What do you perceive in the future for philosophy of computing and information studies?
Nick Bellorini: Since I think some of the topics and problems in this area are of enduring interest, I think the field has a quite bright future. Whether it becomes an established sub-discipline will presumably depend in part on the quality of the work being done but also, importantly, upon the degree to which it becomes an area of undergraduate study in its own right. We hope the Guide will contribute to this process.

The Blackwell Guide to the Philosophy of Computing and Information
Review of “Mind and AI”
Susan A. J. Stuart
University of Glasgow
s.stuart@philosophy.arts.gla.ac.uk
In the Preface to this volume (pp.xi – xiv), Floridi states that this is the first systematic attempt to map the philosophical issues that are raised by the information revolution and how they might be addressed; it is an “exploration as much as an introduction,” adding later that “Each chapter is meant to provide not only a precise, clear, and accessible introduction but also a substantial and constructive contribution to the current debate.” I have borne these claims in mind when reviewing the Mind and AI chapters, and I can, very happily, confirm that they have been achieved.

The Mind and AI section comprises two chapters, the first written by James H. Fetzer, titled, “The Philosophy of AI and its Critique,” and the second written by Brian P. McLaughlin, titled, “Computationalism, Connectionism, and the Philosophy of Mind.”
Fetzer’s chapter covers an immense amount of ground in twenty-one sections, from Turing through, for example, Folk Psychology and Eliminative Materialism, to Semiotic Systems and the Hermeneutic Critique. As an introduction to Philosophy of AI, and how we have moved on from Descartes to think of the mind in both computational and non-computational terms, it is a tour de force.

Fetzer’s writing style is extremely clear, and, although a great deal of what he has to say is distilled from large and very complex areas, the reader is never left behind or bamboozled by technical vocabulary. Indeed, one of the delights of this chapter is the skillful way in which the author leads the reader, via a neat system of hypotheses, on a chronological journey through the various theories that have been proposed as explanations of the mind. We begin with the Cartesian hypothesis that “(Conscious) human minds are thinking things” and conclude, by way of—amongst others—Newell and Simon’s symbol system hypothesis and Fodor’s language of thought hypothesis, that “Thinking things, including human beings, are semiotic systems.” It’s a fantastic chapter for anyone writing an introductory course on the Philosophy of AI; each of the hypotheses is the subject for a lecture, and a ten-lecture course falls into your lap.

The element of this chapter that is as much exploration as substantial and constructive contribution to the current debate is the critique of the theories and hypotheses that runs throughout. This is subtle, yet it provides the reader with a steadying hand to guide her through the debate, reminding her from time to time not to get too comfortable with any particular explanation, that they all have weaknesses.

At the last moment, in a section on intelligent machines, Fetzer strikes a warning note that we, the reader, should not be seduced, like so many before us, into confusing intelligent machines for thinking things (p.133). That we should still need a warning of this sort is indication enough that these issues are alive and well and still very far from resolution.

Fetzer’s chapter is followed by McLaughlin’s with his focus trained on computationalist and connectionist approaches to understanding intelligence and the creation of artificial intelligence. The form of McLaughlin’s chapter almost mirrors Fetzer’s with sections and, in this case, subsections that are short and reasonably self-contained. They range from The Computational Theory of Mind, through The Symbol-system Paradigm, and The Connectionist Computational Paradigm, to the final section, How are the Paradigms Related? It is a chapter that has been laid out both chronologically and—it certainly feels—incrementally, to maximize the benefit to the reader and to guide the novice reader through the acquisition of difficult concepts, for example, function, computation, algorithm, and architecture, and their relation one to the other.

It is a much more technical read than the previous chapter and requires a greater commitment to concentration from the reader. But it has many rewards. I particularly enjoyed the concise, yet wonderfully clear, subsection on Fodor’s language of thought hypothesis and psychosemantics. They provided both exposition of the thesis and responses to it in the forms of Searle’s “Chinese Room,” Block’s dual-aspect semantics, and Kripke’s radical indeterminism of symbols (p.140-1). It is very measured writing and fully complements the very different style and approach of “The Language of Thought” section (p.126) written and responded to by Fetzer in the preceding chapter.

It is a text that is so well-composed and knowledgeable that it imbues in the reader a feeling of trust. So, when McLaughlin systematically cauterizes all manner of blind alleys and wild assumptions, the sorts of ideas that an enthusiastic newcomer might have felt inclined to have taken up, the reader is disinclined to question his judgement. But this does not make it a text of which we should be suspicious; on the contrary, it makes it an excellent tool for a guiding readers at any level through a noisy and sometimes overly-heated debate. The more sophisticated the reader, the more sophisticated his response, and he may not always agree with McLaughlin’s decisions. But any reader worth his salt, who hasn’t yet achieved a level of sophistication which gives him the confidence to disagree, will be excited by what he read and want to pursue the topic in greater detail.

The final section offers many directions in which the debate might proceed, a number of them in some hybrid form of symbolic and subsymbolic computation with any number of possible functional architectures, but, ultimately, McLaughlin ends on a similar note to Fetzer, though not expressed as a warning. None of the issues with which he engages has been, or is even close to being, resolved, and, “What the future will bring remains to be seen.” (p.148).

Part III: Mind and AI is essential to the book’s project as a whole, since the subjects of mind and artificial intelligence contain at their heart the main concepts of this book—computation, complexity, system, and information.

The Blackwell Guide to the Philosophy of Computing and Information

Review of “Science and Technology”

Marvin Croy
University of North Carolina–Charlotte
mjrcroy@email.uncc.edu

This section contains four chapters. Three of these explore the borderlands of science, philosophy, mathematics, and territories of IT. Three chapters, by Paul Thagard, Timothy Colburn, and Patrick Grim, focus on methodology and force a reconsideration of philosophy itself. A fourth article, by Carl Mitcham, both chronicles and exemplifies philosophical reflection upon technology in general and IT in particular. These four chapters exemplify philosophical inquiry directed both at IT and at itself, and they reinforce the notion that philosophy is always self-critical.

In “Philosophy and Information Technology,” and in previous work, Mitcham distinguishes two major traditions within the philosophy of information technology. One emerges from engineering. This tradition takes humanity as being essentially technological and proposes to advance technology through all aspects of human existence. A second tradition emerges from within the humanities. It takes technology to be only one aspect of the human and proposes to limit technological projects to appropriately restricted applications. Mitcham traces the conflict between these “expansionist” and “limitationist” traditions as they appear within theories of IT and within conflicting definitions of information (uninterpreted signal transmission versus semantically useful resource). This discussion is followed by a historical outline of ambivalent attitudes toward the development of information-related technologies (particularly printing and modes of interpretation). The expansionist outlook emphasizes the promise of artificial intelligence (AI) and automated agents. The limitationist perspective is more skeptical and presses pointed questions concerning particular topics within ethics.
(privacy, mediation of human interaction, software reliability),
politics (equity of Internet access, social justice, democracy),
psychology-epistemology (information overload, making
sense of the world), and psychology-anthropology (living the
"virtual" life).

This questioning ultimately turns metaphysical, and
Mitcham follows its path to the work of Heidegger, one
proponent of the view that IT warps human existence at the
deepest level. Mitcham briefly, but clearly, elaborates the main
corresponding to various philosophical pursuits. However, the question
ultimately raises is whether computational methods provide
new means of achieving traditional philosophical aims or
whether their acceptance requires a reconceptualization of
philosophy itself. In “Computational Modeling as a
Philosophical Methodology,” Patrick Grim presents computer
modeling as a new means of achieving the philosophy of
as logical systems, ontological hierarchies, and various analytical
tools can play a role in this abstraction process, Colburn
envisions a close and fruitful connection between philosophy
and computer science. Such an interaction can result in the
design of new computational objects and methods, including
new modeling languages. In sum, Colburn convincingly
demonstrates the value of philosophical reflection upon
computer science and provides insightful examples of such
reflection.

In “Computing in the Philosophy of Science,” Paul Thagard,
a pioneer in this area, presents computational modeling as an
addition to the standard methods of philosophical investigation
such as logical analyses and historical accounts. Thagard
explicates three computational frameworks (cognitive
modeling, engineering AI, and the theory of computation) for
approaching problems within the philosophy of science. For
each of these approaches, recent advances are abstracted in
relation to two significant aspects of science: discovery and
evaluation. Criteria for assessing the results of these
computational endeavors are also presented. Thagard conveys
a palpable sense that valuable progress is being made on
multiple fronts. His description of the benefits of
computational modeling is similar to that outlined by Grim,
emphasizing increased precision and more extensive testing of
a theory’s principles and assumptions (determining, for
example, whether a certain kind of behavior can be produced by
a certain kind of system).

In presenting computational methods as supplementing
existing methods within the philosophy of science, Thagard
opens the door to meta-methodological inquiries. These
inquiries concern the appropriate application of and
relationship between various methods. Logical analysis and
historical case study, for example, have been taken by many
to represent a clash between normative ideals and empirical
explanations. The resolution of this issue is itself part of the
philosophy of science, as evidenced in the writings of Hempel,
Kuhn, Popper, Lakatos, and others. Thagard makes two
references that reveal the stakes involved in adopting a new
philosophical method such as cognitive simulation. First,
Thagard portrays cognitive simulation as part of naturalistic
epistemology in which the study of knowledge is “closely tied
to human psychology.” This raises questions about the nature
of these ties. How and to what degree is philosophy empirical?
How do its empirical aspects relate to (and possibly interact with)
its normative tasks? Is a naturalistic turn required for
accepting the philosophical relevance of cognitive simulation?
This line of inquiry is also suggested by Thagard’s reference to
Aaron Sloman’s groundbreaking work in this area. Anyone familiar with the intricacies of Sloman’s Philosophy and the Computer Revolution is well aware of the magnitude of the challenge involved in finding common ground between philosophy and science. This challenge is worth taking up, and Thagard shows the value of building upon Sloman’s heroic efforts.

It should be clear, however, that advocating new methods within an established discipline lays down a challenge. Either one must show that the new methods can achieve existing goals better than or as well as existing methods, or one must show that the new methods achieve new goals. Any advocacy of new goals requires a reconceiving of the discipline. This is part of what makes reflection upon information and IT so rewarding. Constructing the philosophy of information and orienting ourselves towards newly spawned methods challenges us to clarify our conception of what philosophy is and how we can better pursue it and how we renew ourselves as philosophers in the process. In these chapters, Colburn, Grim, Mitcham, and Thagard speak from their intellectual homes and accentuate the promise of the challenges ahead.

**The Blackwell Guide to the Philosophy of Computing and Information**

**Review of “Cybernetics” and “Artificial Life”**

**Ron Barnette**  
Valdosta State University  
rbarnett@valdosta.edu

Philosophers have pursued the question of human nature since their beginnings and have expressed many accounts to explain what makes us different from non-humans. One is obviously reminded of the Socratic dictum, “Know Thyself.” In late pre-Darwinian days of the early Enlightenment, Descartes introduced a distinction between human beings and machines, based upon his belief that humans enjoy a true autonomy and free choice, unlike mechanisms, which include animals. Not soon to follow, however, La Mettrie, a fellow Frenchman, claimed that humans themselves were machines with a particularly complex physical organization. Since then, the question of intelligent human life and the possibility of mechanistic explanations have given rise to a wealth of thought, in particular, in post-Darwinian twentieth century debates. Since the 1940s and continuing today, central concepts in the study of common human/machine behaviors and abilities, and their differences, have been explored by Cybernetics (earlier), and Artificial Life (later). In this general review, two articles will be discussed, as they describe and analyze these developments: chapter 14, Cybernetics, by Roberto Cordeschi, and chapter 15: Artificial Life, by Mark A. Bedau.

An inspiring development in interdisciplinary studies of human behavior and its unique character was characterized by cybernetics, a term first used by Norbert Wiener in 1947. As Cordeschi notes, Wiener’s definition of cybernetics as the study of “control and communication in the animal and the machine” captures the ambition to develop a unified theory of the behavior of living organisms and machines viewed as systems governed by the same physical laws. But how is self-control, and consistent regulation of behavior, possible in a machine? Cybernetics explores these questions directly, as Cordeschi develops and describes the concepts of self-regulation, feedback control, and homeostatic systems, critical to cybernetics. Earlier milestone work is discussed on purposive behavior, arguably central for understanding intelligent living systems, as well as how cybernetics proposed a scientifically grounded explanation of intentional, purposive behavior in standard (efficient) causal terms, unlike the discredited views of Aristotelian final causes, which were often associated with the notion of purpose in behavioral explanations. Intentionality, so characteristic of intelligent life, had its early twentieth century scientific studies in the context of cybernetics, as newly-emerging interdisciplinary studies tackled the nature of intentional action and purposive behavior, so hotly debated in philosophy and psychology. While not mentioned in the essay, these earlier efforts to account for purposive behavior mechanistically became integral to the important literature surrounding Action Theory, so prevalent in the 1960s and 1970s.

The interaction with logic and neuroscience is another feature of classical cybernetics, and Cordeschi provides a helpful discussion of the critical developments of what would be characterized later in the twentieth century as artificial neural networks and computer science, as new fields of study emerged from cybernetics’ interdisciplinary evolution. Key figures and publications are discussed, as this important history of the search for the nature of intelligent systems unfolds, including their conceptual models of mechanism for mechanism learning, trial-and-error adaptation, and the critical notion of self-organization in the context of a system’s ever-changing environmental confrontations. But as the new disciplines were spawned, there emerged a new object to challenge classical cybernetics and to model and represent intelligent communication systems mechanically: the computer, and the birth of Artificial Intelligence (AI) research...and, eventually, what would be the new cybernetics of cognitive science. Pioneers of AI criticized the inadequacy of the artifacts proposed by cybernetics, such as neural networks or systems with rather rudimentary self-organizing capabilities, in their attempts to simulate cognitive processes: self-regulation, while essential, was not sufficient to account for full-blown purpose and intentionality. Hierarchies of computational structures, guided by logical algorithms, with evolving strategies, symbolic structures, and selection mechanisms, enhanced the AI model of intelligence, and Cordeschi describes clearly this transition in theoretical developments in the new information age, as he characterizes the onslaught of new cybernetics. Still, the early AI approach, with its new computational model, characterized an intelligent system, and, by extension, intelligent life, as governed by a “top-down” research strategy akin to the current computation of a central processing unit, or ultimate control management, in our personal computers. This earlier computational architecture was to evolve, too, as models of intelligent systems explored new avenues of theoretical understanding of essential human behavior that required many “subsystems,” which required adaptable levels of controlling mechanisms and functions. What eventually emerged was a model of low-level routines, which characterized intelligent agency as being composed of functionally distinct control levels, which was ignored, according to Cordeschi, in earlier cybernetics and AI. The new cybernetics would need to address how these control levels would act, with proper self-regulative controls, on an environment without being supervised by a centralized control and action-planning center. A rather ambitious program, indeed. What would evolve in the 1980s would be a so-called “parallel distributed processing” (PDP) model, as modern AI expands its search for intelligent life, and as the new cybernetics is transformed radically.
Cybernetics has interesting concluding sections on early attempts to apply cybernetics research to general social sciences, with their wider ranging, different interdisciplinary projects, notably, in sociology and economics. Weiner, who considered cybernetics as "nothing if it is not mathematical," interestingly enough (historically), was troubled by cybernetics' encouraging a trendy fashion, already rampant (according to him), consisting of the inappropriate use of mathematics in the human sciences: "a flood of superficial and ill-considered work." A most poignant observation in 1964, in light of academic directions and critical reviews of the social sciences since then, especially regarding the latter's attempts to provide more mathematical, quantitative studies of social, political, and economic behaviors. Readers will appreciate the historical references and reconstruction of theoretical developments in this critical juncture in the search for intelligent life pioneered by cybernetics.

Mark A. Bedau, in Artificial Life, explores a myriad of philosophical topics as he describes and analyzes this recent development in understanding life…intelligent and otherwise. But just what is Artificial Life, theoretically? Etymologically, that term was coined by Christopher Langton in 1989, and, as Bedau notes, Artificial Life, or ALife, is a broad, interdisciplinary endeavor that studies life and life-like processes through simulation and analysis. Critical to these studies happens to be a central goal of cybernetics as well, namely, self-regulation and the capacity to maintain an internal equilibrium in the face of dynamical changes in the external environment. Hence the obvious connections with other evolving interdisciplinary. However, ALife's broader set of goals expands on this capacity and broadens philosophy's quest into the fundamental aspects of life and what makes it unique, and this chapter explores the roots of ALife's interdisciplinary beginnings, its methodology, and the numerous components of the research program and its implications. Situated in contemporary philosophy, biology, artificial intelligence, physics, and mathematics, these disciplines are discussed intricately vis-à-vis their relationships with ALife...including a very interesting claim (with support) that John von Neumann implemented the first ALife model!

The computer model methodology of ALife is vital for several reasons. The utilization of parallel distributed processing (PDP) models in connectionist and neural network approaches enables a "bottom-up" architecture for ALife studies, where micro-level entities are allowed to continually affect the context of their own behavior, where their own "decisions" and behaviors are based on information only about their own local situation. As such, the ALife model represents nature's examples in living systems. Unlike the traditional "top-down" approaches of AI models, Bedau explains lucidly how lower-level "bottom-up" models are characterized as "agent-based," or "individual-based." As such, the whole living system's behavior is represented in the model only indirectly; it arises out of the interactions of a collection of directly represented local parts. This notion is critical to ensuing discussions on hierarchies of levels of organization, self-organization and control, adaptation, and emergence...concepts that have been integral to the search for intelligent life.

Bedau discusses attempts to model behaviors of living systems and their critical emergent phenomena and discusses the attending problems. Self-organization is the most primitive phenomenon explored by some ALife models of biological systems, while others attempt to model interactions between different types of organisms or agents. It is in the latter where life-like phenomena are studied in systems represented in John Conway's Game of Life studies, which are described clearly in the research of evolutionary self-organization. Since both living systems and ALife models are commonly said to exhibit emergent phenomena, it is valuable to the discussion that Bedau includes a helpful analytical section on emergence, distinguishing between emergent properties, emergent entities, and emergent phenomena, which is developed clearly. Basically, an entity with an emergent property is construed to be an emergent entity, and an emergent phenomenon involves an emergent entity possessing an emergent property, as he characterizes these concepts.

The problem of emergence for ALife is this: How to explain emergent phenomena, characterized as being (a) constituted by, and generated from, underlying phenomena, yet (b) being autonomous from these underlying phenomena? In other words, "How can something be autonomous from underlying phenomena if it is constituted by and generated from them?" Bedau's developments of three positions (including his own) on this question are most informative to the ALife development on emergence, and his discussion is valuable in the context of the debates surrounding adaptation, perhaps the most important concept in evolution and the understanding of intelligent life. It is also important to note that ALife's methodology is very important, indeed, to the traditional "mind/machine" debates in philosophy in this regard. The grounding of mentality in physical structures and their functions is not unlike ALife's approach to emergence and adaptation, as it turns out, and the computational study of the latter plays an active role in defining new approaches for future philosophical debates over supervenience, reduction, complexity, and mentality. As such, ALife offers an expanded research tool in understanding and appreciating intelligent life.

Bedau presents and develops a very illuminating discussion about controversial claims and arguments over adaptationism: the thesis that the activity of pursuing adaptive explanations of biological traits, so commonly claimed, is a legitimate part of empirical science. Clear differences exist on this thesis and are distinguished. Likewise, there is a remarkable section on evolutionary progress, and the debates over evolution, complexity, and progressive advantages are argued differently by Stephen Jay Gould and by Daniel Dennett. These sections raise a host of critical issues directly related to ALife studies, and Bedau makes a convincing case for the claim that ALife can make a number of contributions to these debates, as empirical studies can be brought to bear on relevant ALife computational models and which can, in fact, test the theoretical thought experiments presented by the disputants. In fact, the achievement of such a goal, vis-à-vis Gould's hypothetical "Tape of Life" experiment, is one of the key open problems in ALife, according to Bedau, as he describes the research program.

The very nature of life, and whether this is a scientific or philosophical matter, is discussed in the context of three prominent views about the issue, which Bedau develops and discusses clearly. As with other debates, he claims that the advent of ALife has revitalized investigations into these topics, including the important question about life itself. As such, ALife both fosters a broad perspective of life, through synthesis, modeling, and simulation, and just might offer the potential to create radically new forms of life. As he argues, "In the final analysis, the nature of life will be settled by whatever provides the best explanation of the rich range of natural phenomena that seem to characterize living systems." And Bedau's own support of ALife results are indeed paramount in this regard, as its interdisciplinary promise is rich in its scope and methodology.

However accurately computers model life and simulate living entities, the question inevitably arises: Can a computer simulation of a living system ever literally be alive? A "yes"
answer supports the position of Strong Artificial Life, and Bedau’s development of this controversy is exceptionally illuminating in this concluding section. In fact, he presents a lucid argument against standard criticisms of Strong ALife as he discusses the salient features of an expanded Turing Test for ALife, a distinction between functionalism and computationalism, adaptive intelligence, and representationalism. He makes a good case, indeed, that the debate about Strong ALife will enliven and inform many such related issues in the philosophy of mind and artificial intelligence. Claiming that ALife and philosophy are natural partners, Bedau articulates succinctly how ALife’s computational methodology is “a direct and natural extension of philosophy’s traditional methodology of a priori thought experiments,” as ALife models explore empirically the consequences of certain simple ideas or premises. “They are ‘thought experiments’ explored with the help of a computer,” he concludes.

The search for life and intelligence by machine continues…from earlier twentieth-century cybernetics and its developments through Artificial Intelligence and Artificial Life and beyond. Socrates’ prescription of “Know Thyself” has indeed enjoyed a fascinating evolution of method and reflection upon who we are.