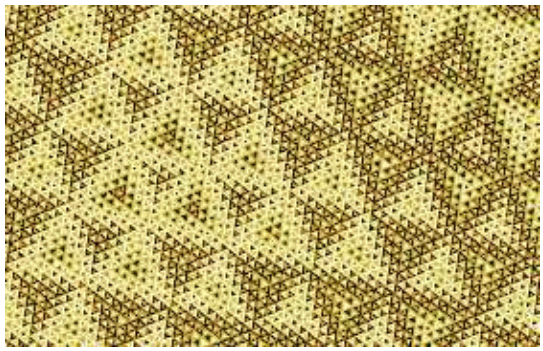
SynergieWissen
Kulturgeschichtliches Forum am ZfL Berlin



Emergenz und Innovation
Dynamiken des Neuen in Natur und Kultur

Workshop am 2. November 2012
Textvorlagen

Achim Stephan (Universität Osnabrück): Spielarten der Emergenz

Text: Achim Stephan: Zur Rolle des Emergenzbegriffs in der Philosophie des Geistes und in der Kognitionswissenschaft, in: Dieter Sturma (Hg.), Philosophie und Neurowissenschaften, Frankfurt am Main 2006, S. 146–166.

Georg Toepfer (ZfL Berlin): Emergenz und Organismus

Text: Mark A. Bedau: Is weak emergence just in the mind? Minds & Machines, 18.4 (2008), S. 443-459.

Jörg Richter (ZfL Berlin): Fortschritt durch Mutation

Text: Susan Glaspell: The Verge (1921). Plays by Susan Glaspell. Ed. with an introd. by C. W. E. Bigsby, Cambridge 2006 (Auszüge).

Die Neurowissenschaften halten das psychophysische Problem für weitgehend geklärt und unterstellen, dass auch die philosophischen Fragen nach Selbstbewusstsein, Willensfreiheit und menschlichen Handlungen durch die neurowissenschaftlichen Forschungsergebnisse im Wesentlichen beantwortet seien. Damit verbinden sich Forderungen nach einem neuen Menschenbild und nach praktischen Konsequenzen für Erziehung und Rechtsprechung.

Die Beiträge dieses Bandes unterziehen die philosophischen wie auch die neurowissenschaftlichen Lösungsvorschläge zum psychophysischen Problem einer eingehenden Prüfung und loten dabei auch die Bedingungen für einen fruchtbaren Austausch zwischen Philosophie und Neurowissenschaften aus – ein Austausch, der den jeweiligen disziplinären und methodischen Eigenheiten hinreichend Rechnung trägt.

Dieter Sturma, geboren 1953, ist Professor für Philosophie an der Universität Duisburg-Essen.

Philosophie und Neurowissenschaften

Herausgegeben von
Dieter Sturma

Suhrkamp

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der
Deutschen Nationalbibliografie
http://dnb.ddb.de

suhrkamp taschenbuch wissenschaft 1770
Erste Auflage 2006

Suhrkamp Taschenbuch Verlag

Druck: Nomos Verlagsgesellschaft, Baden-Baden
Printed in Germany
Umschlag nach Entwürfen von
Willy Fleckhaus und Rolf Staudt
ISBN 3-518-29370-2

2 3 4 5 6 – 11 10 09 08 07 06

K

Inhalt

Dieter Sturma

Zur Einführung: Philosophie und Neurowissenschaften[1]

Die Neurowissenschaften treten gegenwärtig mit einer Vielzahl von Nachrichten über Erträge und Folgen ihrer Forschungen an die Öffentlichkeit. Es wird über therapeutische Durchbrüche genauso berichtet wie über Experimente, denen zu entnehmen sei, dass es keine menschliche Willensfreiheit gebe. Dem sachlichen Bezug nach berühren diese Nachrichten theoretisch wie praktisch den Kern personalen Lebens. Viele Neurowissenschaftler sehen sich aufgrund dieser thematischen Nähe imstande, Kants vierte Frage »Was ist der Mensch?« zu beantworten – eine Frage, die bislang in die Zuständigkeit der Philosophie gefallen ist.

In etlichen Stellungnahmen wird zudem der Eindruck erweckt, die naturwissenschaftliche Aufklärung des menschlichen Bewusstseins sei im Prinzip abgeschlossen, und es käme nur noch darauf an, in Detailanalysen einzelne Funktionsweisen des menschlichen Gehirns aufzuhellen. Der Philosophie wird in diesem Zusammenhang allenfalls eine untergeordnete Rolle zugebilligt, denn zur Lösung der noch offenen Probleme könne sie ohnehin nichts beitragen. Diese Diagnose geht mit einer oftmals ausdrücklich formulierten Enttäuschung über die Philosophie einher, die in den letzten zwei Jahrtausenden keine Fortschritte bei der Erforschung des Bewusstseins erbracht hätte.[2]

Die Neurowissenschaften zielen auf ein neues Menschenbild, wobei sie den Geisteswissenschaften zuweilen durchaus ein Mitspracherecht einräumen.[3] Bei genauerer Betrachtung bleibt allerdings unklar, welches Menschenbild überhaupt revidiert werden soll. Verwiesen wird auf dualistische Modelle der Trennung von Geist und Körper, von denen aber eher zweifelhaft ist, ob sie überhaupt für das

1 Ich danke Francesca Raimondi, Katinka Schulte-Ostermann und Ulrich Steckmann für die redaktionelle Bearbeitung des Bandes.

2 Vgl. Bennett und Hacker, »Philosophie und Neurowissenschaft«, in diesem Band S. 22.

3 Siehe »Das Manifest«, S. 37.

Achim Stephan

Zur Rolle des Emergenzbegriffs in der Philosophie des Geistes und in der Kognitionswissenschaft

In der Philosophie des Geistes und in der Kognitionswissenschaft stößt der Begriff der Emergenz seit den frühen 1990er-Jahren auf zunehmendes Interesse.[1] Ihm wird zugetraut, sowohl besondere philosophische Schwierigkeiten wie etwa das Problem der phänomenalen Qualitäten – das so genannte *hard problem of consciousness* – treffend zu klassifizieren als auch nicht explizit programmierte, sondern im Rahmen von Selbstorganisationsprozessen entstehende Verhaltensweisen künstlicher Systeme adäquat zu charakterisieren. Eine nähere Betrachtung zeigt allerdings, dass es keinen einheitlichen Emergenzbegriff gibt, der beiden Aufgaben gerecht wird. Die »Kandidaten« für Emergenz in Philosophie des Geistes und Kognitionswissenschaft sind zu verschieden voneinander: Mentale Phänomene, die sich als sperrig gegenüber reduktiven Erklärungsversuchen erweisen, erfordern für ihre Klassifikation einen sehr starken Emergenzbegriff, der zwar im Rahmen metaphysischer Fragestellungen große Bedeutung für die Philosophie des Geistes hat, in den übrigen Disziplinen der Kognitionswissenschaft (und in den Naturwissenschaften) aber so gut wie keine Rolle spielt. Stattdessen geht es dort in der Regel um Verhaltensweisen komplexer dynamischer Systeme, die auf der Makroebene erstaunliche Muster und Regelmäßigkeiten zeigen, ohne sich jedoch als widerspenstig gegenüber reduktiven Erklärungen zu erweisen; vielmehr kann deren Verhalten häufig auf einfache Interaktionen einer Vielzahl einfacher Komponenten zurückgeführt werden.

Obzwar die Philosophie des Geistes eine der Disziplinen ist, die mit der Kognitiven Psychologie, Künstlichen Intelligenz, Robotik, Computerlinguistik und den Neurowissenschaften, um nur einige zu nennen, die Kognitionswissenschaft konstituiert, ist es daher sinnvoll, im Rahmen der Emergenz-Debatte die philosophisch interessanten Phänomene gesondert von den übrigen kognitionswissenschaftlich beachtenswerten Phänomenen zu diskutieren.

1 Zu Beginn des 20. Jahrhunderts kam es zu einer ersten Blütezeit für Emergenztheorien: Damals ging es in Philosophie und Biologie darum, einen mittleren Weg zwischen Vitalismus und Mechanismus zu finden. Zur Geschichte und Problematik des Emergentismus vgl. Stephan, *Emergenz*.

1. Spielarten des Emergentismus

Im Folgenden werde ich verschiedene Varianten des Emergentismus vorstellen und diese in Beziehung zu spezifischen Problemen setzen, die für die Philosophie des Geistes und die Kognitionswissenschaft, dort besonders den Konnektionismus und die Robotik, relevant sind. Unter den verschiedenen Spielarten des Emergentismus sind drei Theorien besonders hervorzuheben: der *schwache* Emergentismus, der *synchrone* Emergentismus und der *diachrone* (Struktur-) Emergentismus. Der schwache Emergentismus, der durchaus mit verschiedenen Versionen des Reduktionismus kompatibel ist, stellt die gemeinsame Basis für alle anspruchsvolleren Emergenztheorien dar. Diese ergeben sich aus jenem durch das Hinzufügen stärkerer Thesen. So berücksichtigt der diachrone Emergentismus Aspekte der Neuartigkeit und Unvorhersagbarkeit, der synchrone Emergentismus das Merkmal der Irreduzibilität. Auf die näheren Details dieser Positionen komme ich gleich zurück. Beginnen wir mit dem schwachen Emergentismus.

1. Schwache Emergenz

Der *schwache* Emergentismus vereint verschiedene Minimalanforderungen an emergente Eigenschaften. Die drei Thesen, die ihn konstituieren – die *These des physischen Monismus*, die *These der systemischen Eigenschaften* und die *These der synchronen Determiniertheit* –, sind auch mit einer reduktionistischen Position vereinbar.

Das erste Merkmal ist eine These über die *Beschaffenheit der Systeme*, die emergente Eigenschaften haben. Sie besagt, dass die Träger der emergenten Eigenschaften ausschließlich aus physischen Entitäten bestehen. Mögliche Kandidaten für emergente Eigenschaften wie »lebendig zu sein«, »einen Dur-Dreiklang zu hören« oder »starke Angst zu spüren« werden nach der *These des physischen Monismus*

allein durch physische Systeme mit einer hinreichend komplexen Mikrostruktur instantiiert. Diese schließt zugleich aus, dass für das Haben emergenter Eigenschaften *übernatürliche* Komponenten wie eine *Entelechie* oder eine *res cogitans* verantwortlich sind. Im besonderen bedeutet dies, dass Systeme, die lebendig sind oder Geist haben – seien sie natürliche oder artifizielle Systeme –, aus den gleichen basalen Bausteinen bestehen wie die unbelebten Dinge der Natur.

Physischer Monismus: Die im Universum vorhandenen und entstehenden Systeme setzen sich ausschließlich aus physischen Entitäten zusammen. Auch die als emergent zu charakterisierenden Eigenschaften, Dispositionen, Verhaltensweisen oder Strukturen werden nur von solchen Systemen instantiiert, die ausschließlich aus physischen Komponenten bestehen.

Während die erste These die Diskussion emergenter Eigenschaften und Strukturen in den Rahmen eines physikalistisch-naturalistischen Rahmens stellt, der insbesondere hinsichtlich künstlicher Systeme auch kaum anders zu denken ist, grenzt die zweite These den Typus der Eigenschaften ein, die überhaupt als mögliche Kandidaten emergenter Phänomene in Frage kommen. Es ist die *These der systemischen Eigenschaften*. Dieser These liegt die Annahme zugrunde, dass die allgemeinen Eigenschaften komplexer Entitäten in zwei verschiedene Gruppen zerfallen:[2] (1) in solche Eigenschaften, die auch einige der Systembestandteile haben, und (2) in solche, die kein Bestandteil des Systems hat. Beispiele für Elemente der ersten Gruppe sind Eigenschaften wie »ausgedehnt zu sein« oder »eine Geschwindigkeit zu haben«. Beispiele für die zweite Gruppe sind u. a. »zu fliegen«, »sich fortpflanzen zu können«, »zu atmen« oder »einen Juckreiz zu spüren«. Diese werden als *systemische* oder kollektive Eigenschaften bezeichnet.

Systemische Eigenschaften: Emergente Eigenschaften sind systemische (oder kollektive) Eigenschaften. Eine Eigenschaft ist genau dann systemisch, wenn ein System als Ganzes sie hat, aber wenn kein Bestandteil des Systems eine Eigenschaft dieses Typs hat.

Dass es sowohl künstliche als auch natürliche Systeme mit systemischen Eigenschaften gibt, dürfte unstrittig sein. Wer dies leugnen wollte, müsste die Auffassung vertreten, dass *alle* Systemeigenschaften bereits von einigen der Systembestandteile exemplifiziert werden. Dem widersprechen zahllose Beispiele.

2 Unter allgemeinen Eigenschaften verstehe ich hier Eigenschaften eines generellen Typs wie »eine Masse« oder »eine Farbe« zu haben, nicht aber spezifische Eigenschaften wie »eine Masse von 64, 58 kg« oder »die Farbe Purpur« zu haben.

Während die erste These den Typ der Bausteine eingrenzt, aus denen sich Systeme mit emergenten Eigenschaften zusammensetzen, und die zweite These den Typ der Eigenschaften näher charakterisiert, die für Emergenz in Frage kommen, spezifiziert die dritte These den Typ des Verhältnisses, das zwischen der Mikrostruktur eines Systems und dessen emergenten Eigenschaften besteht, als eines der synchronen Determiniertheit:

Synchrone Determiniertheit: Die Eigenschaften und Verhaltensdispositionen eines Systems hängen nomologisch von dessen Mikrostruktur, das heißt den Eigenschaften seiner Bestandteile und deren Anordnung ab. Es kann keinen Unterschied in den systemischen Eigenschaften geben, ohne dass es zugleich Unterschiede in der Anordnung oder in den Eigenschaften der Systembestandteile gibt.

Jemand, der bereit wäre, die *These der synchronen Determiniertheit* der Systemeigenschaften abzulehnen, müsste entweder Eigenschaften zulassen, die nicht an die Anordnung der Bestandteile ihres Trägers und deren Eigenschaften gebunden sind, oder annehmen, dass ein anderer, in diesem Falle nichtnatürlicher Faktor für die unterschiedlichen Dispositionen struktur- und baustein-identischer Systeme verantwortlich ist. Beides erscheint ziemlich unplausibel.

Die Kognitionswissenschaft kennt eine ganze Reihe schwach emergenter Eigenschaften. So verfügen konnektionistische Netze als Gesamtsysteme über so bemerkenswerte Fähigkeiten wie das Generieren und Befolgen von Regeln, das Bilden von Schemata und das Erkennen von Mustern – alles Eigenschaften und Verhaltensweisen, die ihre einzelnen Bestandteile (ihre Einheiten und Verbindungen) nicht haben. Auch aus der Robotik und der *Artificial-Life*-Forschung (*A-Life*) sind zahlreiche Systeme bekannt, die über Eigenschaften verfügen, die ihre Bestandteile nicht haben: So können Roboter u. a. Fußball spielen, Autos montieren, bei Operationen assistieren und Sprachspiele beginnen. Im *A-Life* weichen *boids* (künstliche Vögel) in rasanten Schwarmbewegungen verschiedensten Hindernissen aus, andere künstliche »Lebewesen« bauen Nester, wie wir sie sonst nur von Termiten kennen.

Es ist jedoch nicht besonders aufschlussreich, wenn wir erfahren, dass viele Eigenschaften von konnektionistischen Netzen, Robotern

und *A-Life*-Kreaturen im schwachen Sinne emergent sind, denn es gibt schlicht zu viele Eigenschaften, die in diesem Sinne »emergent« genannt werden können. Die Welt ist voller Eigenschaften, die nur auf der Systemebene und nicht auf der Ebene der Systemkomponenten auftreten. Auch die Härte eines Diamanten ist in diesem Sinne eine emergente Eigenschaft, da sie nur dem Gesamtsystem und nicht den Kohlenstoffatomen zukommt, die in etwas anderer Anordnung den sehr viel weicheren Graphit instantiieren. Der schwache Emergentismus »schneidet« die Natur zwar an ihren »Scharnieren«, doch diese sind zu zahlreich: Es ist deshalb nichtssagend, wenn die für Kognitionswissenschaftler besonders interessanten Systemeigenschaften das Prädikat »schwach emergent« erhalten.

Eine Möglichkeit der Verstärkung des schwachen Emergentismus besteht darin, die sowohl im Rahmen evolutionärer Vorgänge als auch bei der Entwicklung neuer Artefakte wichtige *These der Neuartigkeit* hinzuzufügen und den schwachen Emergentismus damit um einen diachronen Aspekt zu ergänzen:

> *Neuartigkeit: Im Universum kommt es immer wieder zur Entstehung von genuin Neuartigem. Bereits bestehende Entitäten fügen sich zu neuen Konstellationen; sie bilden neue Strukturen aus, die neue Systeme mit neuen Eigenschaften und Verhaltensdispositionen konstituieren.*

Zwar werden mit dieser These zugleich alle präformationistischen Positionen ausgeschlossen, das bloße Hinzufügen der Neuartigkeitsthese macht jedoch aus einer schwachen noch keine starke Emergenztheorie, da der reduktive Physikalismus auch mit dieser Spielart des Emergentismus kompatibel bleibt. Erst wenn man eine weitere These, nämlich die der *prinzipiellen Unvorhersagbarkeit*, hinzufügte, erhielte man eine stärkere diachrone Emergenztheorie, die für die Kognitionswissenschaft von Bedeutung sein könnte. Auf diese komme ich weiter unten zurück.

2. Starke Emergenz

Lassen Sie mich aber zunächst etwas über synchrone Emergenztheorien sagen. Diese sind von großer Relevanz für die Diskussion des psychophysischen Problems, insbesondere für die Formulierung und Analyse nichtreduktiv physikalistischer Positionen in der Philosophie des Geistes.[3] Dort stehen Fragen nach dem Verhältnis von mentalen und physischen Eigenschaften im Zentrum des Interesses – zum Beispiel: Können mentale Eigenschaften wie das Haben von *intentionalen* oder *phänomenalen* Zuständen durch Rekurs auf eine physische Basis reduktiv erklärt werden? Antworten wir mit »Nein«, so vertreten wir eine starke emergentistische Position; wir behaupten, dass mentale Eigenschaften *irreduzibel* und damit *synchron emergent* sind. Aber was heißt es genau, dass eine Eigenschaft nicht reduktiv erklärt werden kann?

Die Frage nach reduktiven Erklärungen stellt sich üblicherweise, wenn wir verstehen wollen, weshalb eine bestimmte Entität eine bestimmte Eigenschaft hat, und zwar eine Eigenschaft, die in der Regel nur dem Systemganzen zugeschrieben wird. In trivialen Fällen genügt die einfache Addition der entsprechenden Eigenschaften der Komponenten. Auf diese Weise ergibt sich das Gewicht eines Fahrzeugs als die Summe des Gewichtes seiner Teile. Das Fahrverhalten des Wagens in engen Kurven – eine ungleich interessantere Eigenschaft – lässt sich dagegen nicht so einfach ableiten. Denn hier kommt es entscheidend darauf an, wie die einzelnen Bestandteile, insbesondere die Teile des Fahrwerks angeordnet sind, welche Eigenschaften diese haben und wie sie infolgedessen untereinander und mit der Straße interagieren. Ergibt sich das Fahrverhalten unter Zugrundelegung der allgemein geltenden Naturgesetze aus diesen Parametern, so gilt es als reduktiv erklärt. Noch komplizierter gestalten sich solche Erklärungen für die Eigenschaften sich selbst organisierender dynamischer Systeme, insbesondere für Lebewesen.

Das Ziel einer reduktiven Erklärung besteht also darin, die interessierende systemische Eigenschaft allein durch Rekurs auf die Bestandteile des Systems, deren Eigenschaften und Dispositionen, sowie deren Anordnung zu erklären (und gegebenenfalls vorherzu-

3 Im Rahmen seiner Diskussion des *Mechanismus* als einer metaphysischen Theorie formulierte erstmals Charles D. Broad explizit eine synchrone Emergenztheorie. Er charakterisierte sowohl verschiedene chemische als auch biologische und psychische Eigenschaften als mechanisch nicht erklärbare, emergente Eigenschaften (vgl. Broad, *The Mind and Its Place in Nature*, Kapitel II, sowie Stephan, *Emergenz*, § 3.5). Nach den großen Fortschritten in den Naturwissenschaften schieden jedoch die ursprünglich für emergent gehaltenen chemischen und biologischen Eigenschaften aus dem Kandidatenkreis für synchrone Emergenz aus; vgl. McLaughlin, »The Rise and Fall of British Emergentism«.

sagen). Sie ist erfolgreich, wenn sie die folgenden drei Bedingungen erfüllt:

(1) die zu reduzierende Eigenschaft ist funktional (re)konstruierbar;
(2) es lässt sich zeigen, dass die funktionale/kausale Rolle der zu reduzierenden systemischen Eigenschaft durch die Interaktionen der Bestandteile des Systems erfüllt wird;
(3) das Verhalten der Teile innerhalb des Systems ergibt sich aus deren Eigenschaften und Verhalten in Isolation oder in einfacheren Systemen.

Manche Autoren, zu denen u. a. Jaegwon Kim und Joseph Levine zählen, verzichten auf die letzte dieser drei Bedingungen. In jedem Falle ist jedoch erforderlich, dass wir die zu erklärende Eigenschaft beziehungsweise das zu erklärende Verhalten zunächst in der richtigen Weise begrifflich »präparieren«.[4] So fordert Kim als ersten Schritt auf dem Weg zu einer reduktiven Erklärung:

> Funktionalisiere die Eigenschaft, die reduziert werden soll, das heißt charakterisiere die Eigenschaft anhand ihrer kausalen Rolle. Diese Charakterisierung soll in Begriffen der Eigenschaften der Basisebene vorgenommen werden.[5]

Gesucht ist demnach eine funktionale Charakterisierung der reduktiv zu erklärenden Eigenschaft, auf die wir uns gewöhnlich mit einem Begriff beziehen, der Eigenschaften auf der systemischen Ebene klassifiziert, wobei uns spezifische Muster (wie zum Beispiel Schwarmbewegungen) Kenntnis von einer Instantiierung dieser Eigenschaft geben. Der Vorschlag der begrifflichen Präparierung hat das Ziel, den Übergang von der Komponenten-Ebene zur System-Ebene zu ermöglichen. Schlägt diese fehl, so scheitert die angestrebte reduktive Erklärung.

Auch Levine favorisiert im Rahmen seiner Explikation des *explanatory-gap*-Argumentes eine Unterteilung der Reduktionsprozedur in zwei Stufen:[6]

4 «To reduce a property *M* to a domain of base properties we must first ›prime‹ *M* for reduction by construing, or reconstruing, it relationally or extrinsically. This turns *M* into a relational/extrinsic property« (Kim, *Mind in a Physical World*, S. 98).

5 Kim, »Emergenz«, S. 156.

6 Sowohl Kim als auch Levine verzichten auf die dritte Reduktionsbedingung. Wie sich im nächsten Abschnitt zeigen wird, ist diese im Rahmen ihrer Zielsetzung nicht nötig.

> Stage 1 involves the (relatively? quasi?) *a priori* process of working the concept of the property to be reduced ›into shape‹ for reduction by identifying the causal role for which we are seeking the underlying mechanisms. Stage 2 involves the empirical work of discovering just what those underlying mechanisms are.[7]

Reduktive Erklärungen können in zwei gegenläufigen Richtungen gegeben werden. Ist bereits bekannt, dass ein System *S* eine bestimmte systemische Eigenschaft *E* hat, so besteht die Aufgabe darin, unter Berücksichtigung der Mikrostruktur *MS*(*S*) von *S*, der basalen Naturgesetze und der Interaktionsgesetze, die für die Komponenten K_i von *S* gelten, zu zeigen, dass *S* Eigenschaft *E* haben muss. Dabei ist von adäquaten begrifflichen Präparierungen von *E* Gebrauch zu machen. Ist dagegen (wie häufig bei Artefakten) ein System *S* in Entwicklung oder mit seiner Mikrostruktur *MS*(*S*) gegeben, so besteht die Aufgabe darin, unter Berücksichtigung der (geplanten) Mikrostruktur *MS*(*S*) von *S*, der basalen Naturgesetze sowie der Interaktionsgesetze, die für die Komponenten K_i von *S* gelten, zu ermitteln, ob *S* die (gewünschte) Makroeigenschaft *E* hat. Auch in diesem Fall ist von adäquaten begrifflichen Präparierungen von *E* Gebrauch zu machen.

Scheitern reduktive Erklärungen aus prinzipiellen Gründen, so ist die reduktiv zu erklärende systemische Eigenschaft irreduzibel und damit *synchron* (und das heißt: in einem starken Sinne) *emergent*.

Ausgehend von den drei Bedingungen für reduktive Erklärungen gibt es folglich auch drei verschiedene Möglichkeiten, weshalb eine Eigenschaft *irreduzibel* sein kann:

> *Irreduzibilität: Eine systemische Eigenschaft ist irreduzibel, wenn sie (1) nicht funktional (re)konstruierbar ist, oder wenn (2) nicht gezeigt werden kann, dass die Systemkomponenten die (re)konstruierte funktionale Rolle erfüllen; oder wenn sich (3) das Verhalten der Systemkomponenten, über dem die systemische Eigenschaft supervenient, nicht aus dem Verhalten der Komponenten ergibt, das diese in Isolation oder in einfacheren Konfigurationen zeigen.*

Die *These der Irreduzibilität* präzisiert in moderner Terminologie die gleichen Bedingungen für synchrone Emergenz, die implizit bereits in Broads Unterscheidung zwischen mechanistischen und emergentistischen Theorien enthalten sind. In einer geradezu »klassisch« zu

7 Levine, »On Leaving Out What It's Like«, S. 132.

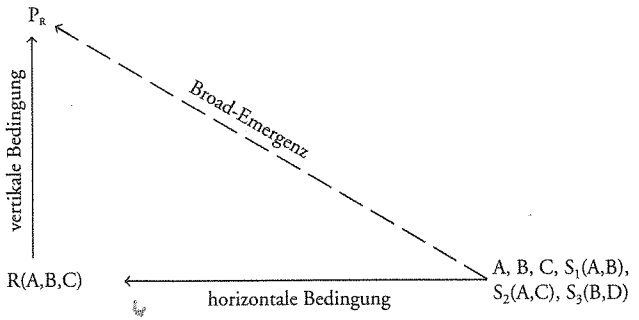


Abb. 1: P_R bezeichnet die reduktiv zu erklärende systemische Eigenschaft; *A*, *B* und *C* referieren auf die Teile, aus denen das System *R(A,B,C)* besteht; und $S_1(A,B)$, $S_2(A,C)$ und $S_3(B,C)$ stehen für einfachere Systeme als *R(A,B,C)*. Die Diagonale entspricht Broads Idee der synchronen Emergenz, der vertikale Pfeil repräsentiert Bedingung (2), der horizontale Pfeil die in der Philosophie des Geistes üblicherweise nicht in Betracht gezogene Bedingung (3).

nennenden Textpassage aus *The Mind and its Place in Nature* heißt es:

Put in abstract terms the *emergent theory* asserts (i) that there are certain wholes, composed (say) of constituents A, B, and C in a relation R to each other; (ii) that all wholes composed of constituents of the same kind as A, B, and C in relations of the same kind as R have certain characteristic properties; (iii) that A, B, and C are capable of occurring in other kinds of complex where the relation is not the same kind as R; and (iv) that the characteristic properties of the whole R(A,B,C) *cannot*, even in theory, *be deduced* from the most complete knowledge of the properties of A, B, and C in isolation or in other wholes which are not of the form R(A,B,C). The *mechanistic theory* rejects the last clause of this assertion.[8]

Das oben stehende Diagramm (*Abbildung 1*), das ich leicht abgewandelt aus Boogerd u. a. (2005) entnehme, kontrastiert in einer schematischen Darstellung Broads Definition der synchronen Emergenz mit den drei explizit formulierten Bedingungen der Irreduzibilitätsthese.[9]

8 Broad, *The Mind and Its Place in Nature*, S. 61 (Hervorhebung und explizite Aufzählung A. S.).

9 Eine ausführlichere Analyse dieser Beziehungen findet sich in Boogerd u. a., »Emergence«. Darin gehen wir auch auf einige detailreichere Überlegungen Broads ein, wie sie unter anderem in seinem zeitlich früheren Aufsatz »Mechanical Explanation and its Alternatives« zu finden sind.

2. Emergenz in der Philosophie des Geistes und in der Kognitionswissenschaft

Unter den reduktiv zu erklärenden Eigenschaften, Dispositionen und Verhaltensweisen nehmen mentale Eigenschaften eine Sonderstellung ein: Sie gelten als auffallend widerspenstig. Einige Autoren zweifeln gar daran, dass es jemals gelingen wird, bestimmte mentale Eigenschaften wie zum Beispiel das Haben phänomenaler Erlebnisse reduktiv zu erklären.[10] Wer an einer reduktiven Erklärung mentaler Phänomene interessiert ist, der sollte daher im Prinzip *alles* zur Verfügung stellen, was überhaupt als Reduktionsbasis dienen kann, also: das vollständige physische System mit seinen Komponenten, deren Arrangement und Interaktionen sowie verhaltensrelevante Eigenschaften der Umwelt. Hierin liegt auch der Grund, weshalb in der Philosophie des Geistes der dritten (horizontalen) Bedingung für Emergenz, die die Reduktionsbasis weiter einschränkt, keine Beachtung geschenkt wird. Aber selbst wenn wir alles über die neuronalen Korrelate bewusster psychischer Vorgänge wüssten, selbst wenn wir angeben könnten, wie sich jedes einzelne Neuron und jede einzelne Synapse verhält, scheint nicht ausgemacht zu sein, dass wir dann in der Lage wären, die mit diesen Vorgängen korrelierten mentalen Eigenschaften reduktiv zu erklären. Der Grund dafür sind nicht etwa mangelnde Einsichten der Neurowissenschaftler, sondern vielmehr, dass es nicht zu gelingen scheint (und zwar im Prinzip nicht), phänomenale Qualitäten begrifflich adäquat über ihre funktionale Rolle zu »präparieren«. Dies ist zumindest der Tenor von Levines Analyse:

What seems to be responsible for the explanatory gap, then, is the fact that our concepts of qualitative character do not represent, at least in terms of their psychological contents, causal roles. [...] Thus, to the extent that there is an element in our concept of qualitative character that is not captured by features of its causal role, to that extent it will escape the explanatory net of a physicalistic reduction.[11]

Die Qualia-Debatte verdeutlicht, dass zumindest in der Philosophie

10 Vgl. Stephan, »Phänomenaler Pessimismus« und »Phänomenale Eigenschaften«.

11 Levine, »On Leaving Out«, S. 134.

des Geistes ein starker Begriff von Emergenz unverzichtbar ist. Er behält auch dann noch seine klassifikatorische Bedeutung für nichtreduktiv physikalistische Positionen, wenn es gelingen sollte, die explanatorische Lücke teilweise zu schließen.[12] In der Kognitionswissenschaft scheint ein starker (synchroner) Emergenzbegriff freilich ohne Anwendung zu sein. Wie ich bereits an anderem Ort zeigte,[13] haben konnektionistische Netzwerke keine synchron emergenten Eigenschaften, sie sind nicht in einem starken Sinne emergent. Alle interessanten Eigenschaften wie das Befolgen von Regeln, die Schemabildung und das Erkennen von Mustern sind Eigenschaften, die vollständig aus der Struktur des Netzes, den Eigenschaften seiner Einheiten (ihrer Aktivierungsfunktion) und den Eigenschaften ihrer Verbindungen (der Verteilung der Gewichte) abgeleitet werden können. Ebenso wenig haben *A-Life*-Kolonien und Roboter synchron emergente Eigenschaften. Fähigkeiten wie das Einsammeln von Dosen, das Beginnen von Sprachspielen oder das Assistieren bei Operationen sind über ihre funktionale Rolle spezifiziert und vollständig reduktiv erklärbar. Das mitunter erstaunliche Makroverhalten von Kolonien künstlicher Lebewesen, das den Schwarmbewegungen von Vögeln und Fischen oder dem Verhalten beim Bau von Termitennestern gleicht, wird zum Teil durch ganz einfache Interaktionsmuster erzeugt. Für die von Craig Reynolds[14] kreierten *boids* gelten drei Verhaltensregeln, auf die sich das gesamte Schwarmverhalten zurückführen lässt: (1) Vermeide Zusammenstöße, (2) gleiche deine Geschwindigkeit der der benachbarten *boids* an, und (3) bleibe in der Nähe der benachbarten *boids*. Von analoger Einfachheit sind die Regeln, die Mitchel Resnick[15] zufolge genügen, um höchst komplexe Termitenbauten ent-

12 Durchaus Erfolg versprechend sind Überlegungen von Susan Hurley und Alva Noë (»Neural Plasticity and Consciousness«, S. 132), im Hinblick auf phänomenale Qualitäten die *intramodale* Vergleichslücke (Warum gibt eine bestimmte neuronale Aktivität Anlass zu einer *Rot*-Empfindung und nicht zu einer *Grün*-Empfindung?) und die *intermodale* Vergleichslücke (Warum gibt eine bestimmte neuronale Aktivität Anlass zu einer *visuellen* Empfindung und nicht zu einer *auditiven* Empfindung?) zu schließen. Zunächst unberührt bleibt davon allerdings die *absolute* explanatorische Lücke (Warum gibt eine bestimmte neuronale Aktivität überhaupt Anlass zu einer Empfindung?).

13 Stephan, *Emergenz*, Kapitel 17.

14 Vgl. Reynolds, »Flocks, Herds and Schools«.

15 Vgl. Resnick, *Turtles, Termites, and Traffic Jams*.

stehen zu lassen. Die systemischen Eigenschaften sind in diesen Fällen also gerade nicht unerklärbar. Verblüffend ist vielmehr, nach welch einfachen Regeln sie zu realisieren sind.

Für die Kognitionswissenschaft sind somit weder der schwache Emergentismus noch der starke (synchrone) Emergentismus besonders hilfreich, wenngleich aus entgegengesetzten Gründen. Abgesehen vom gut begründeten Gebrauch des starken Emergenzbegriffs in der Philosophie des Geistes findet dieser in den anderen Bereichen der Kognitionswissenschaft keine weitere Verwendung. Obgleich auch der synchrone Emergentismus die Natur an einem ihrer »Scharniere zu schneiden« scheint, könnten phänomenale Erlebnisse die einzigen Eigenschaften sein, die sich in diesem Sinne als stark emergent erweisen. Der synchrone Emergentismus ist deshalb nur dann für die Kognitionswissenschaft von Bedeutung, wenn sie sich selbst mit phänomenalen Qualitäten befasst, also zum Beispiel mit der Frage auseinander setzt, ob auch künstliche Systeme eine Innen-Perspektive oder Empfindungen haben können. In allen anderen Fragen ist der starke Emergentismus von keiner Relevanz für die Kognitionswissenschaft. Die übrigen von ihr studierten Phänomene sind nicht stark emergent. Und dies festzustellen ist noch nicht einmal sehr aufschlussreich. Denn es gibt nahezu keine systemische Eigenschaft, die stark emergent ist.

3. Keine Emergenz für *boids* und Roboter?

Sollten wir daraus nun schließen, dass es entgegen der Annahme vieler Vertreter der *A-Life*-Forschung und der Kognitionswissenschaft keine »Emergenz« für *boids* und Roboter gibt? Boden[16] hatte den Begriff der Emergenz einmal zu den »Schlüsselbegriffen der *A-Life*-Forschung« gezählt. Und auch viele andere prominente Kognitionswissenschaftler wie Luc Steels, Peter Cariani oder David E. Rumelhart und James L. McClelland,[17] um nur einige zu nennen, hatten dem Begriff der Emergenz einiges zugetraut. Aus der bisherigen Untersuchung lässt sich jedoch das Fazit ziehen, dass der Begriff

16 Vgl. Boden, »Introduction«.

17 Vgl. Steels, »Towards a Theory of Emergent Functionality« und »The Artificial Life Roots of Artificial Intelligence«; Cariani, »Emergence and Artificial Life« oder Rumelhart und McClelland, »PDP Models«.

der schwachen Emergenz zu schwach ist (es gibt zu viele systemische Eigenschaften) und dass der synchrone Emergenzbegriff zu stark ist (es gibt zu wenige synchron emergente Eigenschaften), um für die Kognitionswissenschaft und die *A-Life*-Forschung relevant zu sein. Nun könnte man geneigt sein zu schließen, dass sich die »Emergenz-Freunde« unter den Kognitionswissenschaftlern gehörig getäuscht haben – hinsichtlich der Aussagekraft des Emergenzbegriffs in ihrer Disziplin. Zutreffender dürfte freilich die Annahme sein, dass die meisten von ihnen – wenn auch nicht explizit formuliert – einfach einen anderen Emergenzbegriff vor Augen hatten. Zu fragen ist dann, welcher Begriff von Emergenz in der Lage ist, die *Welt* der Artefakte wie die der Roboter und der *A-Life*-Kreaturen an aussagekräftigeren Stellen zu »schneiden«.

Einen guten Anknüpfungspunkt bietet hierfür Andy Clark, der in *Mindware* vier alternative Konzeptionen vorstellt und diskutiert, die seiner Ansicht nach in der einschlägigen Literatur zumindest implizit eine Rolle gespielt haben: (1) Emergenz als *kollektive Selbstorganisation*, (2) Emergenz als *nichtprogrammierte Funktionalität*, (3) Emergenz als *interaktive Komplexität* und (4) Emergenz als *nichtkomprimierbare Entwicklung*.[18]

Emergenz im Sinne *kollektiver Selbstorganisation* liege dann vor, wenn durch die Interaktionen sehr vieler gleichartiger Komponenten höchst bemerkenswerte Eigenschaften auf der Systemebene realisiert werden: zum Beispiel Schwarmbewegungen durch *boids* oder Termitenbauten durch andere *A-Life*-Kreaturen. Allgemein werden Phänomene dieser Art von der Synergetik (der »Lehre vom Zusammenwirken«) untersucht. Weitere bekannte Beispiele sind die Bildung von Mustern (Konvektionsrollen) in einseitig erhitzten Flüssigkeiten, die Strukturbildung beim Schleimpilz oder das Auftreten des gegenüber dem Lampenlicht völlig neuartigen Laserlichts.[19] In mathematischen Beschreibungen gelingt es häufig, die unübersehbare Zahl von möglichen Variablen auf der Komponentenebene auf wenige »kollektive Variablen« (die so genannten Ordnungsparameter) zu reduzieren und dennoch zu adäquaten Beschreibungen der untersuchten Phänomene zu gelangen. Aus der epistemischen Möglichkeit, die zur Systembeschreibung relevante Information auf wenige Ordnungsparameter komprimieren zu können, haben einige Autoren weit reichende metaphysische Schlüsse gezogen und den Ordnungsparametern kausale Kräfte zugeschrieben. So behauptet Haken in seinen eher populärwissenschaftlichen Werken, dass die auf der Systemebene zu lokalisierenden Ordnungsparameter im Modus »zirkulärer Kausalität« die Systemkomponenten »versklaven«. Diese stark metaphorische, in der Sache aber irreführende Redeweise suggeriert, dass die Komponenten durch ihr Zusammenwirken Ordnungsparameter etablieren, die dann ihrerseits die Komponenten zu einem bestimmten Verhalten zwingen. Die Ordnungsparameter erwiesen sich damit als emergente Eigenschaften mit »abwärts wirkenden« kausalen Kräften. Aber diese Sichtweise ist falsch. Wie das Beispiel der *boids* zeigt, ist es weder der Schwarm als solcher noch einer seiner »Ordnungsparameter«, die einem einzelnen Vogel vorschreiben, wie er zu fliegen hat. Vielmehr sind die den Schwarm konstituierenden Vögel vom Flugverhalten ihrer jeweiligen Nachbarn wechselseitig abhängig. Genau dies belegen die einfachen Verhaltensregeln, mit deren Hilfe Reynolds Vogelschwärme auf virtuelle Reisen schickte.

Welcher Begriff von Emergenz lässt sich dann aber aus diesen Beispielen extrahieren? Clarks Explikationsvorschlag ist eher ernüchternd:

> An emergent phenomenon is thus any interesting behavior that arises as a direct result of multiple, self-organizing interactions occurring in a system of simple elements.[20]

Auch wenn es verblüffen mag, welch »interessante« Muster und Strukturen spontan und ohne intendiertes Design aus dem bloßen Zusammenwirken einer Vielzahl ähnlicher Komponenten von selbst entstehen können, hat der von Clark vorgeschlagene Emergenzbegriff keine theoretische Relevanz. Er bleibt ganz einer subjektiv-phänomenologischen Ebene verhaftet und hebt sich kaum von seinem umgangssprachlichen Gebrauch ab, wonach er so viel bedeutet wie »unerwartet auftauchen oder in Erscheinung treten«. Darüber hinaus hat er keine Anwendung bei Robotern und konnektionistischen Netzen, denn dort kommt es entscheidend auf deren Design beziehungsweise die in Trainingsphasen aufgebauten Strukturen an.

18 Vgl. Clark, *Mindware*, S. 112-117.

19 Vgl. Haken, *Erfolgsgeheimnisse der Natur* sowie Stephan, *Emergenz*, § 18.1.

20 Vgl. Clark, *Mindware*, S. 114.

Der zweite Vorschlag geht ursprünglich auf Steels[21] zurück und expliziert Emergenz als *nichtprogrammierte Funktionalität*. Darunter ist zielgerichtetes adaptives Verhalten zu verstehen, das aus zumeist iterierten Interaktionen des Gesamtsystems mit seiner Umgebung entsteht und nicht allein das Ergebnis systeminterner Steuerungsprozesse oder expliziter Programmierung ist. Die im System repräsentierten Verhaltensoptionen referieren dabei nicht auf das *de facto* verfolgte Ziel (zum Beispiel das Einer-Wand-Folgen), sondern enthalten wesentlich einfachere Anweisungen (wie »bounce and veer« – bewege dich mit einem spezifischen Drall, erreichst du ein Hindernis, so drehst du dich unter einem bestimmten Winkel, der dem Drall entgegenwirkt, setze deine Bewegung mit dem ursprünglichen Drall fort usw.). Erreicht ein mit der *Bounce-and-veer*-Technik ausgestatteter Roboter eine Wand, so folgt er dieser. Sein spezifisches Verhalten (Geschwindigkeit, Kursabweichung) kann jedoch nur indirekt beeinflusst werden, da es nicht von zentralen oder expliziten Kontrollstrukturen abhängt – es basiert auf »unbeherrschten Variablen«. Dennoch muss es aus Sicht eines Ingenieurs weder unerwartet noch ungeplant sein. Es ist vielmehr ein einfaches Beispiel für *embodied and embedded cognition* – das adaptive Verhalten zeigt sich in einer geeigneten Umwelt, in der das System mit seinem Körper fortgesetzt interagiert. Wahrscheinlich ist es diese Beobachtung: Es ist das Realisieren eines zielgerichteten Verhaltens, das die intern repräsentierten Zielvorgaben transzendiert und nur aus dem Zusammenspiel des Systems mit seiner Umgebung entspringt, das dem Verhalten einen Anschein von *Emergenz* verleiht. In gleicher Weise dürfte auch das Verhalten von Insekten nichtprogrammierte Funktionalität besitzen. Dennoch ist es, wenn man die Mechanismen kennt, vollständig reduktiv erklärbar.

Ob das Merkmal der *nichtprogrammierten Funktionalität* zu einem theoretisch gehaltvollen Begriff führen kann, der zwischen schwacher und starker Emergenz anzusiedeln wäre, muss sich freilich erst noch zeigen. Ich habe allerdings Bedenken. Einerseits dürfte es schwer werden, einen solchen Emergenzbegriff bei nichtmenschlichen Organismen adäquat anzuwenden, denn es bedürfte jeweils eines Vergleiches der im System repräsentierten Ziele und Strategien

21 Vgl. Steels, »Towards a Theory of Emergent Functionality« und »Artificial Life Roots of Artificial Intelligence«.

mit den in einer passenden Welt *de facto* erreichten. Andererseits scheinen unzählige menschliche Verhaltensweisen in diesem Sinne emergent zu sein. Die im Zuge emotionaler Reaktionen ablaufenden Körperveränderungen, um nur ein Beispiel zu nehmen, haben offenbar interne und externe Signalfunktionen, die als solche jedoch nicht in den zugrunde liegenden neuronalen und neurochemischen Mechanismen repräsentiert werden: Es bedarf des Anderen, der unsere Mimik und Gestik zu lesen versteht, damit er das Signalisierte versteht. Sind deshalb emotional kommunizierte Botschaften emergent?

Clark selbst scheint Emergenz eher als *interaktive Komplexität* konzipieren zu wollen. Nach dieser Idee sind diejenigen systemischen Eigenschaften, Muster und Verhaltensweisen als emergent auszuzeichnen, die durch komplexe, zumeist zyklische Interaktionen der Systemkomponenten instantiiert werden. Die Beachtung der *Komplexität* der Interaktionen erlaube zudem einen graduellen Emergenzbegriff. Danach ist Verhalten, das aus einer Abfolge linearer Interaktionen resultiert wie der *Bounce-and-veer*-Technik des einer Wand folgenden Roboters, nur *schwach* emergent. Phänomene, die auf multiplen nichtlinearen und zeitlich asynchronen Interaktionen mit positivem Feedback beruhen wie Konvektionsrollen bei einseitig erhitzten Flüssigkeiten, seien dagegen *stark* emergent.[22] Ohne Frage ist es ein lohnendes Ziel, einen graduellen Emergenzbegriff zu entwickeln, der die Komplexität von Strukturen und Interaktionen abbildet. Clarks Vorschlag hat jedoch die unplausible Konsequenz, dass ausgerechnet Konvektionsrollen und Schwarmbewegungen, die sich auf sehr einfache Interaktionsregeln zurückführen lassen, stark emergent wären. Komplex sind in diesen Fällen nicht die Strukturen oder das Verhalten der zu untersuchenden Systeme, sondern die mathematischen Beschreibungen, die nötig wären, das Systemverhalten auf der Ebene unzählig vieler einfacher Komponenten zu erfassen. Meines Erachtens ist es nicht sinnvoll, hier überhaupt von Emergenz in einem anderen als in dem von mir als *schwach* bezeichneten Sinne zu sprechen.

22 Vgl. Clark, *Mindware*, S. 115 f. Die von Clark als schwach beziehungsweise stark emergent ausgezeichneten Phänomene sind gemäß der von mir eingeführten Terminologie nur als *schwach* emergent anzusehen. Da sie im Prinzip reduktiv erklärt werden können, sind sie nicht *stark* emergent.

Der vierte, von Clark nach eigener Auskunft nur der Vollständigkeit halber vorgestellte Vorschlag thematisiert Emergenz als *nichtkomprimierbare Entwicklung*. Diese Idee stammt von Mark Bedau, der die folgende Definition vor allem an *Game-of-Life*-Konfigurationen illustriert hatte:[23]

> Macrostate *P* of *S* with microdynamic *D* is *weakly emergent* if *P* can be derived from *D* and *S*'s external conditions but only by simulation.[24]

Als »schwach emergent« werden damit diejenigen systemischen Zustände *P* eines Systems *S* ausgezeichnet, die nur durch eine Simulation aller auf der Komponentenebene stattfindenden Interaktionen sowie der äußeren Einflüsse auf *S* abgeleitet werden können. Nach Clarks Ansicht ist dieser Vorschlag zu restriktiv,[25] da er die von ihm selbst als *stark emergent* charakterisierten Phänomene (Schwarmbewegungen, Konvektionsrollen) noch nicht einmal als *schwach emergent* auszeichnen würde. In diesen Fällen ist es nämlich möglich, die interessierenden systemischen Zustände durch eine geeignete Wahl der Ordnungsparameter adäquat zu beschreiben. Eine vollständige Simulation der ihnen zugrunde liegenden Mikrodynamik ist nicht nötig. Da ich, wie gerade ausgeführt, selbst Bedenken gegenüber dem von Clark favorisierten Begriff der (starken) Emergenz habe, halte ich die deflationären Konsequenzen, die sich aus Bedaus Begriffsanalyse für die von Clark für *stark emergent* gehaltenen Phänomene ergeben, eher für begrüßenswert.

Bedau charakterisiert den von ihm explizierten Begriff als »schwach« emergent, um ihn von den starken Varianten abzugrenzen, die in der Philosophie des Geistes Verwendung finden. Explizit bezieht er sich dabei auf Timothy O'Connor, der systemische Eigenschaften als *stark emergent* definiert, wenn sie über irreduzible und abwärts gerichtete kausale Kräfte verfügen.[26] Mein eigener Vorschlag begreift *starke* Emergenz zwar über das Merkmal der Irreduzibilität und nicht über abwärts gerichtete Kausalität; da die von Bedau diskutierten Eigenschaften jedoch reduktiv erklärt werden können,

23 Das *Game of Life* ist ein von dem Mathematiker Conway entworfener zweidimensionaler zellulärer Automat, in dem sich mit einfachen Regeln komplexe Strukturen und Mechanismen definieren lassen, die interessante Ähnlichkeiten zu biologischen Systemen aufweisen.

24 Vgl. Bedau, »Weak Emergence«, S. 378.

25 Vgl. Clark, *Mindware*, S. 116.

26 Vgl. O'Connor, »Emergent Properties«, S. 97.

sind sie auch in dem von mir definierten Sinne nicht *stark* emergent. Gleichwohl ist der von Bedau vorgeschlagene Begriff der *schwachen* Emergenz anspruchsvoller als der, den ich allen weiteren Spielarten des Emergentismus zugrunde gelegt habe. Denn sehr viele systemische Eigenschaften, die nach meinem Vorschlag schwach emergent sind, können auch ohne Simulation der Mikrodynamik reduktiv erklärt werden und sind damit nach Bedau noch nicht einmal schwach emergent.

Bedaus schwache Emergenztheorie hat große Gemeinsamkeiten mit einer Theorie der Emergenz, die ich an anderer Stelle durch die *These der Struktur-Unvorhersagbarkeit* charakterisiert und als *diachronen Struktur-Emergentismus* bezeichnet habe.[27] Diese Variante der Unvorhersagbarkeit hat zwar in der klassischen emergenztheoretischen Literatur keine Rolle gespielt, gewinnt jedoch gerade durch die in der Robotik und der *A-Life*-Forschung untersuchten Phänomene zunehmend an Bedeutung:

> *Struktur-Unvorhersagbarkeit: Das Entstehen neuartiger Strukturen ist prinzipiell unvorhersagbar (und damit emergent), wenn die Bildung der Strukturen den Gesetzen des deterministischen Chaos folgt. Ebenso wenig sind die Eigenschaften vorhersagbar, die durch die neuen Strukturen instantiiert werden.*

Da es aus theoretischer Perspektive wenig interessant wäre, wenn die Bildung einer neuen Struktur nur deshalb als unvorhersagbar gelten müsste, weil ihr Entstehen indeterminiert wäre, haben sich Emergentisten gefragt, ob auch determiniert ablaufende Strukturbildungen *prinzipiell unvorhersagbar* sein können. Sie sind es dann, wenn sie nach Gesetzen verlaufen, die dem deterministischen Chaos zuzurechnen sind. Ein wesentlicher Befund der Chaostheorie ist nämlich, dass es (sogar sehr einfache) mathematische Funktionen gibt, deren iteriertes Verhalten nicht vorhergesagt werden kann. Erst das Aufkommen »experimenteller Computer-Mathematik« hat zum Beispiel die Eigenschaften von verschiedenen logistischen Funktionen preisgegeben. Deren intramathematische Unvorhersagbarkeit hängt mit dem irregulären Verhalten dieser Funktionen zusammen, bei dem minimal verschiedene Startwerte zu radikal verschiedenen Funktionsverläufen führen können. Entsprechend gilt, dass Systeme, deren Entwicklung den Gesetzen des deterministi-

27 Vgl. Stephan, *Emergenz*, § 3.6 und 18.3.

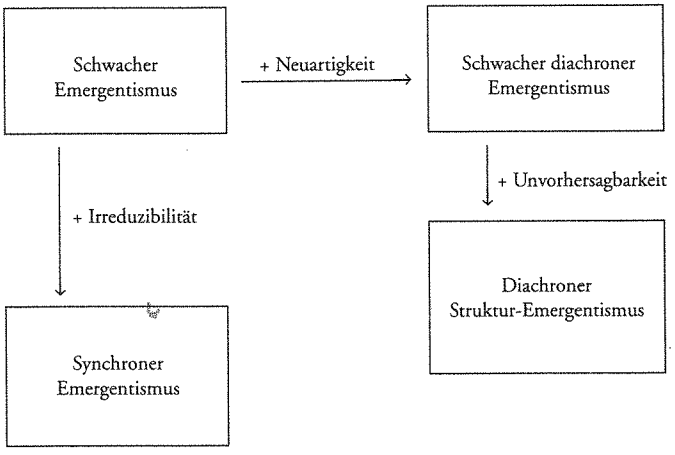


Abb. 2

schen Chaos folgt, prinzipiell unvorhersagbare Strukturen ausbilden und damit emergentes Verhalten zeigen können. Und hier treffen sich Bedaus begriffliche Vorstellungen mit meinen: Will man Phänomene beschreiben und vorhersagen, die den Gesetzen des deterministischen Chaos folgen, so sind vollständige Simulationen der mathematischen Funktionen auf der Komponentenebene gefordert; es stehen keine abkürzenden Informationskomprimierungen durch Ordnungsparameter zur Verfügung. Bedau hat die Fruchtbarkeit seines Ansatzes für die *A-Life*-Forschung am *Game of Life* vorgeführt. In der Kognitionswissenschaft könnte die *These der Struktur-Unvorhersagbarkeit* besonders bei der Charakterisierung *evolutionärer Architekturen* Anwendung finden – denn diese entwickeln sich häufig in einer Art und Weise, die prinzipiell unvorhersagbar ist.

Mit dem Begriff der Emergenz als *Struktur-Unvorhersagbarkeit* liegt folglich eine echte Alternative vor, die zwischen *schwacher* und *starker* Emergenz anzusiedeln ist und die *Welt der Artefakte* in der Kognitionswissenschaft an interessanten Scharnieren »schneidet«. Der *diachrone Struktur-Emergentismus* ist zwar eine anspruchsvolle Form des Emergentismus, aber durchaus noch mit reduktiven Erklärungen vereinbar: Sollten sich evolutionäre Architekturen ausbilden, die vor ihrer tatsächlichen Entwicklung unvorhersehbar waren, so ist nicht ausgeschlossen, dass sich ihre Eigenschaften und Verhaltensweisen aus der Struktur des Systems reduktiv erklären lassen.

Abbildung 2 stellt zusammenfassend die logischen Beziehungen dar, die zwischen den von mir vorgeschlagenen Spielarten des Emergentismus bestehen.

Der *schwache diachrone* Emergentismus entsteht aus dem *schwachen* Emergentismus durch Hinzufügen der Neuartigkeitsthese. Beide Versionen sind mit dem reduktiven Physikalismus kompatibel und wenig aussagekräftig. Der (starke) *synchrone* Emergentismus geht aus dem schwachen Emergentismus durch Hinzufügen der Irreduzibilitätsthese hervor; er wird vor allem in der Philosophie des Geistes benötigt. Der *Struktur-Emergentismus* entsteht aus dem *schwachen diachronen* Emergentismus durch Hinzufügen der These der Struktur-Unvorhersagbarkeit. Er ist vereinbar mit reduktiv physikalistischen Positionen und insofern schwächer als der synchrone Emergentismus, könnte aber einen breiten Anwendungsbereich in der Kognitionswissenschaft und *A-Life*-Forschung haben.

Literatur

Bedau, Mark A., »Weak Emergence«, in: *Philosophical Perspectives* 11, 1997, S. 375-399.

Boden, Margaret, »Introduction«, in: *The Philosophy of Artificial Life*, hg. v. Margaret Boden, Oxford 1996, S. 1-35.

Boogerd, Fred C., u. a., »Emergence and its Place in Nature: A Case Study of Biochemical Networks«, in: *Synthese* 145, 2005, S. 131-164.

Broad, Charles D., »Mechanical Explanation and its Alternatives«, in: *Proceedings of the Aristotelian Society* 19, 1919, S. 86-124.

–, *The Mind and Its Place in Nature*, London 1925.

Cariani, Peter, »Emergence and Artificial Life«, in: *Artificial Life II*, hg. v. Christopher G. Langton u. a., Reading, Mass. 1991, S. 775-797.

Clark, Andy, *Mindware. An Introduction to the Philosophy of Cognitive Science*, Oxford 2001.

Haken, Hermann, *Erfolgsgeheimnisse der Natur. Synergetik: Die Lehre vom Zusammenwirken*, Frankfurt am Main 1988.

Hurley, Susan, und Alva Noë, »Neural Plasticity and Consciousness«, in: *Biology and Philosophy* 18, 2003, S. 131-168.

Kim, Jaegwon, *Mind in a Physical World*, Cambridge, Mass. 1998.
–, »Emergenz, Reduktionsmodelle und das Mentale«, in: *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?*, hg. v. Michael Pauen und Achim Stephan, Paderborn 2002, S. 148-164.
Levine, Joseph, »On Leaving Out What It's Like«, in: *Consciousness*, hg. v. Martin Davies und Glyn W. Humphreys, Oxford 1993, S. 121-136.
McLaughlin, Brian, »The Rise and Fall of British Emergentism«, in: *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*, hg. v. Ansgar Beckermann, Hans Flohr und Jaegwon Kim, Berlin 1992, S. 49-93.
O'Connor, Timothy, »Emergent Properties«, in: *American Philosophical Quarterly* 31, 1994, S. 91-104.
Resnick, Mitchel, *Turtles, Termites, and Traffic Jams: Explorations in Massively Parallel Microworlds*, Cambridge, Mass. 1994.
Reynolds, Craig, »Flocks, Herds, and Schools: A Distributed Behavioral Model«, in: *Computer Graphics* 21, 1987, S. 25-34.
Rumelhart, David E., und James L. McClelland, »PDP Models and General Issues in Cognitive Science«, in: *Parallel Distributed Processing. Explorations in the Microstructure of Cognition.* Bd. 1, hg. v. David E. Rumelhart und James L. McClelland, Cambridge, Mass. 1986, S. 110-146.
Steels, Luc, »Towards A Theory of Emergent Functionality«, in: *From Animals to Animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior*, hg. v. Jean-Arcady Meyer u. a., Cambridge, Mass. 1991, S. 451-461.
–, »The Artificial Life Roots of Artificial Intelligence«, in: *Artificial Life* 1, 1994, S. 75-110.
Stephan, Achim, *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation*, Dresden 1999 (2. Auflage: Paderborn 2005).
–, »Phänomenaler Pessimismus«, in: *Phänomenales Bewusstsein – Rückkehr zur Identitätstheorie?*, hg. v. Michael Pauen und Achim Stephan, Paderborn 2002, S. 342-363.
–, »Phänomenale Eigenschaften, phänomenale Begriffe und die Grenzen reduktiver Erklärung«, in: *Grenzen und Grenzüberschreitungen. XIX. Deutscher Kongress für Philosophie. Vorträge und Kolloquien*, hg. v. Wolfram Hogrebe in Verbindung mit Joachim Bromand, Berlin 2004, S. 404-416.

Ralph Schumacher

Die prinzipielle Unterbestimmtheit der Hirnforschung im Hinblick auf die Gestaltung schulischen Lernens

1. Einleitung

Keine empirische Wissenschaft, die neue Bereiche der Wirklichkeit erschließt, kommt ohne die philosophische Reflexion ihrer Grundlagen aus. Auf Philosophie zu verzichten bedeutet in diesem Fall, die Grundlagen der betreffenden Disziplin ungeprüft vorauszusetzen. Dies mag in einigen begrenzten Bereichen keinen großen Schaden anrichten. Aber insgesamt betrachtet schränkt es die wissenschaftliche Vorstellungskraft erheblich ein, wenn sich das Denken im Rahmen überkommener philosophischer Voraussetzungen bewegt. In diesem Aufsatz soll am Beispiel der Beziehung der empirischen Hirnforschung zur psychologischen und pädagogischen Lehr-Lern-Forschung gezeigt werden, dass die kritische Reflexion ontologischer und wissenschaftstheoretischer Grundlagen der Hirnforschung erforderlich ist, um die Bedeutung neurophysiologischer Einsichten für die Debatten um die Verbesserung des Schulunterrichts richtig einschätzen zu können.

Unter dem Oberbegriff »das psychophysische Problem« werden eine ganze Reihe miteinander verwandter Fragestellungen verstanden, zu denen vor allem die Folgenden zählen:

(1) Gehören geistige und physikalische beziehungsweise neurophysiologische Zustände zu grundsätzlich verschiedenen ontologischen Kategorien?
(2) Wie lässt sich die Interaktion zwischen geistigen und physikalischen beziehungsweise neurophysiologischen Zuständen erklären?
(3) Lassen sich geistigen Zuständen kausale Eigenschaften zuschreiben, ohne sie damit auf physikalische oder neurophysiologische Zustände zu reduzieren?
(4) Können wir die Intentionalität geistiger Zustände im Rahmen einer naturalistischen Theorie des Geistes erklären?

Minds & Machines (2008) 18:443–459
DOI 10.1007/s11023-008-9122-6

Is Weak Emergence Just in the Mind?

Mark A. Bedau

Received: 4 December 2007 / Accepted: 5 May 2008 / Published online: 4 November 2008

Abstract Weak emergence is the view that a system's macro properties can be explained by its micro properties but only in an especially complicated way. This paper explains a version of weak emergence based on the notion of explanatory incompressibility and "crawling the causal web." Then it examines three reasons why weak emergence might be thought to be just in the mind. The first reason is based on contrasting mere epistemological emergence with a form of ontological emergence that involves irreducible downward causation. The second reason is based on the idea that attributions of emergence are always a reflection of our ignorance of non-emergent explanations. The third reason is based on the charge that complex explanations are anthropocentric. Rather than being just in the mind, weak emergence is seen to involve a distinctive kind of complex, macro-pattern in the mind-independent objective micro-causal structure that exists in nature. The paper ends by addressing two further questions. One concerns whether weak emergence applies only or mainly to computer simulations and computational systems. The other concerns the respect in which weak emergence is dynamic rather than static.

Keywords Weak emergence · Epistemological emergence · Dynamic emergence · Computational emergence · Micro-causal network · Micro-causal web · Explanatory incompressibility

M. A. Bedau (✉)
Department of Philosophy, Reed College, Vollum College Center 141, 3203 SE Woodstock Blvd., Portland, OR 97202-8199, USA
e-mail: mab@reed.edu

M. A. Bedau
European School of Molecular Medicine, Milan, Italy

The notion of emergence is enjoying a renaissance in philosophy and science today,[1] but it remains haunted by the worry that all apparent emergence in nature is really in one way or another just in the mind.[2] If emergent phenomena are just in the mind, then they are not real and objective phenomena; they have no independent ontological existence; they have no independent causal power; they have no objective reality outside the mind. The charge that emergence is just in the mind has recently been directed specifically at weak emergence, a view which I have developed and defended elsewhere,[3] and which I will explain and defend here. My main goal here is to show that weak emergence is not just in the mind. Rather, it is a distinctive kind of complex, macro-pattern in the mind-independent objective micro-causal structure that exists in nature. My secondary goal here is to answer two further questions about weak emergence. One concerns whether weak emergence is limited to computer simulations and computational systems. The other concerns how weak emergence is dynamic rather than static.

Weak emergence (as I intend it) can be explained in various ways.[4] My explanation here uses the concept of *explanatory incompressibility*. Previously I have defined weak emergence using the concept of underivability except by simulation (Bedau 1997; 2003). These two definitions are similarly indirect, and they are essentially equivalent. I support them both. Here I focus on explanatory incompressibility in order to highlight that weak emergence applies not just to computer simulations. It is true that weak emergence has a special connection with computer simulations, as we will see below, but it also applies equally well to a great many natural systems, especially those that motivated the original discussions of emergence by the British emergentists (McLaughlin 1992).

My defense of weak emergence will turn on its distinctively *dynamical* nature. In contrast to most recent discussions of emergence in philosophy, weak emergence is concerned with the complex dynamic processes by which certain global phenomena are generated. The dynamical nature of weak emergence helps explain the distinctive role that computer simulations play in both explaining natural emergent phenomena and artificially generating interesting new examples of emergent phenomena. It also helps explain why emergence is not merely in the mind. For if the dynamical causal processes distinctive of weak emergence are not merely in the mind, then neither is weak emergence itself.

[1] See, e.g., Feltz et al. (2006), Kistler (2006), Bedau and Humphreys (2008).

[2] Some who claim that weak emergence is just in the mind view this in a positive light (e.g., Newmann 1996 and McIntyre 1998), on the grounds that if weak emergence is just in the mind, then it surely exists and is real, and it has no objectionable metaphysical baggage. In this paper I am arguing that weak emergence is *not* just in the mind; it is real and objective in nature. Specifically, it is a certain kind of complex micro-causal network that has unpredictable macro effects. In addition, this kind of weak emergence has no metaphysical baggage. Each instance of weak emergence is entirely consistent with any reasonable form of naturalism.

[3] Bedau (1997, 2003).

[4] The different variants of weak emergence in the literature include those by Wimsatt (1986, 1997, 2000), Rueger (2000a, b), Boogerd et al. (2005), and Bedau (1997, 2003). Different conceptions of weak emergence focus on different kinds of explanatory complexity, but all agree that weak emergence involves some distinctive kind of explanatory complexity.

Weak Emergence as Explanatory Incompressibility

Emergence always involves a certain kind of relationship between global or macro-phenomena and local or micro-phenomena. Specifically, emergent macro-phenomena somehow both depend on, and are autonomous from, micro-phenomena. Dependence and autonomy can be given different interpretations, and different interpretations lead to different conceptions of emergence, including different conceptions of weak emergence.

The characteristic feature of weak emergence, in general, is that the macro is ontologically and causally reducible to the micro *in principle*, but the reductive micro-explanation is *especially complex*.[5] Different kinds of explanatory complexity create different versions of weak emergence.[6] Elsewhere I have characterized weak emergence as underivability without simulation.[7] Here I shift terms slightly and replace derivations with explanations, and replace macro-states that are underivable except by simulation with macro-states that have only incompressible explanations. Throughout I tacitly assume that we seek only true, exact, and complete explanations of how macro properties are generated from prior micro properties over time.

My definition of weak emergence is this, specifically: If P is a macro-property of some system S, then P is weakly emergent if and only if P is generatively explainable from all of S's prior micro-facts but only in an incompressible way. This definition defines weak emergent macro-phenomena by the distinctive way in which we explain how they are generated from underlying micro-states.

The basic idea of an *incompressible* generative explanation is simple. An explanation is generative just in case it exactly and correctly explains how macro-events unfold over time, how they are generated dynamically. The temporally (discrete or continuous) generative explanation assumes complete information about both the micro-causal dynamics that drive the system and the system's earlier micro-states and boundary conditions. The explanation works simply by tracing through the temporal details in the complex web of micro-level causal interactions that ultimately generate the macro-events. This kind of explanation is appropriate for any system with (global) macro-features that depend on (local) micro-features in certain complex ways. In particular, it is appropriate if we can describe the system's macro-features *at a given time* by appropriately conjoining or aggregating or summing the (local) micro-features that exist at the same time. This is a synchronic

[5] See, e.g., Simon (1996).

[6] Unfortunately, the phrase "weak emergence" is used in different ways by different philosophers. For example, what I call "weak emergence" differs from what is called "weak emergence" by Stephan (2006) or by K. Balog (personal communication), and it does not apply to typical Braitenberg's vehicles (an example of weak emergence from A. Beckerman, personal communication).

[7] See Bedau (1997, 2003). I should note that my earlier use of the word "derivation" to define weak emergence does not imply any connection with, or dependence on, a Hempelian Deductive-Nomological account of explanation, which is famous for its central role for derivations (Hempel 1965). Rather, I am content with a much more loose and informal notion of explanation. Nevertheless, explanation and deduction still use and rely on various derivations.

reduction of macro to micro. Now, by starting with a completely specified initial condition, and by tracing forwards in time through the network of local micro-level causal interactions, the system's macro-features (which are aggregation of micro-features at a given time) can be explained from immediately preceding aggregations of micro-features. Explaining the generation of a system's macro-behavior by aggregating and iterating the earlier local micro-interactions over time I shall describe, for short, as *crawling the micro-causal web*.

Incompressible explanations cannot be replaced without explanatory loss by shorter explanations that avoid crawling the causal web. If an explanation of some macro-property of some system is incompressible, then there is no short-cut generative explanation of that macro-property that is true, complete, accurate, and can avoid crawling the causal web.[8] This explains the temporal signature of incompressible explanations: Explaining later behavior requires additional explanatory effort (Crutchfield et al. 1986).

On the other hand, if the explanation is *compressible*, then explaining the macro-property arbitrarily far into the future takes some fixed and finite amount of explanatory effort, no matter how far into the future your explanation reaches. The required explanatory effort is capped; explaining later behavior takes no more explanatory effort than explaining earlier behavior. A compressible explanation can achieve these economies because it avoids the incrementally growing cost of crawling the causal web.

Examples can vividly convey what a compressible explanation is, and the simplest and clearest examples involve cellular automata. A cellular automaton (or CA) is a regular spatial lattice of "cells," each of which can be in one of a finite number of states (e.g., "alive" or "dead"). The lattice typically has 1, 2, or 3 spatial dimensions. The state of each cell in the lattice is updated simultaneously in discrete time steps. Each cell is a finite state machine that outputs the next state of the cell given as input the states of the cells within some finite, local neighborhood of the lattice. Typically all cells in the lattice are governed by the same finite state machine, which typically is deterministic.

Now, consider the exceedingly trivial cellular automaton that I call *All Life*, which is completely governed by the following very simple local causal rule: A cell is alive at a given time whether or not it or any of its neighbors were alive or dead at the previous moment. It is trivial to give a short-cut explanation of any macro-property of All Life arbitrarily far into the future, because one can see that all cells will be alive at all future times, no matter what the initial aggregate local configuration of the cellular automaton.

All Life is atypical compared with the cellular automata usually discussed, such as the Game of Life (GOL). The most interesting CAs are those like GOL that are

[8] The explanation can apply to indeterministic systems by including complete information about the system's inderministic micro-state changes in the information from which the explanation is sought. System boundary conditions are handled in a similar fashion. See Bedau (1997). My notion of incompressible explanation is closely connected with Chaitin's notion of random sequence (Chaitin 1975, 1988) and Wolfram's subsequent notion of incompressible computation (Wolfram 1985, 2002), as well as the notion of a dynamical system that must be simulated to discover its generic properties (Crutchfield et al. 1986).

complex and known to require incompressible explanation.[9] The behavior of complex cellular automata typically cannot be explained except by crawling the causal web. The same holds throughout many other kinds of computational systems, such as soft artificial life systems like Ray's Tierra, Holland's Echo, and Packard's Bugs.[10] The more simulations of natural complex adaptive systems you study, the more weak emergence you find. This is a contingent empirical claim, but it is still true.

One might ask exactly much micro-causal complexity is sufficient for weak emergence, and exactly how much is necessary. This question makes sense if there is a bright line separating weak emergent properties from merely resultant properties, but the truth is more complicated. Weak emergence comes in degrees. Assad and Packard (1992, p. 232) describe a scale for degrees of emergence, ranging over behavior that is "immediately deducible upon inspection of the specification or rules generating it", to behavior that is "deducible in hindsight from the specification after observing the behavior", and continuing to behavior that is "deducible in theory, but its elucidation is prohibitively difficult", and finally reaching behavior that is "impossible to deduce from the specification". Explanatory incompressibility can be arrayed into similar stages. So, since weak emergence depends on explanatory incompressibility, weak emergence also comes in stages or degrees.[11] The paradigm case of weak emergence involves properties with incompressible explanations. A lower degree of emergence involves properties with compressible explanations that are so complicated that in practice no one can use the explanation except with a computer simulation. A higher degree of emergence involves properties that can be simulated but not in any finite simulation. These examples illustrate some of the sorts of ways in which weak emergence comes in degrees.[12]

My indirect definition of weak emergence applies to all systems with macro-behavior with only incompressible explanations. Now, if a system's behavior requires incompressible explanations, that is presumably because of the complexity of the system's micro-causal interactions. If we could directly identify what it is about micro-causal interactions that make them incompressible, then we might be able to construct a direct definition of weak emergence. In the meantime, experience in the field with our indirect definition suggests two conclusions about the intrinsic properties of micro-causal dynamics that require incompressible explanations. First, complex systems with weak emergent behavior typically involve massively parallel

[9] The classic reference on the Game of Life is Berlekamp et al. (2004), and Wolfram (1985, 2002) are important references on cellular automata in general. For more on emergence and cellular automata, see Bedau (1997, 2003) and the references therein.

[10] E.g., Bedau (1997) reviews the supple adaptation to the edge of disorder that emerges from Packard's Bugs model.

[11] See, also, Bedau (2003, p. 163).

[12] In this paper I will not take a stand on how to measure amounts of weak emergence. The most precise and explicit formal definition of amounts of weak emergence known to me is due to Paul Hovda (2008). Hovda defines the amount of simulation effort needed to derive something. This formalism could be interpreted as the amount of effort required for something's generative explanation.

micro-level populations of independent and autonomous agents that interact with their neighbors and their local environment. Second, the interactions among the agents and their environments are typically non-linear and synergistic, so that the behavior of an agent is highly sensitive to its local context, including other agents. These two factors make the behavior of complex systems' impossible to predict, even given complete prior micro information, short of crawling the causal web. On the other hand, crawling the micro-causal web provides an explanation that is guaranteed to be perfect, at least in principle.

The possibility of completely explaining weak emergent phenomena by crawling the causal web entails that weak emergence is consistent with reductionism. Many philosophers and scientists assume that emergence and reduction are *incompatible*. One typical form of reductionism is mereological supervenience (Kim 1978), the view that wholes are completely determined, ontologically and causally, by their parts. And it is certainly true that some kinds of emergence are incompatible with reduction; for example, strong forms of emergence are often defined in terms of reductive failure (e.g., Kim 1999). However, weak emergence differs from strong emergence because it is *consistent* with many forms of reduction.[13] To see this, consider ontological, causal, and explanatory reductions, in turn (which, respectively, involve reducing ontologies, causal relations, and explanations). Each concrete physical embodiment of weak emergence is ontologically nothing more than some kind of aggregation of smaller embodied objects. For example, the ontological substance of a traffic jam is nothing more than a certain kind of aggregation of cars on a road, and the ontological substance of a vesicle is nothing more than a certain kind of closed spherical bilayer aggregation of amphiphilic molecules in water. Furthermore, the causes and effects of each concrete instance of any kind of weak emergent macro-phenomenon are reducible to the iteration of the aggregation of the causes and effects operating at the micro-level, at least in principle. So, each example of macro-level weak emergence is ontologically and causally reducible to micro-level phenomena. However, in practice, typically nobody can understand or follow such a micro-causal reduction unless they simulate the micro-causal web on a computer, because the micro-level causal web is so complex. In a wealth of interesting cases, studied in fields like soft artificial life, computer simulations make it possible to crawl the causal web.[14]

The distinction between explanations or reductions that hold only *in principle*, versus those that also hold *in practice*, deserves further discussion. A reductive generative explanation of macro from micro might exist, in principle, but be unhelpful for explaining weak emergent phenomena, in practice, for a variety of reasons. One is that some relevant micro-level details required for the explanation might be unknown and inaccessible. Furthermore, even if all those details were known, the explanation might still be too complex and tedious for anyone to work through without the aid of something like a computer simulation. Nevertheless,

[13] A point often emphasized by Wimsatt (1986, 1997, 2000).

[14] It should be noted that in many cases we still do not know how to explain some natural regularities or patterns we seem to see in nature. One good example is the arrow of complexity in the evolution of life on Earth (Bedau, forthcoming).

given enough time and patience, anyone could work through all those micro-level explanatory details, as least in principle.[15] And working through those details is exactly what a computer simulation does. So, weak emergent phenomena always have complete and accurate explanations solely from micro-phenomena, at least in principle. These explanations rely on complete prior micro-level information, and they necessarily proceed by crawling the causal web. It is easy to see why it is typically impossible for anyone to grasp or understand how the emergent phenomena unfold from the micro, in practice, without resorting to computer simulations. This leads to the special connection between weak emergence and computer simulations, discussed below.

The distinction between explanation and reduction *in principle* and *in practice* helps explain how weak emergence fits the two hallmarks of emergence we mentioned earlier: the dependence of the macro on the micro, and the autonomy of macro from the micro. In cases of weak emergence, the macro depends on the micro because, in principle, each instance of the macro ontologically and causally is nothing more than the aggregation of micro-causal elements. For example, the ontological and causal state of a cellular automaton macro structure is nothing more than the aggregation of the ontological and causal states of its micro constituents. At the same time, weak emergence exhibits a kind of macro autonomy because of the incompressibility of the micro-causal generative explanation of the macro structure. Because the explanation is incompressible, it is useless in practice (except in so far as it serves as the basis for a good simulation of the system).

The subtle way in which weak emergence balances principles and practices is summarized with the awkward but apt notion of *in principle irreducibility in practice*. Although weak emergence phenomena have a true, complete, and exact micro-level generative explanations, at least in principle, incompressibility makes the explanations of little use, in practice. In practice, we have no alternative but to simulate the systems micro-level behavior, if we want to observe what macro behavior will emerge. This is a practical limitation, a limitation on irreducibility in practice. Furthermore, this practical limitation holds *in principle* for any naturalistic epistemic agent that is trying to explain the behavior of complex systems. We can put these points together by saying that weak emergent phenomena are in principle irreducible in practice.

As an aside, we should note that in many contexts an especially important subspecies of weak emergence is *robust* weak emergence. Weak emergence is robust when it involves causally salient law-like patterns involving weak emergent macro-properties (Bedau 2003). These robust emergent patterns recur in regular statistical patterns. Being typical or generic, they have some explanatory force. Many properties of the emergent patterns are insensitive to the details of the local micro-interactions that produce the patterns, so the emergent patterns have multiple realizations. One interesting special case of robust weak emergence are the physical systems that exhibit what physicists call “universal” behavior, especially around

[15] The computer-generated proof of the four colors theorem is one specific kind of example of a proof that one could work through in principle but not in practice.

phase transitions, such as when a solid melts into a liquid.[16] Physicists in some instances have mathematically proved that the critical behavior of some large class of physical systems is insensitive to almost all details about the system, but in most cases one has merely empirical evidence that a physical system exhibits universal behavior. Nevertheless, this empirical evidence can be very strong.[17] Since the evidence is empirical, sometimes we are wrong when we think we have strong evidence that a system's behavior is weakly emergent. But this is not a weakness in the notion of weak emergence. It is the expected consequence of all empirically justified claims.

Why Weak Emergence is Not Just in the Mind

There are a number of reasons why weak emergence might seem to be just in the mind, and I will discuss some of the most important and influential ones. I will argue that weak emergence goes beyond our minds and concerns actual objective causal relations in nature. In that sense, weak emergence is not just in the mind.

The first worry is especially simple to explain: The *existence* of in principle irreducible downward causation is an ontological matter, because it involves the real existence of a certain kind of causal process. Many people, especially in the philosophy of mind, are interested in a strong form of emergence that entails the existence of this kind of in principle irreducible downward causation. For example, Silberstein and McGeeve (1999) define "ontological emergence" as "features of systems of wholes that possess causal capacities not reducible to any of the intrinsic causal capacities of the parts nor to any of the (reducible) relations between the parts" (p. 182), and continue: "Emergent properties are properties of a system taken as a whole which exert a causal influence on the parts of the system consistent with, but distinct from, the causal capacities of the parts themselves" (p. 182). This statement illustrates how emergence sometimes is equated only with strong emergence, specifically, the sort that involves in principle irreducible downward causation.

Now, weak differs from strong emergence on precisely this point, for weak emergence *bars* in principle irreducible downward causation. Weak emergence does involve a certain kind of downward causation, and that kind of downward causation is irreducible in practice, due to explanatory incompressibility. But weakly emergent phenomena can always be given an explanation by crawling the causal web.[18] The web could always be crawled in principle, given complete information about the initial conditions and boundary conditions, and given enough time and effort. This is what computer simulations do. This difference between reducibility *in principle* and *in practice* is the difference between strong and weak emergence.

By defining their terms appropriately, Silberstein and McGeeve (1999) brand weak emergence as merely "epistemological" and not genuinely "ontological".

[16] See, e.g., Laughlin and Pines (2000), Laughlin (2006). Batterman (2002) is the first philosopher to emphasize the connection between emergence and universality in physics.

[17] See, e.g., Stanley (1971).

[18] Bedau (2003) elaborates this claim.

They define epistemological emergence to apply to any property that is "reducible to or determined by the intrinsic properties of the ultimate constituents of the objects or system" but "is very difficult for us to explain, predict or derive… on the basis of the ultimate constituents" (p. 186). A reductive consequence for macro phenomena immediately follows: "*In principle*, in such cases the higher-level feature, rule or law is a logical consequence of some lower-level feature, rule or law" (186, emphasis added). Nevertheless, micro-reductive explanations fail in practice, and "at each stage, entirely new laws, concepts and generalizations will be necessary (*though not in principle*) to explain or predict the phenomena with relative ease" (186, emphasis added).

Note that all of these properties of "epistemological" emergence apply to weak emergence. I have stressed how weak emergent macro phenomena in principle can always be explained solely from complete prior micro phenomena, by crawling the causal web. But those explanations (or reductions) exist only in principle. In practice, the explanations are so incompressible that we can explain the emergent phenomena only if we resort to computer simulations or to appeal to empirically justified macro-level patterns, regularities, or laws. So, weak emergence meets the definition of "epistemological" emergence from Silberstein and McGeever.

But this definition does not provide a reason to conclude that weak emergence is merely in the mind. It is just a definition; it has no force as an argument. Silberstein and McGeever *define* "epistemological" emergence as any form of emergence that rejects in principle irreducible downward causation, and according to that definition weak emergence is "epistemological." But it does not follow that weak emergence is *merely* in the mind or *merely* epistemological. Weak emergent phenomena might also have ontological, non-epistemological aspects. In particular, the distinctively incompressible micro-causal explanations of weak emergence presumably are due to a distinctively incompressible form of micro-causal structure in reality. It is presumably not an accident that one sort of micro-causal structure is incompressible and another sort is compressible. So, though weak emergence meets the Silberstein and McGeever definition of "epistemological" emergence, weak emergence is not *merely* epistemological. It is not *just* in the mind. Instead, weak emergence results from incompressible macro-level structure in the network of micro-level causal connections. This causal web is embodied and brought to life in real ontological substances with real causal powers, and it really generates certain macro-level ontological and causal phenomena.

Let us now turn to a second reason for thinking that emergence is merely in the mind, due to Hempel (Hempel and Oppenheim 1965). Hempel construed emergence as irreducibility of the macro from the micro given the full resources of the best scientific theories of the day. He presumed that all apparently emergent phenomena are merely apparent, and have a true, reductive and non-emergent explanation. This implies that attributions of emergence are merely admissions of our ignorance of the true, reductive and non-emergent explanation. If our best scientific theories construe certain phenomena as emergent (because in principle they are irreducible), that does not show us anything about nature. Rather, it just shows that we need a better scientific theory. You can sum up Hempel's complaint against emergence this way:

If our best scientific theories imply the existence of weak emergent phenomena, that merely reflects our ignorance of their true, non-emergent explanation.

Cellular automata provide an especially clear and simple illustration why Hempel's complaint does not apply to weak emergence. Consider the Game of Life (GOL).[19] We know the complete micro-theory for the behavior of each cell in the GOL: A cell at a given moment is alive just in case it was alive at the previous moment and had two or three living neighbors, or it was dead at the previous moment and had exactly three living neighbors. There is no ignorance whatsoever in our understanding of the rules that completely determine the micro-behavior of any call in the GOL at any time, because the behavior of every cell in the GOL is a trivial application of the GOL birth-death rule. Thus, any weak emergence that exists in the GOL is not merely the result of our ignorance of true non-emergent explanations; it is not just something we have imagined. If it exists, it is real and objective. The explanatory incompressibility in the GOL arises from the context-sensitivity of the birth-death rule; living cells arise or persist only if their immediate neighborhood contains just the right level of living cells, and neither too few nor too many. The GOL is also synergistic; the effect of the state of a given cell depends on the states of neighboring cells, and this is symmetric. In addition, the rule enables the existence of emergent macro-level causal structures, such as so-called "gliders" that propagate and interact in a family of reactions. A universal Turing machine can even be constructed in the GOL (Berlekamp et al. 2004). Gliders enable CAs to be programmed to perform various desired forms of complex parallel computations, such as density classification and synchronization (Crutchfield and Mitchell 1995; Crutchfield et al. 2003). In general, these interesting macro-behaviors of CAs cannot be explained even given the complete and accurate theory governing micro-state behavior (along with contingent information about initial micro-states and boundary conditions), except, of course, by crawling the causal web. This shows that the behavior of those CAs is weakly emergent. In the GOL we know that this kind of weak emergence is not a sign of our ignorance of some true, underlying, non-emergent explanation of the macro-behaviors. Emergent macro-behaviors have no non-emergent, compressible explanation, but they can always be explained by crawling the causal web. So, contemporary echoes of Hempel's criticism provide no reason to think that weak emergence is merely in the mind.

There is a third reason why weak emergence might seem to be merely in the mind: Weak emergence is anthropocentric and concerns a limitation in human epistemological capacities. In fact, weak emergence is defined in this paper in terms of its incompressible explanations, and elsewhere in terms of its underivability without simulation. Both definitions are indirect and identify emergent phenomena in terms of our distinctive epistemological relationship with them (the incompressibility of their explanation, and their underivability except by simulation). Since weak emergence is defined by reference to these human epistemological limitations, doesn't it follow that weak emergence is just in our minds?[20]

[19] For details about the GOL, see Bedau (1997, 2003) and the references cited therein.

[20] I might mention that the indirectness itself of these definitions of weak emergence does not make the definitions dubious. Indirect definitions can still be perfectly useful and accurate.

In truth, though, weak emergence has nothing to do with specifically human epistemological limitations. Weak emergence does not involve phenomena that are too complex for humans to explain but simple enough for smarter naturalistic epistemic agents to explain.[21] Rather, when weak emergence arises, the actual underlying local micro-causal processes are so complex that, in principle, complete and accurate explanations of macro-behavior are all incompressible. The emergent phenomena that arise from complex synergistic micro-causal explanations are explanatorily incompressible for any naturalistic epistemic agent. No matter how fast and infallibly inferences are made, no matter how perfect memories remain, naturalistic epistemic agents trying to explain weak emergent phenomena can produce only incompressible explanations. The practical limitations of explanations of weak emergence apply in principle to any epistemic agent; this is in principle inexplicability in practice.

If something has an indirect epistemological definition, it does not follow that it is just in the mind. Instead, the indirect epistemological definition is produced by and reflects a distinctive underlying ontological status or structure in nature. Incompressibility of explanations is a consequence of the objective complexity of the local micro-causal interactions that are ultimately generating the emergent behavior being explained. The micro-causal web is real and objective, and the incompressible causal pathways of weak emergent phenomena have a distinctive epistemological consequence. Note that the explanatory incompressibility that defines weak emergence applies to the explanations of any naturalistic epistemic agent, in principle. Just like us, any non-human epistemic agent will have to work through the objective complexity of the local micro-causal interactions. Thus, weak emergence is not merely in the mind, but refers to objective complexity in the objective natural world that is in principle irreducible in practice.

We can now turn to two further questions about weak emergence: in what way weak emergence applies only or mainly to computer simulations, and in what way weak emergence is inherently dynamic rather than static. It turns out that answering these questions has a connection with why weak emergence is not just in the mind.

Computer Simulations and Weak Emergence

Weak emergence has an especially close connection with computer simulations and computational systems, but the link is sometimes misunderstood. Some of the best examples of weak emergence come from computer simulations, and my previous definitions of weak emergence rely centrally on the notion of "underivability except by simulation" (Bedau 1997, 2003).[22] Someone might incautiously infer that weak

[21] By "naturalistic epistemic agent" I mean one with no magical abilities, such as an infinite amount of storage space.

[22] See also Humphreys's (2007a) discussion of computational emergence.

emergence applies merely or primarily to computer simulations or other kinds of computational systems. But that would be a mistake.

In truth, explanatory incompressibility typifies the behavior of a great many natural systems. For example, macro-level traffic jams are composed of a loose and changing group of individual micro-level automobiles, and traffic jams exhibit interesting macro-behavior. For example, jams suddenly spontaneously form when the traffic density crosses a critical value, and jams move slowly backwards in the traffic flow. It is easy to explain these macro-behaviors by iterating and aggregating all the simple local interactions among individual vehicles (Nagel and Rasmussen 1996; Sugiyama et al. 2008), but as far as anyone knows it is impossible to give any short-cut explanation of this behavior from complete information about micro-states and boundary conditions.

The same holds for the behavior of many natural chemical processes, such as the self-assembly, growth, and subsequent division of vesicles formed from amphiphiles in appropriate aqueous solutions (Hanczyc et al. 2003). Likewise, explanatory incompressibility seems to characterize a vast number of global properties of complex systems in molecular and cellular biology, including regulatory gene networks, metabolic networks, and the process by which proteins fold into three-dimensional structures. The same can be said for many systems studied by psychology and the social sciences. So, according to our best current explanations of complex systems, weak emergence applies throughout nature. It is not limited only or mainly to computational systems and computer simulations.

At the same time, one must acknowledge that computer simulations and computational systems have two important roles in helping us to understand weak emergence. First, certain computational systems produce some of the most striking *examples* of weak emergent phenomena. Computer simulations of complex, non-linear, dynamical, hierarchical systems in nature comprise one class of computational embodiment of weak emergence. Another class consists of complex computational systems that are not simulations of something else but are studied in their own right. Both kinds of computational systems have a massively parallel architecture with non-linear local interactions. Cellular automata and other software systems studied in artificial life provide plenty of good examples. Those familiar with these computational systems know that the global patterns they produce comprise many interesting and vivid examples of weak emergent phenomena.

Note that these patterns and regularities produced by computational systems are not mere simulations of emergent phenomena. Rather, they are computational embodiments of real emergent phenomena. That is, the computer produces something that is weakly emergent in its own right. If the computer happens to be simulating some natural system, that natural system might also exhibit its own emergent phenomena. Further, if the simulation is accurate in the relevant respects, it might explain why and how the natural system's phenomena is weakly emergent. But the computer simulation itself, considered as an object it its own right, is also exhibiting emergent behavior.

This points to a distinctive role that computer simulations play in our *evidence for* weak emergence in complex natural systems.[23] We typically study the behavior of complex systems by computer simulations, because we typically have no practical alternative. It is no accident that computer simulations fill the study of complex natural systems in virtually all disciplines. Computer simulations provide our only useful evidence about how complex systems will behave, about what global patterns emerge from their myriad micro-interactions.[24] That is why it is possible to define weak emergence as that which is "underivable except by simulation" (Bedau 1997, 2003).

The evidence for weak emergence provided by computer simulations, like other empirical evidence, can be misleading. So our beliefs about which systems are weakly emergent can be mistaken, when empirical evidence leads us astray. If we discover that there is a compressible explanation for some complex behavior that we thought was weakly emergent, this is not a flaw with the notion of weak emergence. It shows merely that we were wrong about an example of weak emergence. The possibility of this kind of error is an expected consequence of the empirical nature of the evidence for emergence in simulations.

The indirect epistemological role of computer simulations in explaining weak emergence might fuel a revival of the belief that weak emergence is in some sense merely epistemological. But this would be a mistake. The weak emergence exhibited in jamming traffic and dividing vesicles is not merely epistemological. Traffic jams and vesicles require incompressible explanations because of their objective, intrinsic micro-causal complexity. Traffic jams and vesicles are not just in the mind.

[23] Different kinds of computational systems have been called "simulations" so I should clarify what I mean. The simulations I have in mind are those that crawl the causal web (recall above) and generate global properties out of myriad local interactions. In addition to cellular automata, so-called "agent-based" models are good examples of simulations that crawl the causal web. They explicitly describe how local causal processes unfold over time, and global properties are merely certain kinds of aggregations of local properties.

[24] A tangential issue arises here: How can we tell if a computer model corresponds to reality, especially if the model is much simpler than the natural system being studied? This complex issue is beyond the scope of this footnote, but I would like to mention one point—that some computer simulations aim to explain only certain very general and robust global patterns and regularities in the behavior of certain complex systems. They do not attempt to explain the system's detailed behavior. Further, sometimes a complex system's robust global patterns and regularities are due to relatively simple and abstract features of the system; many of the details about the system do not materially affect its robust global behavior. In these cases, a very simple and abstract model can adequately explain the system's robust behavior.

One example might be Shelling's famous simple models of social phenomena such as segregated neighborhoods (Schelling 1968). These models abstract away from almost all the details of actual social neighborhoods. But they preserve certain key property—such as each agent's awareness of the social class of its local neighbours and itself, and its preference for local neighbors of the same social class—and the models explain how global segregation can result merely from those simple facts. Furthermore, you can empirically test whether people actually do know the social classes of their immediate neighborhoods, so even simple models can be empirically grounded.

Weak Emergence and the Dynamics of Causal Processes

Weak emergence has a distinctive *dynamical* nature.[25] Most recent philosophical discussions of emergence concern the static, synchronic relationships between different kinds of instantaneous phenomena or states (e.g., McLaughlin 1997, Kim 1999). The canonical example concerns someone's mental state (or some aspect of it) at some moment is thought to emerge from the person's brain state *at the same moment*. That kind of emergence is static, because it concerns states and conditions that are synchronic, i.e., that all happen at the same moment. By contrast, weak emergence concerns the processes by which certain global phenomena are generated from an aggregation of local phenomena. These generative processes are essentially dynamic. They happen over time, and are caused by local interactions. So, something is weakly emergent at a given moment not merely because of the aggregation of its local states at that time, but rather because of how its global states arose from previous aggregations of previous local states. That is, weak emergence is a historical property; it requires having a certain kind of causal history.

We noted above that the micro-level causal processes that underlie weak emergence form a large, context sensitive, massively parallel network of local causal interactions. Many of those interactions are nonlinear. In general, the only way to explain the system's eventual global (macro) behavior from the sum of its local (micro) behavior is to crawl the causal web. Starting with a completely specified initial micro-condition, one propagates the local micro influences forward in time, and then aggregates the local micro results into the global state of the system. You can think of this as "deriving" the system's later global states from its earlier micro states. These complex causal networks require a distinctive incompressible kind of explanation. The local causal networks really involve a large number of components that interact locally and nonlinearly, sometimes in a heterogeneous variety of ways. The net effect is that the system's global behavior is unpredictable, except by crawling the causal web.

Complex micro-causal webs exhibit different kinds of dynamical behavior. Sometimes the causal dynamics produces a global equilibrium state that is essentially constant and fixed over time. Sometimes the causal dynamics produce a chaotically changing sequence of global states, which remain far from equilibrium and behave unpredictably. Sometimes weak emergent global dynamics are robust, indicating a generic statistical regularity in the dynamical emergent behavior of a class of complex systems.

The momentary, static global state of a system exhibiting weak emergence is trivially derivable from the system's local state at that same moment. For the global static state is nothing more than simply the sum or aggregation of all the local states at that time. Return to our earlier examples of weak emergence: the formation and behavior of traffic jams, and the spontaneous self-assembly, growth and division of vesicles. Note that both are dynamic causal processes that unfold in time. The causal processes that generate traffic jams and cause vesicles to grow and divide are not

[25] Rueger (2000b) and Humphreys (2007a, 2007b) stand out in contemporary philosophical discussions for their focus on dynamic forms of emergence.

just in the mind. They really exist and really involve a complex network of causal relations among micro-level entities. Nature is full of this sort of complex causal systems, especially those parts of nature that are alive, or have a mind, or involve social relations and technology. In this way, the dynamic nature of weak emergence underscores why weak emergence is not merely in the mind.

Conclusions

This paper has defended a form of weak emergence that is based on the notion of explanatory incompressibility. This weak emergence is clearly metaphysically innocent and consistent with any reasonable form of naturalism, for it rests entirely on merely the dynamical micro-causal processes that underlie and generate complex phenomena. These dynamical causal processes occur not just in computer simulations or computer systems; they also occur in a vast number of the complex systems in the natural world.

So, weak emergence is not just in our minds. It concerns not just how we explain things. It is produced not merely by human mental, explanatory, or epistemic limitations. It does not mask ignorance of true non-emergent explanations. Rather, weak emergence is an objective phenomenon that exists in nature. Any naturalistic epistemic agent who tries to explain it will have to use incompressible explanations. Weak emergence is the macro-level mark of incompressible complexity in a network of micro-causal interactions. When the objective micro-causal web is sufficiently complex, all explanations of its macro-behavior are incompressible. The resulting weak emergence is not just in the mind

Acknowledgement I am grateful for helpful comments from Kellyn Bardeen, Paul Humphreys, and Philippe Huneman.

References

Assad, A., & Packard, N. H. (1992). Emergence. In M. A. Bedau & P. Humphreys (Eds.). (2008). *Emergence: Contemporary readings in philosophy and science* (pp. 231–234). Cambridge: MIT Press; page citations refer to this printing. Originally appeared as Sect. 2 of "Emergent colonization in an artificial ecology". In F. Varela & P. Bourgine (Eds.), *Towards a practice of autonomous systems: Proceedings of the First European Conference on Artificial Life* (pp. 143–152). Cambridge, MA: The MIT Press.

Batterman, R. (2002). *The devil in the details*. New York: Oxford University Press.

Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives* 11, 375–399.

Bedau, M. A. (2003). Downward causation and autonomy in weak emergence. *Principia Revista Internacional de Epistemologica, 6*, 5–50. (Reprinted in M. A. Bedau & P. Humphreys (Eds.). (2008). *Emergence: Contemporary readings in philosophy and science*. Cambridge: MIT Press, page citations refer to this printing).

Bedau, M. A. (Forthcoming). The evolution of complexity. In T. Pradeu, et al. (Eds.), *Mapping the future of biology: Evolving concepts and theories*. Heidelberg: Springer.

Bedau, M. A., & Humphreys, P. (Eds.). (2008). *Emergence: Contemporary readings in philosophy and science*. Cambridge: MIT Press.

Berlekamp, E. R., Conway, J. H., & Guy, R. K. (2004). What is life? Chapter 25 in *Winning ways for your mathematical plays* (2nd ed., Vol. 4, pp. 927–961). Wellseley, MA: AK Peters.

Boogerd, F. C., Bruggeman, F. J., Richardson, R. C., Stephan, A., & Westerhoff, H. V. (2005). Emergence and its place in nature: A case study of biochemical networks. *Synthese, 145*, 131–164.
Chaitin, G. J. (1975, May). Randomness and mathematical proof. *Scientific American, 232*, 47–53.
Chaitin, G. J. (1988, July). Randomness in arithmetic. *Scientific American, 259*, 80–85.
Crutchfield, J. P., Farmer, J. D., Packard, N. H., & Shaw, R. S. (1986). Chaos. *Scientific American, 255*, 46–57.
Crutchfield, J. P., & Mitchell, M. (1995). The evolution of emergent computation. *Proceedings of the National Academy of Sciences USA, 92*, 10742–10746.
Crutchfield, J. P., Mitchell, M., & Das, R. (2003). Evolutionary design of collective computation in cellular automata. In J. P. Crutchfield, & P. K. Schuster (Eds.), *Evolutionary dynamics—exploring the interplay of selection, neutrality, accident, and function* (pp. 361–411). New York: Oxford University Press.
Feltz, B., Crommelinck, M., & Goujon, P. (Eds.). (2006). *Self-organization and emergence in life sciences*. Heidelberg: Springer.
Hanczyc, M. M., Fujikawa, S. M., & Szostak, J. W. (2003). Experimental models of primitive cellular components: Encapsulation, growth, and division. *Science, 302*, 618–622.
Hempel, C. (1965). Aspects of scientific explanation. In C. Hempel (Ed.), *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 331–496). New York: Free Press.
Hempel, C., & Oppenheim, P. (1965). On the idea of emergence. In C. Hempel (Ed.), *Aspects of scientific explanation and other essays in the philosophy of science* (pp. 258–264). New York: Free Press.
Hovda, P. (2008). Quantifying weak emergence. *Minds and Machines* (this issue). doi:10.1007/s11023-008-9123-5
Humphreys, P. (2007a). Computational and conceptual emergence (preprint).
Humphreys, P. (2007b). Pattern emergence (preprint).
Kim, J. (1978). Supervenience and nomological incommensurables. *American Philosophical Quarterly, 15*, 149–156.
Kim, J. (1999). Making sense of emergence. *Philosophical Studies, 95*, 3–36.
Kistler, M. (Ed.). (2006). New perspectives on reduction and emergence in physics, biology, and psychology. *Synthese (Special Issue), 151*(3).
Laughlin, R. (2006). *A different universe: Reinventing physics from the bottom down*. New York: Basic Books.
Laughlin, R., & Pines, D. (2000). The theory of everything. *Proceedings of the National Academy of Science USA, 97*, 28–31.
McIntyre, L. (1998). Complexity: A philosopher's reflections. *Complexity, 3*, 26–32.
McLaughlin, B. P. (1992). The rise and fall of British Emergentism. In A. Beckerman, H. Flohr, & J. Kim (Eds.), *Emergence or reduction? Essays on the prospects of nonreductive physicalism* (pp. 49–93). Berlin: Walter de Gruyter.
McLaughlin, B. P. (1997). Emergence and supervenience. *Intellectica, 25*, 25–43.
Nagel, K., & Rasmussen, S. (1996). Particle hopping models and traffic flow theory. *Physical Review E, 53*, 4655–4672.
Newmann, D. V. (1996). Emergence and strange attractors. *Philosophy of Science, 63*, 245–261.
Rueger, A. (2000a). Robust supervenience and emergence. *Philosophy of Science, 67*, 466–489.
Rueger, A. (2000b). Physical emergence, diachronic and synchronic. *Synthese, 124*, 297–322.
Schelling, T. (1968). *Micromotives and macrobehavior*. New York: Norton.
Silberstein, M., & McGeeve, J. (1999). The search for ontological emergence. *Philosophical Quarterly, 49*, 182–200.
Simon, H. (1996). Alternative views of complexity. In his *The sciences of the artificial* (3rd ed.). Cambridge: MIT Press.
Stanley, H. E. (1971). *Introduction to phase transitions and critical phenomena*. New York: Oxford University Press.
Stephan, A. (2006). The dual role of 'emergence' in the philosophy of mind and in cognitive science. *Synthese, 151*, 485–498.
Sugiyama, Y., Fukui, M., Kikuchi, M., Hasebe, K., Nakayama, A., Tadaki, S., & Nakayama, A. (2008). Traffic jams without bottlenecks—experimental evidence for the physical mechanisms of the formation of a jam. *New Journal of Physics, 10*, 033001. doi:10.1088/1367-2630/10/3/033001
Wimsatt, W. C. (1986). Forms of aggregativity. In A. Donagan, A. N. Perovich Jr., & M. V. Wedin (Eds.), *Human nature and natural knowledge* (pp. 259–291). Dordrecht: Reidel.

Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science, 64*, S372–S384.
Wimsatt, W. C. (2000). Emergence as nonaggregativity and the biases of reductionisms. *Foundations of Science, 5*, 269–297.
Wolfram, S. (1985). Undecidability and intractability in theoretical physics. *Physical Review Letters, 54*, 735–738.
Wolfram, S. (2002). *A new kind of science*. Champaign, IL: Wolfram Media.

Susan Glaspell

The Verge

Prefatory Remarks to the Excerpt [JTR]:

The reading I give you is the excerpt of a drama that is at the very heart of American expressionist theatre. It was first performed at the Provincetown Playhouse on November 14, 1921. The Playhouse, founded by Susan Glaspell and her husband George Cram Cook, was crucial in establishing this new theatre aesthetics. To give you the gist of the play: The central motif of the three-act drama is the botanist Claire Archer's struggle with the social conventions of her time, often shown through the repeated emphasis put on patterns and acts of pattern-breaking on the symbolic level. Thus, the first act is set in Claire's experimental greenhouse where, according to the set description, "[t]he frost has made patterns on the glass as if—as Plato would have it—the patterns inherent in abstract nature and behind all life had to come out, not only in the creative heat within, but in the creative cold on the other side of the glass. And the wind makes patterns of sound around the glass house." Against the backdrop of the patterned set design, the first act shows how Claire's seemingly secluded life as a horticultural scientist is invaded by domestic life. Deprived of heating at their living quarters during a stormy winter morning, her three male companions, her friend Tom, her lover Dick, and her husband Harry as well as her daughter from a previous marriage, Elizabeth, seek their way to the greenhouse to have breakfast, for it is there where all the heat has been ordered by Claire. During a tumultuous breakfast, Claire successively rejects the intruders' claims in regard to her roles as a women, lover, wife and mother with references to her horticultural work. In the second act, set in her study in the family mansion that is also structured by a "a marvellous pattern on the curved wall", Claire is presented again defending herself against similar demands, this time, among others, by her sister Adelaide who wants her to appreciate the satisfaction of being part of "the main body" and "having one's roots in the big common experiences". Adelaide's intervention, however, fails, so that Claire is introduced to a neurologist who then affirms the suspicion of her insanity that has already been brought up by both her sister and husband. The final act returns to the greenhouse to show the success of Claire's plant experiments. When the Breath of Life, Claire's mutated plant, ultimately proves its stability as a new species, Claire, instead of celebrating her success, kills Tom, the one male to whom she has been able to confide her unconventional ideas. She does so after he had pledged his love to her, and this is precisely what Claire rejects as a last captivating emotion that would tie her to the social sphere. The drama ends then with Claire singing the Puritan hymn "Nearer My God to Thee" in the presence of the remaining, flabbergasted men.
Just in case you like a complete reading, the full play can be conveniently accessed at http://www.gutenberg.org/etext/10623.]

From **Act One**:

[…]

CLAIRE: What are you doing here?

HARRY: Getting breakfast. (*all the while doing so*)

CLAIRE: I'll not have you in my place!

HARRY: If you take all the heat then you have to take me.

CLAIRE: I'll show you how I have to take you. (*with her hands begins scooping upon him the soil* ANTHONY *has prepared*)

HARRY: (*jumping up, laughing, pinning down her arms, putting his arms around her*) Claire—be decent. What harm do I do here?

CLAIRE: You pull down the temperature.

HARRY: Not after I'm in.

CLAIRE: And you told Tom and Dick to come and make it uneven.

HARRY: Tom and Dick are our guests. We can't eat where it's warm and leave them to eat where it's cold.

CLAIRE: I don't see why not.

HARRY: You only see what you want to see.

CLAIRE: That's not true. I wish it were. No; no, I don't either. (*she is disturbed—that troubled thing which rises from within, from deep, and takes* CLAIRE. *She turns to the Edge Vine, examines. Regretfully to* ANTHONY, *who has come in with a plant*) It's turning back, isn't it?

ANTHONY [Claire’s assistant – JTR]: Can you be sure yet, Miss Claire?

CLAIRE: Oh yes—it's had its chance. It doesn't want to be—what hasn't been.

[…]

ANTHONY: But you're doing it this time, Miss Claire. When Breath of Life opens—and we see its heart—

(CLAIRE *looks toward the inner room. Because of intervening plants they do not see what is seen from the front—a plant like caught motion, and of a greater transparency than plants have had. Its leaves, like waves that curl, close around a heart that is not seen. This plant stands by itself in what, because of the arrangement of things about it, is a hidden place. But nothing is between it and the light*.)

CLAIRE: Yes, if the heart has (*a little laugh*) held its own, then Breath of Life is alive in its otherness. But Edge Vine is running back to what it broke out of.

HARRY: Come, have some coffee, Claire.

(ANTHONY *returns to the inner room, the outer door opens*. DICK *is hurled in*.)

CLAIRE: (*going to the door, as he gasps for breath before closing it*) How dare you make my temperature uneven! (*she shuts the door and leans against it*)

DICK: Is that what I do?

(*A laugh, a look between them, which is held into significance*.)

HARRY: (*who is not facing them*) Where's the salt?

DICK: Oh, I fell down in the snow. I must have left the salt where I fell. I'll go back and look for it.

CLAIRE: And change the temperature? We don't need salt.

HARRY: You don't need salt, Claire. But we eat eggs.

CLAIRE: I must tell you I don't like the idea of any food being eaten here, where things have their own way to go. Please eat as little as possible, and as quickly.

HARRY: A hostess calculated to put one at one's ease.

CLAIRE: (*with no ill-nature*) I care nothing about your ease. Or about Dick's ease.

DICK: And no doubt that's what makes you so fascinating a hostess.

CLAIRE: Was I a fascinating hostess last night, Dick? (*softly sings*) 'Oh, night of love—' (*from the Barcorole of 'Tales of Hoffman'*)

HARRY: We've got to have salt.

(*He starts for the door.* CLAIRE *slips in ahead of him, locks it, takes the key. He marches off, right*.)

CLAIRE: (*calling after him*) That end's always locked.

DICK: Claire darling, I wish you wouldn't say those startling things. You do get away with it, but I confess it gives me a shock—and really, it's unwise.

CLAIRE: Haven't you learned that the best place to hide is in the truth? (*as* HARRY *returns*) Why won't you believe me, Harry, when I tell you the truth—about doors being locked?

HARRY: Claire, it's selfish of you to keep us from eating salt just because you don't eat salt.

CLAIRE: (*with one of her swift changes*) Oh, Harry! Try your egg without salt. Please—please try it without salt! (*an intensity which seems all out of proportion to the subject*)

HARRY: An egg demands salt.

CLAIRE: 'An egg demands salt.' Do you know, Harry, why you are such an unseasoned person? 'An egg demands salt.'

HARRY: Well, it doesn't always get it.

CLAIRE: But your spirit gets no lift from the salt withheld.

HARRY: Not an inch of lift. (*going back to his breakfast*)

CLAIRE: And pleased—so pleased with itself, for getting no lift. Sure, it is just the right kind of spirit—because it gets no lift. (*more brightly*) But, Dick, you must have tried your egg without salt.

DICK: I'll try it now. (*he goes to the breakfast table*)

CLAIRE: You must have tried and tried things. Isn't that the way one leaves the normal and gets into the byways of perversion?

HARRY: Claire.

DICK: (*pushing back his egg*) If so, I prefer to wait for the salt.

HARRY: Claire, there is a *limit*.

CLAIRE: Precisely what I had in mind. To perversion too there is a limit. So—the fortifications are unassailable. If one ever does get out, I suppose it is—quite unexpectedly, and perhaps—a bit terribly.

HARRY: Get out where?

CLAIRE: (*with a bright smile*) Where you, darling, will never go.

HARRY: And from which you, darling, had better beat it.

CLAIRE: I wish I could. (*to herself*) No—no I don't either

(*Again this troubled thing turns her to the plant. She puts by themselves the two which* ANTHONY *covered with paper bags. Is about to remove these papers*. HARRY *strikes a match*.)

CLAIRE: (*turning sharply*) You can't smoke here. The plants are not used to it.

HARRY: Then I should think smoking would be just the thing for them.

CLAIRE: There is design.

HARRY: (*to* DICK) Am I supposed to be answered? I never can be quite sure at what moment I am answered.

(*They both watch* CLAIRE, *who has uncovered the plants and is looking intently into the flowers. From a drawer she takes some tools. Very carefully gives the rose pollen to an unfamiliar flower—rather wistfully unfamiliar, which stands above on a small shelf near the door of the inner room*.)

DICK: What is this you're doing, Claire?

CLAIRE: Pollenizing. Crossing for fragrance.

DICK: It's all rather mysterious, isn't it?

HARRY: And Claire doesn't make it any less so.

CLAIRE: Can I make life any less mysterious?

HARRY: If you know what you are doing, why can't you tell Dick?

DICK: Never mind. After all, why should I be told? (*he turns away*)

(*At that she wants to tell him. Helpless, as one who cannot get across a stream, starts uncertainly*.)

CLAIRE: I want to give fragrance to Breath of Life (*faces the room beyond the wall of glass*)—the flower I have created that is outside what flowers have been. What has gone out should bring fragrance from what it has left. But no definite fragrance, no limiting enclosing thing. I call the fragrance I am trying to create Reminiscence. (*her hand on the pot of the wistful little flower she has just given pollen*) Reminiscent of the rose, the violet, arbutus—but a new thing—itself. Breath of Life may be lonely out in what hasn't been. Perhaps some day I can give it reminiscence.

DICK: I see, Claire.

CLAIRE: I wonder if you do.

HARRY: Now, Claire, you're going to be gay to-day, aren't you? These are Tom's last couple of days with us.

CLAIRE: That doesn't make me especially gay.

HARRY: Well, you want him to remember you as yourself, don't you?

CLAIRE: I would like him to. Oh—I would like him to!

HARRY: Then be amusing. That's really you, isn't it, Dick?

DICK: Not quite all of her—I should say.

CLAIRE: (*gaily*) Careful, Dick. Aren't you indiscreet? Harry will be suspecting that I am your latest strumpet.

HARRY: Claire! What language you use! A person knowing you only by certain moments could never be made to believe you are a refined woman.

CLAIRE: True, isn't it, Dick?

HARRY: It would be a good deal of a lark to let them listen in at times—then tell them that here is the flower of New England!

CLAIRE: Well, if this is the flower of New England, then the half has never been told.

DICK: About New England?

CLAIRE: I thought I meant that. Perhaps I meant—about me.

HARRY: (*going on with his own entertainment*) Explain that this is what came of the men who made the laws that made New England, that here is the flower of those gentlemen of culture who—

DICK: Moulded the American mind!

CLAIRE: Oh! (*it is pain*)

HARRY: Now what's the matter?

CLAIRE: I want to get away from them!

HARRY: Rest easy, little one—you do.

CLAIRE: I'm not so sure—that I do. But it can be done! We need not be held in forms moulded for us. There is outness—and otherness.

HARRY: Now, Claire—I didn't mean to start anything serious.

CLAIRE: No; you never mean to do that. I want to break it up! I tell you, I want to break it up! If it were all in pieces, we'd be (*a little laugh*) shocked to aliveness (*to* DICK)—wouldn't we? There would be strange new comings together—mad new comings together, and we would know what it is to be born, and then we might know—that we are. Smash it. (*her hand is near an egg*) As you'd smash an egg. (*she pushes the egg over the edge of the table and leans over and looks, as over a precipice*)

HARRY: (*with a sigh*) Well, all you've smashed is the egg, and all that amounts to is that now Tom gets no egg. So that's that.

CLAIRE: (*with difficulty, drawing herself back from the fascination of the precipice*) You think I can't smash anything? You think life can't break up, and go outside what it was? Because you've gone dead in the form in which you found yourself, you think that's all there is to the whole adventure? And that is called sanity. And made a virtue—to lock one in. You never worked with things that grow! Things that take a sporting chance—go mad—that sanity mayn't lock them in—from life untouched—from life—that waits, (*she turns toward the inner room*) Breath of Life. (*she goes in there*)

[…]

CLAIRE: I am not doing any useful beautiful thing.

ELIZABETH: Oh, but you are, mother. Of course you are. Miss Lane [a teacher of Elizabeth – JTR] says so. She says it is your splendid heritage gives you this impulse to do a beautiful thing for the race. She says you are doing in your way what the great teachers and preachers behind you did in theirs.

CLAIRE: (*who is good for little more*) Well, all I can say is, Miss Lane is stung.

ELIZABETH: Mother! What a thing to say of Miss Lane. (*from this slipping into more of a little girl manner*) Oh, she gave me a spiel one day about living up to the men I come from.

(CLAIRE *turns and regards her daughter*.)

CLAIRE: You'll do it, Elizabeth.

ELIZABETH: Well, I don't know. Quite a job, I'll say. Of course, I'd have to do it in my way. I'm not going to teach or preach or be a stuffy person. But now that—(*she here becomes the product of a superior school*) values

have shifted and such sensitive new things have been liberated in the world—

CLAIRE: (*low*) Don't use those words.

ELIZABETH: Why—why not?

CLAIRE: Because you don't know what they mean.

ELIZABETH: Why, of course I know what they mean!

CLAIRE: (*turning away*) You're—stepping on the plants.

HARRY: (*hastily*) Your mother has been working awfully hard at all this.

ELIZABETH: Well, now that I'm here you'll let me help you, won't you, mother?

CLAIRE: (*trying for control*) You needn't—bother.

ELIZABETH: But I *want* to. Help add to the wealth of the world.

CLAIRE: Will you please get it out of your head that I am adding to the wealth of the world!

ELIZABETH: But, mother—of course you are. To produce a new and better kind of plant—

CLAIRE: They may be new. I don't give a damn whether they're better.

ELIZABETH: But—but what are they then?

CLAIRE: (*as if choked out of her*) They're different.

ELIZABETH: (*thinks a minute, then laughs triumphantly*) But what's the use of making them different if they aren't better?

HARRY: A good square question, Claire. Why don't you answer it?

CLAIRE: I don't have to answer it.

HARRY: Why not give the girl a fair show? You never have, you know. Since she's interested, why not tell her what it is you're doing?

CLAIRE: She is not interested.

ELIZABETH: But I am, mother. Indeed I am. I do want awfully to understand what you are doing, and help you.

CLAIRE: You can't help me, Elizabeth.

HARRY: Why not let her try?

CLAIRE: Why do you ask me to do that? This is my own thing. Why do you make me feel I should—(*goes to* ELIZABETH) I will be good to you, Elizabeth. We'll go around together. I haven't done it, but—you'll see. We'll do gay things. I'll have a lot of beaus around for you. Anything else. Not—this is—Not this.

ELIZABETH: As you like, mother, of course. I just would have been so glad to—to share the thing that interests you. (*hurt borne with good breeding and a smile*)

HARRY: Claire! (*which says, 'How can you?'*)

CLAIRE: (*who is looking at* ELIZABETH) Yes, I will try.

TOM: I don't think so. As Claire says—anything else.

ELIZABETH: Why, of course—I don't at all want to intrude.

HARRY: It'll do Claire good to take someone in. To get down to brass tacks and actually say what she's driving at.

CLAIRE: Oh—*Harry*. But yes—I will try. (*does try, but no words come. Laughs*) When you come to say it it's not—One would rather not nail it to a cross of words—(*laughs again*) with brass tacks.

HARRY: (*affectionately*) But I want to see you put things into words, Claire, and realize just where you are.

CLAIRE: (*oddly*) You think that's a—good idea?

ELIZABETH: (*in her manner of holding the world capably in her hands*) Now let's talk of something else. I hadn't the least idea of making mother feel badly.

CLAIRE: (*desperately*) No, we'll go on. Though I don't know—where we'll end. I can't answer for that. These plants—(*beginning flounderingly*) Perhaps they are less beautiful—less sound—than the plants from which they diverged. But they have found—otherness, (*laughs a little shrilly*) If you know—what I mean.

TOM: Claire—stop this! (*To* HARRY) This is wrong.

CLAIRE: (*excitedly*) No; I'm going on. They have been shocked out of what they were—into something they were not; they've broken from the forms in which they found themselves. They are alien. Outside. That's it, outside; if you—know what I mean.

ELIZABETH: (*not shocked from what she is*) But of course, the object of it all is to make them better plants. Otherwise, what would be the sense of doing it?

CLAIRE: (*not reached by* ELIZABETH) Out there—(*giving it with her hands*) lies all that's not been touched—lies life that waits. Back here—the old pattern, done again, again and again. So long done it doesn't even know itself for a pattern—in immensity. But this—has invaded. Crept a little way into—what wasn't. Strange lines in life unused. And when you make a pattern new you know a pattern's made with life. And then you know that anything may be—if only you know how to reach it. (*this has taken form, not easily, but with great struggle between feeling and words*)

HARRY: (*cordially*) Now I begin to get you, Claire. I never knew before why you called it the Edge Vine.

CLAIRE: I should destroy the Edge Vine. It isn't—over the edge. It's running, back to—'all the girls'. It's a little afraid of Miss Lane, (*looking sombrely at it*) You are out, but you are not alive.

ELIZABETH: Why, it looks all right, mother.

CLAIRE: Didn't carry life with it from the life it left. Dick—you know what I mean. At least you ought to. (*her ruthless way of not letting anyone's feelings stand in the way of truth*) Then destroy it for me! It's hard to do it—with the hands that made it.

DICK: But what's the point in destroying it, Claire?

CLAIRE: (*impatiently*) I've told you. It cannot create.

DICK: But you say you can go on producing it, and it's interesting in form.

CLAIRE: And you think I'll stop with that? Be shut in—with different life—that can't creep on? (*after trying to put destroying hands upon it*) It's hard to—get past what we've done. Our own dead things—block the way.

TOM: But you're doing it this next time, Claire, (*nodding to the inner room*.) In there!

CLAIRE: (*turning to that room*) I'm not sure.

TOM: But you told me Breath of Life has already produced itself. Doesn't that show it has brought life from the life it left?

CLAIRE: But timidly, rather—wistfully. A little homesick. If it is less sure this time, then it is going back to—Miss Lane. But if the pattern's clearer now, then it has made friends of life that waits. I'll know to-morrow.

ELIZABETH: You know, something tells me this is *wrong*.

CLAIRE: The hymn-singing ancestors are tuning up.

ELIZABETH: I don't know what you mean by that, mother but—

CLAIRE: But we will now sing, 'Nearer, my God, to Thee: Nearer to—'

ELIZABETH: (*laughingly breaking in*) Well, I don't care. Of course you can make fun at me, but something does tell me this is wrong. To do what—what—

DICK: What God did?

ELIZABETH: Well—yes. Unless you do it to make them better—to *do* it just to do it—that doesn't seem right to me.

CLAIRE: (*roughly*) 'Right to you!' And that's all you know of adventure—and of anguish. Do you know it is you—world of which you're so true a flower—makes me have to leave? You're there to hold the door shut! Because you're young and of a gayer world, you think I can't *see* them—those old men? Do you know why you're so sure of yourself? Because you can't *feel*. Can't feel—the limitless—out there—a sea just over the hill. I will not stay with you! (*buries her hands in the earth around the Edge Vine. But suddenly steps back from it as she had from* ELIZABETH) And I will not stay with *you! (grasps it as we grasp what we would kill, is trying to pull it up. They all step forward in horror. ANTHONY is drawn in by this harm to the plant*)

ANTHONY: Miss Claire! Miss Claire! The work of years!

CLAIRE: May only make a prison! (*struggling with* HARRY, *who is trying to stop her*) You think I too will die on the edge? (*she has thrown him away, is now struggling with the vine*) Why did I make you? To get past you! (*as she twists it*) Oh yes, I know you have thorns! The Edge Vine should have thorns, (*with a long tremendous pull for deep roots, she has it up. As she holds the torn roots*) Oh, I have loved you so! You took me where I hadn't been.

ELIZABETH: (*who has been looking on with a certain practical horror*) Well, I'd say it would be better not to go there!

CLAIRE: Now I know what you are for! (*flings her arm back to strike* ELIZABETH with the Edge Vine)

HARRY: (*wresting it from her*) Claire! Are you mad?

CLAIRE: No, I'm not mad. I'm—too sane! (*pointing to* ELIZABETH—*and the words come from mighty roots*) To think that object ever moved my belly and sucked my breast! (ELIZABETH *hides her face as if struck*)

[…]

From Act Two:

[…]

CLAIRE: I've known a few moments that were life. Why don't they help me now? One was in the air. I was up with Harry—flying—high. It was about four months before David [Claire’s son who had died in infancy – JTR] was born—the doctor was furious—pregnant women are supposed to keep to earth. We were going fast—I *was* flying—I had left the earth. And then—within me, movement, for the first time—stirred to life far in air—movement within. The man unborn, he too, would fly. And so—I always loved him. He was movement—and wonder. In his short life were many flights. I never told anyone about the last one. His little bed was by the window—he wasn't four years old. It was night, but him not asleep. He saw the morning star—you know—the morning star. Brighter—stranger—reminiscent—and a promise. He pointed—'Mother', he asked me, 'what is there—beyond the stars?' A baby, a sick baby—the morning star. Next night—the finger that pointed was—(*suddenly bites her own finger*) But, yes, I am glad. He would always have tried to move and too much would hold him. Wonder would die—and he'd laugh at soaring, (*looking down, sidewise*) Though I liked his voice. So I wish you'd stay near me—for I like your voice, too.

TOM: Claire! That's (*choked*) almost too much.

CLAIRE: (*one of her swift glances—canny, almost practical*) Well, I'm glad if it is. How can I make it more? (*but what she sees brings its own change*) I know what it is you're afraid of. It's because I have so much—yes, why shouldn't I say it?—passion. You feel that in me, don't you? You think it would swamp everything. But that isn't all there is to me.

TOM: Oh, I know it! My dearest—why, it's because I know it! You think I *am*—a fool?

CLAIRE: It's a thing that's—sometimes more than I am. And yet I—I am more than it is.

TOM: I know. I know about you.

CLAIRE: I don't know that you do. Perhaps if you really knew about me—you wouldn't go away.

TOM: You're making me suffer, Claire.

CLAIRE: I know I am. I want to. Why shouldn't you suffer? (*now seeing it more clearly than she has ever seen it*) You know what I think about you? You're afraid of suffering, and so you stop this side—in what you persuade yourself is suffering, (*waits, then sends it straight*) You know—how it is—with me and Dick? (*as she sees him suffer*) Oh, no, I don't want to hurt you! Let it be you! I'll teach you—you needn't scorn it. It's rather wonderful.

TOM: Stop that, Claire! That isn't you.

CLAIRE: Why are you so afraid—of letting me be low—if that is low? You see—(*cannily*) I believe in beauty. I have the faith that can be bad as well as good. And you know why I have the faith? Because sometimes—from my lowest moments—beauty has opened as the sea. From a cave I saw immensity.

My love, you're going away—
Let me tell you how it is with me;
I want to touch you—somehow touch you once before I die—
Let me tell you how it is with me.
I do not want to work,
I want to be;
Do not want to make a rose or make a poem—
Want to lie upon the earth and know. (*closes her eyes*)
Stop doing that!—words going into patterns;
They do it sometimes when I let come what's there.
Thoughts take pattern—then the pattern is the thing.
But let me tell you how it is with me. (*it flows again*)
All that I do or say—it is to what it comes from,
A drop lifted from the sea.
I want to lie upon the earth and know.
But—scratch a little dirt and make a flower;
Scratch a bit of brain—something like a poem. (*covering her face*)
Stop *doing* that. Help me stop doing that!

TOM: (*and from the place where she had carried him*)
Don't talk at all. Lie still and know—
And know that I am knowing.

CLAIRE:
Yes; but we are so weak we have to talk;
To talk—to touch.
Why can't I rest in knowing I would give my life to reach you?
That has—all there is.
But I must—put my timid hands upon you,
Do something about infinity.
Oh, let what will flow into us,
And fill us full—and leave us still.
Wring me dry,
And let me fill again with life more pure.
To know—to feel,
And do nothing with what I feel and know—
That's being good. That's nearer God.
(*drenched in the feeling that has flowed through her—but surprised helpless*) Why, I said your thing, didn't I? Opened my life to bring you to me, and what came—is what sends you away.

TOM: No! What came is what holds us together. What came is what saves us from ever going apart. (*brokenly*) My beautiful one. You—you brave flower of all our knowing.

CLAIRE: I am not a flower. I am too torn. If you have anything—help me. Breathe, Breathe the healing oneness, and let me know in calm. (*with a sob his head rests upon her*)

CLAIRE: (*her hands on his head, but looking far*) Beauty—you pure one thing. Breathe—Let me know in calm. Then—trouble me, trouble me, for other moments—in farther calm. (*slow, motionless, barely articulate*)

TOM: (*as she does not move he lifts his head. And even as he looks at her, she does not move, nor look at him*) Claire—(*his hand out to her, a little afraid*) You went away from me then. You are away from me now.

CLAIRE: Yes, and I could go on. But I will come back, (*it is hard to do. She brings much with her*) That, too, I will give you—my by-myself-ness. That's the uttermost I can give. I never thought—to try to give it. But let us do it—the great sacrilege! Yes! (*excited, she rises; she has his hands, and bring him up beside her*) Let us take the mad chance! Perhaps it's the only way to save—what's there. How do we know? How can we know? Risk. Risk everything. From all that flows into us, let it rise! All that we never thought to use to make a moment—let it flow into what could be! Bring all into life between us—or send all down to death! Oh, do you know what I am doing? Risk, risk everything, why are you so afraid to lose? What holds you from me? Test all. Let it live or let it die. It is our chance—our chance to bear—what's there. My dear one—I will love you so. With all of me. I am not afraid now—of—all of me. Be generous. Be unafraid. Life is for *life*—though it cuts us from the farthest life. How can I make you know that's true? All that we're open to—(*hesitates, shudders*) But yes—I will, I will risk the life that waits. Perhaps only he who gives his loneliness—shall find. You never keep by holding, (*gesture of giving*) To the uttermost. And it is gone—or it is there. You do not know and—that makes the moment—(*music has begun—a phonograph downstairs; they do not heed it*) Just as I would cut my wrists—(*holding them out*) Yes, perhaps this lesser thing will tell it—would cut my wrists and let the blood flow out till all is gone if my last drop would make—would make—(*looking at them fascinated*) I want to see it doing that! Let me give my last chance for life to—

(*He snatches her—they are on the brink of their moment; now that there are no words the phonograph from downstairs is louder. It is playing languorously the Barcarole; they become conscious of this—they do not want to be touched by the love song.*)

CLAIRE: Don't listen. That's nothing. This isn't that, (*fearing*) I tell you—it isn't that. Yes, I know—that's amorous—enclosing. I know—a little place. This isn't that, (*her arms going around him—all the lure of 'that' while she pleads against it as it comes up to them*) We will come out—to radiance—in far places (*admitting, using*) Oh, then let it be that! Go with it. Give up—the otherness. I will! And in the giving up—perhaps a door—we'd never find by searching. And if it's no more—than all have known, I only say it's worth the allness! (*her arms wrapped round him*) My love—my love—let go your pride in loneliness and let me give you joy!

From Act Three:

[…]

CLAIRE: You're not going away?

TOM: Not without you, Claire. And you and I will be together. Is that—what you wanted?

CLAIRE: Wanted? (*as if wanting is something that harks far back. But the word calls to her passion*) Wanted! (*a sob, hands out, she goes to him. But before his arms can take her, she steps back*) Are you trying to pull me down into what I wanted? Are you here to make me stop?

TOM: How can you ask that? I love you because it is not in you to stop.

CLAIRE: And loving me for that—would stop me? Oh, help me see it! It is so important that I see it.

TOM: It is important. It is our lives.

CLAIRE: And more than that. I cannot see it because it is so much more than that.

TOM: Don't try to see all that it is. From peace you'll see a little more.

CLAIRE: Peace? (*troubled as we are when looking at what we cannot see clearly*) What is peace? Peace is what the struggle knows in moments very far apart. Peace—that is not a place to rest. Are you resting? What are you? You who'd take me from what I am to something else?

TOM: I thought you knew, Claire.

CLAIRE: I know—what you pass for. But are you beauty? Beauty is that only living pattern—the trying to take pattern. Are you trying?

TOM: Within myself, Claire. I never thought you doubted that.

CLAIRE: Beauty is it. (*she turns to Breath of Life, as if to learn it there, but turns away with a sob*) If I cannot go to you now—I will always be alone.

(TOM *takes her in his arms. She is shaken, then comes to rest.*)

TOM: Yes—rest. And then—come into joy. You have so much life for joy.

CLAIRE: (*raising her head, called by promised gladness*) We'll run around together. (*lovingly he nods*) Up hills. All night on hills.

TOM: (*tenderly*) All night on hills.

CLAIRE: We'll go on the sea in a little boat.

TOM: On the sea in a little boat.

CLAIRE: But—there are other boats on other seas, (*drawing back from him, troubled*) There are other boats on other seas.

TOM: (*drawing her back to him*) My dearest—not now, not now.

CLAIRE: (*her arms going round him*) Oh, I would love those hours with you. I want them. I want you! (*they kiss—but deep in her is sobbing*) Reminiscence, (*her hand feeling his arm as we touch what we would remember*) Reminiscence. (*with one of her swift changes steps back from him*) How dare you pass for what you're not? We are tired, and so we think it's you. Stop with you. Don't get through—to what you're in the way of. Beauty is not something you say about beauty.

TOM: I say little about beauty, Claire.

CLAIRE: Your life says it. By standing far off you pass for it. Smother it with a life that passes for it. But beauty—(*getting it from the flower*) Beauty is the humility breathed from the shame of succeeding.

TOM: But it may all be within one's self, dear.

CLAIRE: (*drawn by this, but held, and desperate because she is held*) When I have wanted you with all my wanting—why must I distrust you now? When I love you—with all of me, why do I know that only you are worth my hate?

TOM: It's the fear of easy satisfactions. I love you for it.

CLAIRE: (*over the flower*) Breath of Life—you here? Are you lonely—Breath of Life?

TOM: Claire—hear me! Don't go where we can't go. As there you made a shell for life within, make for yourself a life in which to live. It must be so.

CLAIRE: As you made for yourself a shell called beauty?

TOM: What is there for you, if you'll have no touch with what we have?

CLAIRE: What is there? There are the dreams we haven't dreamed. There is the long and flowing pattern, (*she follows that, but suddenly and as if blindly goes to him*) I am tired. I am lonely. I'm afraid, (*he holds her, soothing. But she steps back from him*) And because we are tired—lonely—and afraid, we stop with you. Don't get through—to what you're in the way of.

TOM: Then you don't love me?

CLAIRE: I'm fighting for my chance. I don't know—which chance.

(*Is drawn to the other chance, to Breath of Life. Looks into it as if to look through to the uncaptured. And through this life just caught comes the truth she chants.*)

I've wallowed at a coarse man's feet,

I'm sprayed with dreams we've not yet come to.

I've gone so low that words can't get there,

I've never pulled the mantle of my fears around me

And called it loneliness—And called it God.

Only with life that waits have I kept faith.

(*with effort raising her eyes to the man*)

And only you have ever threatened me.

TOM: (*coming to her, and with strength now*) And I will threaten you. I'm here to hold you from where I know you cannot go. You're trying what we can't do.

CLAIRE: What else is there worth trying?

TOM: I love you, and I will keep you—from fartherness—from harm. You are mine, and you will stay with me! (*roughly*) You hear me? You will stay with me!

CLAIRE: (*her head on his breast, in ecstasy of rest. Drowsily*) You can keep me?

TOM: Darling! I can keep you. I will keep you—safe.

CLAIRE: (*troubled by the word, but barely able to raise her head*) Safe?

TOM: (*bringing her to rest again*) Trust me, Claire.

CLAIRE: (*not lifting her head, but turning it so she sees Breath of Life*) Now can I trust—what is? (*suddenly pushing him roughly away*) No! I will beat my life to pieces in the struggle to—

TOM: To *what*, Claire?

CLAIRE: Not to stop it by seeming to have it. (*with fury*) I will keep my life low—low—that I may never stop myself—or anyone—with the thought it's what *I* have. I'd rather be the steam rising from the manure than be a thing called beautiful! (*with sight too clear*) Now I know who you are. It is you puts out the breath of life. Image of beauty—*You fill the place—should be a gate.* (*in agony*) Oh, that it is *you*—fill the place—should be a gate! My darling! That it should be you who—(*her hands moving on him*) Let me tell you something. Never was loving strong as my loving of you! Do you know that? Oh, know that! Know it now! (*her arms go around his neck*) Hours with you—I'd give my life to have! That it should be you—(*he would loosen her hands, for he cannot breathe. But when she knows she is choking him, that knowledge is fire burning its way into the last passion*) It *is* you. It is you.

TOM: (*words coming from a throat not free*) Claire! What are you doing? (*then she knows what she is doing*)

CLAIRE: (*to his resistance*) No! You are *too much*! You are *not enough*. (*still wanting not to hurt her, he is slow in getting free. He keeps stepping backward trying, in growing earnest, to loosen her hands. But he does not loosen them before she has found the place in his throat that cuts off breath. As he gasps*)

Breath of Life—my gift—to you!

(*She has pushed him against one of the plants at right as he sways, strength she never had before pushes him over backward, just as they have struggled from sight. Violent crash of glass is heard.*)

TOM: (*faint smothered voice*) *No*. I'm—hurt.

CLAIRE: (*in the frenzy and agony of killing*) Oh, gift! Oh, gift! (*there is no sound.*

CLAIRE *rises—steps back—is seen now; is looking down*) Gift.

(*Like one who does not know where she is, she moves into the room—looks around. Takes a step toward Breath of Life; turns and goes quickly to the door. Stops, as if stopped. Sees the revolver where the Edge Vine was. Slowly goes to it. Holds it as if she cannot think what it is for. Then raises it high and fires above through the place in the glass left open for ventilation*. ANTHONY *comes from the inner room. His eyes go from her to the body beyond*. HARRY *rushes in from outside*.)

HARRY: Who fired that?

CLAIRE: I did. Lonely.

(*Seeing* ANTHONY'S *look*, HARRY *'s eyes follow it*.)

HARRY: Oh! What? What? (DICK *comes running in*) Who? Claire!

(DICK *sees—goes to* TOM)

CLAIRE: Yes. I did it. MY—Gift.

HARRY: Is he—? He isn't—? He isn't—?

(*Tries to go in there. Cannot—there is the sound of broken glass, of a position being changed—then* DICK *reappears*.)

DICK: (*his voice in jerks*) It's—it's no use, but I'll go for a doctor.

HARRY: No—no. Oh, I suppose—(*falling down beside* CLAIRE—*his face against her*) My darling! How can I save you now?

CLAIRE: (*speaking each word very carefully*) Saved—myself.

ANTHONY: I did it. Don't you see? I didn't want so many around. Not—what this place is for.

HARRY: (*snatching at this but lets it go*) She wouldn't let—(*looking up at* CLAIRE—*then quickly hiding his face*) And—don't you see?

CLAIRE: Out. (*a little like a child's pleased surprise*) Out.

(DICK *stands there, as if unable to get to the door—his face distorted, biting his hand*.)

ANTHONY: Miss Claire! You can do anything—won't you try?

CLAIRE: Reminiscence? (*speaking the word as if she has left even that, but smiles a little*)

(ANTHONY *takes Reminiscence, the flower she was breeding for fragrance for Breath of Life—holds it out to her. But she has taken a step forward, past them all*.)

CLAIRE: Out. (*as if feeling her way*)

Nearer, (*Her voice now feeling the way to it.*)

Nearer— (*Voice almost upon it.*)

—my God, (*Falling upon it with surprise.*)

to Thee, (*Breathing it*.)

Nearer—to Thee,

E'en though it be— (*A slight turn of the head toward the dead man she loves—a mechanical turn just as far the other way.*)

a cross

That (*Her head going down.*)

raises me; (*Her head slowly coming up—singing it.*)

Still all my song shall be,

Nearer, my—

(*Slowly the curtain begins to shut her out. The last word heard is the final* Nearer—*a faint breath from far*.)

CURTAIN