Evan Hepler-Smith

Harvard University Center for the Environment

"A way of thinking backwards": Computing and method in synthetic organic chemistry

Forthcoming in *Historical Studies in the Natural Sciences*, 47 (2018)

March 21, 2018

This article addresses the history of retrosynthetic analysis, a key method that chemists use to plan organic chemical syntheses, and LHASA, a computer program intended to aid chemists in this task. The chemist E. J. Corey developed this method and computer program side-by-side, from the early 1960s through the 1990s. While the LHASA program never came into widespread use, retrosynthetic analysis became a standard method for teaching and practicing synthetic planning, a subject previously taken as resistant to generalization. This article shows how the efforts of Corey and his collaborators to make synthetic planning tractable to computer automation shaped a way of thinking taken up by chemists, unaided by machines. The method of retrosynthetic analysis made chemical thinking (as Corey perceived it) explicit, in accordance with the demands of computing (as Corey and his LHASA collaborators perceived them).

> This book is about making molecules. Or rather it is to help you design your own syntheses by logical and sensible thinking. This is not a matter of guesswork but requires a way of thinking backwards…
> - Warren & Wyatt, *Organic Synthesis: The Disconnection Approach,* 2nd ed. (2008), 1.

This article is about making plans for making molecules. Or rather it is a history of how synthetic organic chemists in the mid-to-late twentieth century thought about thinking, about computing, and about the relationship between the two. The thinking in question involved devising sequences of materials and reactions for building chemical substances, an activity known as *synthetic planning*. Midcentury specialists took synthetic planning to be a matter of creative genius: an irreducible, hard-to-define capacity to choreograph novel chemical transformations. During the early 1960s, the chemist E. J. Corey embarked upon a project to recast this thinking as a general method that could be taught to students and programmed on a computer. Corey refashioned synthetic planning as *retrosynthetic analysis*, a systematic way of thinking backwards.

"Synthesis, the making of molecules," the chemist and essayist Roald Hoffmann has written, "is at the heart of chemistry – the art, craft, business, and science of substances."[1] Synthetic organic chemistry – the construction of chemical substances containing the element carbon – has shaped the modern molecular sciences and the chemically-built

[1] Roald Hoffmann, "Why Think Up New Molecules," in *Roald Hoffmann on the Philosophy, Art, and Science of Chemistry*, ed. Jeffrey Kovac and Michael Weisberg (New York: Oxford University Press, 2012), 114–119, on 114.

world. Plastics, pesticides, drugs, and myriad other synthetic organic chemicals account for roughly three *trillion* dollars of annual revenue for chemical and pharmaceutical firms around the world, and constitute prodigious boons and threats to humans and the environment.[2] Research in organic synthesis has also yielded considerable insight into chemical and biochemical phenomena, and into how they may be controlled. In the year 2000, the authors of a review could plausibly claim that organic synthesis was largely "responsible for some of the most exciting and important discoveries of the twentieth century in chemistry, biology, and medicine."[3]

The synthesis of complex naturally-occurring chemical substances has often been cited as an especially rich source of such dividends. On these grounds, governments and pharmaceutical firms have generously supported research in "total synthesis," a resource-intensive enterprise that only rarely yields commercially-viable processes for preparing the compound in question.[4] Synthetic planning, also known as synthetic design, sets the framework for such synthesis projects, laying out a tentative series of chemical operations by which the chemist will attempt to join and modify selected starting materials to obtain

[2] American Chemistry Council, *2016 Guide to the Business of Chemistry* (Arlington, VA: American Chemistry Council, 2016), 37.

[3] K. C. Nicolaou et al., "The Art and Science of Total Synthesis at the Dawn of the Twenty-First Century," *Angewandte Chemie International Edition* 39, no. 1 (2000): 44–122, on 45. On the synthetic sciences more broadly, see William R. Newman and Bernadette Bensaude-Vincent, *The Artificial and the Natural: An Evolving Polarity* (Cambridge, MA: MIT Press, 2007); Sophia Roosth, *Synthetic: How Life Got Made* (Chicago: University of Chicago Press, 2017).

[4] "Total synthesis," coined by the French chemist Marcellin Berthelot in 1856, denotes the preparation of a compound from simple starting materials (in principle, from chemical elements) via chemical reactions; Alan J. Rocke, *Nationalizing Science: Adolphe Wurtz and the Battle for French Chemistry* (Cambridge, MA: MIT Press, 2001), 248. Most commercial processes for drug production involve partial synthesis (synthesis from a closely-related compound derived from a natural source) or biosynthesis (synthesis by the action of microorganisms or enzymes).

the desired product.[5] In natural products synthesis, a group of practitioners explained, "it is the journey toward the target molecule that becomes the essence and significance of the exercise."[6] In the laboratory, such "journeys" involve substantial labor and organization, skilled manipulations of chemicals and instruments, and tuning of reactions by trial and error. Synthetic planning lays the groundwork for these material stages of synthesis, proposing the routes, detours, and shortcuts that a pathway from starting materials to product might traverse.

E. J. Corey has been one of the foremost contributors to the field of natural products synthesis since the mid-twentieth century. During a decade at the University of Illinois, more than fifty years at Harvard, and a longstanding role as an advisor to the pharmaceutical firm Pfizer, Corey built up a wide-ranging research program. As of 2003, he was the most cited author in the history of the *Journal of the American Chemical Society*.[7] In 1960, however, when Corey arrived at Harvard and was tasked with teaching organic synthesis for the first time, the acknowledged master of the field was another Harvard chemist: R. B. Woodward. Woodward cast first-rate synthetic design as an expression of the singular insights of a particular chemist about structural quirks of a particular chemical. Corey thought differently. Discerning common patterns behind disparate synthetic plans, Corey began a long-term project to develop a general approach to synthetic design that students could learn, chemists could apply, and a computer could be programmed to carry

[5] Following the usage of chemists, I use "synthetic planning" and "synthetic design" interchangeably.

[6] K. C. Nicolaou, E. J. Sorensen, and N. Winssinger, "The Art and Science of Organic and Natural Products Synthesis," *Journal of Chemical Education* 75, no. 10 (1998): 1225–58, on 1232.

[7] P. J. Stang, "124 Years of Publishing Original and Primary Chemical Research," *Journal of the American Chemical Society* 125, no. 1 (2003): 1–8, on 1.

out. The results of Corey's project were a method, which he called retrosynthetic analysis, and a computer program, LHASA.

Corey's method, his computer program, and the relationship between the two present a historical puzzle. Retrosynthetic analysis is often counted among the most significant developments in recent chemistry. In numerous reviews and awards, including Corey's 1990 Nobel Prize, Corey's colleagues have recognized retrosynthetic analysis as a transformative approach to the pedagogy and practice of synthetic planning. Nevertheless, many chemists – including Corey himself – have also characterized retrosynthetic analysis simply as chemical common sense, a method that skilled synthetic chemists employed, perhaps unconsciously, before Corey ever described it.[8]

The LHASA program, in contrast, is generally regarded as a pioneering foray in chemical artificial intelligence that never quite caught on as a practical tool. Chemists working in this field today attribute the (so far) limited application of computing to synthetic design in part to ambitions that outstripped the computing power and human labor needed to realize them. Most of all, they point to the pleasure and pride that chemists have taken in synthetic planning: "the chemist's belief that they could do a better job than any computer program," as one review put it.[9] Still, some have discerned a relationship between the project to get computers to plan syntheses and the method by which chemists began to do so in the wake of this project. LHASA and its successors "may have fallen out of

[8] On the historiography of recent chemistry, see Peter J. T. Morris, "The Fall and Rise of the History of Recent Chemistry," *Ambix* 58, no. 3 (2011): 238–56.

[9] Anthony Cook et al., "Computer-Aided Synthesis Design: 40 Years On," *Wiley Interdisciplinary Reviews: Computational Molecular Science* 2, no. 1 (2012): 79–107, on 81. On the current state of this field (as characterized by contributors to one active project), see Tomasz Klucznik et al., "Efficient Syntheses of Diverse, Medicinally Relevant Targets Planned by Computer and Executed in the Laboratory," *Chem* 4, no. 3 (2018): 522–32; Sara Szymkuć et al., "Computer-Assisted Synthetic Planning: The End of the Beginning," *Angewandte Chemie International Edition* 55, no. 20 (2016): 5904–37.

favour," notes a specialist in chemical expert systems, but "they fundamentally changed the way research chemists think and organic chemistry is taught."[10] The authors of a popular textbook introduce retrosynthetic analysis as an outgrowth of the LHASA project. They write, "[T]he logic survives as a way of planning syntheses used by almost all organic chemists. It is more useful to humans than to machines."[11]

These characterizations of retrosynthetic analysis and the LHASA program pose three historical questions. First, how did Corey recast a way of thinking previously defined by its idiosyncrasy as the expression of a general method, and what role did the LHASA project play in this process? Second, how could retrosynthetic analysis be taken at once as groundbreaking *and* as the expression of a well-established way of thinking about synthetic design? (I will not aim to determine whether the method was "really" novel, but to understand how this curious dual perspective came to be.) Third, how did a method developed expressly for a chemical computer program catch on among human chemists, especially when the program itself did not?

I argue that Corey made synthetic planning subject to method by taking synthetic planning itself – making plans, rather than making molecules – as an object of study. At first, he did so in order to make synthetic planning teachable. Then, in successive and mutually reinforcing projects, he sought to subject synthetic planning to a general method, and he brought together a collection of graduate students and postdoctoral fellows, the "LHASA group," to program a computer to carry out this method. Through their efforts to

[10] Phillip Judson, *Knowledge-Based Expert Systems in Chemistry: Not Counting on Computers* (Cambridge: Royal Society of Chemistry, 2009), 31.

[11] Stuart G. Warren and Paul Wyatt, *Organic Synthesis: The Disconnection Approach*, 2nd ed. (Chichester, UK: Wiley, 2008), 2.

render synthetic planning tractable to computer automation, Corey and the LHASA group developed retrosynthetic analysis. Brought back into the classroom, this method became a way of thinking for human chemists, unaided by machines.

The first part of this essay situates Corey's initial reflections on synthetic design in the historical development of organic synthesis through the mid-twentieth century and the pedagogical challenge posed by Woodward's approach to the subject. The second section follows Corey's turn from pedagogy to practice. Starting from the presumption that existing synthetic plans were expressions of unrecognized but widely-applicable principles, Corey set out a framework for enumerating the elemental units of thinking that went into synthetic design and assembling them into a general method. The third section addresses the early years of the LHASA program, a venture in computer automation that Corey saw as an apposite medium for his method. The members of the LHASA group were the kind of intermediaries that the historian Carsten Reinhardt has called "method-makers": they took up a technology unfamiliar to most of their colleagues and crafted a computer-based instrument, a method (retrosynthetic analysis), and a chemical vernacular for both.[12] The fourth section of this essay addresses how Corey and others made retrosynthetic analysis a standard approach to the pedagogy and practice of synthetic planning, eliding the role of computing in its development.

This article also speaks to fundamental historical questions about the relationship among computing, method-making, artificial intelligence, and scientific thinking.[13] Within

[12] Carsten Reinhardt, *Shifting and Rearranging: Physical Methods and the Transformation of Modern Chemistry* (Sagamore Beach, MA: Science History Publications/USA, 2006).

[13] Similar themes are at issue in the history of early modern "thinking about thinking," especially in mathematics. See, e.g., Lorraine J. Daston, "The Physicalist Tradition in Early Nineteenth Century French Geometry," *Studies in History and Philosophy of Science Part A* 17, no. 3 (1986): 269–95; Matthew L. Jones,

fields such as natural history and government administration, applications of computing appear to have fit into well-established methods of reasoning. In such domains, the historian Jon Agar has argued, "computerization only took place where there were existing material practices of computation."[14] In other fields, the transformation of cognitive tasks into processes that a computer could carry out seems to have generated entirely new methods, for machines, humans, or both. In contemporary biology, for example, "databases and algorithms determine what sorts of objects exist and relationships between them," the historian and ethnographer Hallam Stevens contends.[15] The development of retrosynthetic analysis suggests a middle path. The requirements of writing a computer program – or, more precisely, Corey and his team's interpretation of these requirements – impelled the LHASA group to express existing practices more explicitly. The conclusion of this article suggests how focusing on such examples of computerization as explication may provide a useful perspective on the history of projects to think on, with, or like machines.

Reckoning with Matter: Calculating Machines, Innovation, and Thinking about Thinking from Pascal to Babbage (Chicago: University of Chicago Press, 2016).

[14] Jon Agar, "What Difference Did Computers Make?," *Social Studies of Science* 36, no. 6 (2006): 869–907, on 898–900. See also David Sepkoski, "The Database before the Computer?," *Osiris* 32, no. 1 (2017): 175–201.

[15] Hallam Stevens, *Life out of Sequence: A Data-Driven History of Bioinformatics* (Chicago: University of Chicago Press, 2013), 5; Hallam Stevens, "A Feeling for the Algorithm: Working Knowledge and Big Data in Biology," *Osiris* 32, no. 1 (2017): 151–74. For an analogous argument regarding mathematical theorem-proving, see Stephanie Dick, "AfterMath: The Work of Proof in the Age of Human–Machine Collaboration," *Isis* 102, no. 3 (2011): 494–505; Stephanie Dick, "Of Models and Machines: Implementing Bounded Rationality," *Isis* 106, no. 3 (2015): 623–34.

Teaching the unteachable

In 1957, the American Chemical Society presented its first "Award for Creative Work in Synthetic Organic Chemistry," a prize funded by the Synthetic Organic Chemical Manufacturers' Association.[16] The award marked the ascendance of a perspective casting natural products synthesis as a distinct field at the forefront of organic chemistry, synthesis as an engine of progress for chemical science and industry in general, and creative synthetic planning as the heart of synthesis. The recipient of the 1957 prize, the Harvard chemist Robert Burns Woodward, both exemplified this approach to synthesis and had done much to create it. The origins of retrosynthetic analysis lay in Corey's attempt to grapple with the pedagogical challenge presented by synthetic planning, a subject that Woodward defined both as central to the progress of chemistry and as all but unteachable.

"Wherever the historian has adjudicated the past from the lofty pinnacle of the present, he has over-rated the role of the organic synthesis," the historian John Hedley Brooke observed in 1971.[17] The target of Brooke's critique was the total synthesis of natural products – the default meaning of "organic synthesis" at a moment when Woodward and Albert Eschenmoser were completing their landmark synthesis of vitamin B_{12}. Considered as a broader collection of practices, however, synthesis played a central role in the development of organic chemistry from the origins of the field in the early nineteenth century. While chemists accomplished relatively few total syntheses of natural products in the nineteenth century (by the reckoning of their twentieth-century

[16] "Three New Awards," *Chemical & Engineering News* 33, no. 49 (1955): 5292; "People," *Chemical & Engineering News* 37, no. 49 (1959): 114–26, on 114.

[17] John Hedley Brooke, "Organic Synthesis and the Unification of Chemistry—A Reappraisal," *The British Journal for the History of Science* 5, no. 4 (1971): 363–92, on 375.

successors), they made extensive use of synthesis directed towards exploration of open questions rather than the preparation of specific products. Chemists used synthesis to investigate general patterns of chemical behavior and relationships, the constitution of particular chemical substances, the production of commercially valuable products, and the unification of organic and inorganic chemistry.[18] The planning involved in these syntheses entailed coordinating laboratory instruments, principles of chemical theory, and provisional formula models employed as "paper tools," adjusting each over the course of an investigation to stabilize both empirical results and interpretations.[19] In the process, chemists conceived new standards for isolating pure substances, new characteristic properties to measure, and new experimental instruments.[20] The organic chemists who developed the principles of structure theory – the chemistry of atoms, bonds, and molecules – relied on evidence drawn from exploratory "synthetical experiments."[21] So did those who opposed this theory. Indeed, the most energetic proponent of total synthesis as a foundation for chemical knowledge, the French chemist Marcellin Berthelot, was also the staunchest opponent of atomism and structure theory.[22]

---

[18] C. A. Russell, "The Changing Role of Synthesis in Organic Chemistry," *Ambix* 34, no. 3 (1987): 169–80.

[19] Ursula Klein, *Experiments, Models, Paper Tools: Cultures of Organic Chemistry in the Nineteenth Century* (Stanford: Stanford University Press, 2003).

[20] Catherine M. Jackson, "The Curious Case of Coniine: Constructive Synthesis and Aromatic Structure Theory," in *Objects of Chemical Inquiry*, ed. Ursula Klein and Carsten Reinhardt (Sagamore Beach, MA: Science History Publications/USA, 2014), 61–102; Catherine M. Jackson, "Chemical Identity Crisis: Glass and Glassblowing in the Identification of Organic Compounds," *Annals of Science* 72, no. 2 (2015): 187–205.

[21] Catherine M. Jackson, "Synthetical Experiments and Alkaloid Analogues: Liebig, Hofmann and the Origins of Organic Synthesis," *Historical Studies in the Natural Sciences* 44, no. 4 (2014): 319–63. On the development of structure theory, see Alan J. Rocke, *Image and Reality: Kekulé, Kopp, and the Scientific Imagination* (Chicago: University of Chicago Press, 2010).

[22] Rocke, *Nationalizing Science* (ref. 4), 235–67, 301–31.

Making use of the insights into chemical relationships afforded by structure theory, chemists of the late nineteenth and early twentieth centuries used constructive synthesis and degradative analysis as complementary techniques for determining the structure of organic compounds.[23] The meticulous laboratory operations and elaborate chains of inference that went into structure determination yielded novel chemical generalizations, techniques, and substances – foundations for the further development of organic chemistry and for the emerging synthetic chemical and pharmaceutical industries.[24] In the chemistry of natural product structure determination of the first half of the twentieth century, both the plan of a synthesis and the structural formula assigned to the product were representations of a series of chemical transformations effected by well-understood chemical reactions, underwritten by painstaking labor at the bench. Such projects involved substantial ongoing planning, as attested by the long series of preliminary publications that they often yielded.[25] The outcome of a successful synthesis research project of this sort was a confirmed structural formula and a plan of synthesis, a specific series of laboratory-attested chemical transformations that served as evidence for the formula of the product.

Developments in physical organic chemistry and physical analytical instrumentation during the 1930s – 1950s laid the groundwork for the emergence of natural products synthesis as a distinct field of organic chemistry. Analytical methods using physical instruments – especially infrared spectroscopy, mass spectrometry, and nuclear magnetic

[23] For a detailed comparison of total synthesis and degradative analysis, see Jeffrey I. Seeman, "On the Relationship between Classical Structure Determination and Total Synthesis," forthcoming in *Israel Journal of Chemistry* 57 (2017).

[24] Leo B. Slater, "Woodward, Robinson, and Strychnine: Chemical Structure and Chemists' Challenge," *Ambix* 48, no. 3 (2001): 161–89.

[25] Andre Siegel, "Sir Robert Robinson's 'Anthocyanin Period': 1922–1934 — A Case Study of an Early Twentieth-Century Natural Products Synthesis," *Ambix* 55, no. 1 (2008): 62–82.

resonance spectroscopy – provided rapid access to information about chemical structure previously hard-won over months at the bench, a transformation often called chemistry's "instrumental revolution."[26] Valence-bond and molecular orbital theories adapted quantum physics to provide general accounts of chemical bonding.[27] Mechanistic and steric studies opened the black box of the chemical reaction, shedding light on possible pathways of chemical change and the factors that shaped them.[28] "Method-makers" straddling physics, engineering, and organic chemistry made instrument data and physical theories into useful tools for organic chemists.[29] By making structural formulas more easily accessible and unshackling them from specific series of chemical reactions, physical instruments enabled chemists to treat these diagrams more confidently as representations of molecular reality. Thanks to rules of thumb grounded in quantum, mechanistic, and steric principles, chemists could put these newly mobilized structural formulas to use in proposing novel chemical reactions and predicting their likely outcomes.[30]

R. B. Woodward was among this vanguard of method-makers. As a postdoctoral fellow during the late 1930s and early 1940s and throughout his career, he worked out rules for expressing instrumental data and the predictions of physical theories in terms of

[26] Peter Morris, ed., *From Classical to Modern Chemistry: The Instrumental Revolution* (Cambridge: Royal Society of Chemistry, 2002).

[27] Kostas Gavroglu and Ana Simões, *Neither Physics nor Chemistry: A History of Quantum Chemistry* (Cambridge, MA: MIT Press, 2012).

[28] Mary Jo Nye, *From Chemical Philosophy to Theoretical Chemistry: Dynamics of Matter and Dynamics of Disciplines, 1800-1950* (Berkeley: University of California Press, 1993), 196–223; Jeffrey Allan Johnson, "The Case of the Missing German Quantum Chemists: On Molecular Models, Mobilization, and the Paradoxes of Modernizing Chemistry in Nazi Germany," *Historical Studies in the Natural Sciences* 43, no. 4 (2013): 391–452.

[29] Reinhardt, *Shifting and Rearranging* (ref. 12).

[30] Albert Eschenmoser and Claude E. Wintner, "Natural Product Synthesis and Vitamin B_{12}," *Science* 196, no. 4297 (1977): 1410–20, on 1410–13.

structural organic chemistry.[31] He also demonstrated the value of these methods, using them extensively in directing syntheses of noteworthy and complex chemical substances such as quinine, strychnine, cholesterol, and chlorophyll during the 1940s and 1950s. These achievements (along with considerable rhetorical style and self-regard) made Woodward a leading authority in his field – "the world's leading exponent of organic chemistry," as a pair of colleagues put it.[32]

During the 1950s and 1960s, Woodward promoted an influential conception of "synthesis for its own sake" as a kind of basic science of organic chemistry.[33] Woodward wrote that chemists should not "suppose that the challenge for the hand and the intellect must be less, or the fruits less tantalizing, when chemistry *begins* at the advanced vantage point of an established structure." Chemists could redirect their ingenuity away from determining molecular structures and toward assembling them. Woodward emphasized that chemists could look beyond well-understood reactions in pursuing their synthesis projects, since they could use physical instruments to determine the identity of reaction products at each step. Chemists could therefore use natural products syntheses as an occasion to propose and test entirely new reagents and transformations, yielding a

[31] Leo B. Slater, "Instruments and Rules: R. B. Woodward and the Tools of Twentieth-Century Organic Chemistry," *Studies in History and Philosophy of Science Part A* 33, no. 1 (2002): 1–33.

[32] On Woodward's career and character, see O. Theodor Benfey and Peter John Turnbull Morris, eds., *Robert Burns Woodward: Architect and Artist in the World of Molecules* (Philadelphia: Chemical Heritage Foundation, 2001); Leo Barney Slater, "Organic Synthesis and R. B. Woodward: An Historical Study in the Chemical Sciences" (PhD dissertation, Princeton University, 1997); Jeffrey I. Seeman, "R. B. Woodward: A Larger-than-Life Chemistry Rock Star," *Angewandte Chemie* 129, no. 34 (2017): 10362–79. Quotation from Alexander Todd and John Cornforth, "Robert Burns Woodward," *Biographical Memoirs of Fellows of the Royal Society* 27 (1981): 628–95, on 641.

[33] R. B. Woodward, "Synthesis," in *Perspectives in Organic Chemistry*, ed. Alexander R. Todd (New York: Interscience, 1956), 155–84, on 158.

"dividend of unsought fact" that would advance organic chemistry and its industrial and biomedical applications.[34]

According to Woodward, synthetic planning was the key to this enterprise. "Synthesis must always be carried out by plan," he wrote in 1956, "and the synthetic frontier can be defined only in terms of the degree to which realistic planning is possible." Using quantum, mechanistic, and steric considerations, chemists could assemble series of structural formulas representing plausible, though perhaps novel, steps in the transformation of starting materials into a desired product. Planning could only be a first stage in a synthesis project. Determining specific reagents and conditions to optimize each synthetic step and isolate its products took careful laboratory work, skilled judgment, and inventive use of laboratory instrumentation. Synthetic plans inevitably had to be adjusted as the synthesis was carried out in the laboratory; predictions regarding novel chemical transformations remained uncertain (then and now). However, Woodward referred to the laboratory portions of synthesis as "reduction of plans to practice." In contrast, he cast synthetic planning as the crucial act of creative intellectual authorship that stood behind a natural products synthesis.[35]

Woodward presented the singularity of each natural product and each synthetic plan as the basis for the intellectual appeal of synthesis and its contributions to chemistry. Sometimes, he did so by means of military metaphors: the synthetic chemist as campaigning general, substances yet to be synthesized as "specific objectives," some soon

[34] R. B. Woodward et al., "The Total Synthesis of Strychnine," *Tetrahedron* 19, no. 2 (1963): 247–88, on 248.

[35] Woodward, "Synthesis" (ref. 33), 155–56.

to see "their defenses crumble."[36] Speaking in Bombay, Woodward invoked a more pacific imperialism: "The structure known, but not yet accessible to synthesis, is to the chemist what the unclimbed mountain, the uncharted sea, the untilled field, the unreached planet are to other men."[37] Such images of masculinity, conquest, and empire cast each synthesis as an individual triumph over an individual challenge.[38]

Such singular achievements, Woodward emphasized, gave rise to the unexpected. He referred to chemical theory and experiment as "basic tools" for designing and executing syntheses, distinguishing synthetic planning – the manipulation of sequences of structural formulas on the mind, page, and blackboard – as a distinct aspect of chemical practice. Woodward described synthetic planning as "essentially entirely a creative activity, in which art, design, imagination, and inspiration play a predominant role."[39] This creative activity had to satisfy pragmatic criteria – minimizing the number of steps in the synthesis and maximizing the efficiency (or "yield") of each step. However, Woodward contended that the proper virtue of a synthetic plan lay in its "delightful elements of surprise": the careful selection of a reactant such that a seemingly extraneous structural subunit could later be

[36] R. B. Woodward, "The Total Synthesis of Chlorophyll," *Pure and Applied Chemistry* 2, no. 3–4 (1961): 383–404, on 383; Woodward, "Synthesis" (ref. 33), 158.

[37] R. B. Woodward, "Art and Science in the Synthesis of Organic Compounds: Retrospect and Prospect," in *Pointers and Pathways in Research: Six Lectures in the Fields of Organic Chemistry and Medicine*, ed. Maeve O'Connor (Bombay: CIBA of India Ltd., 1963), 22–41, on 41.

[38] Chemist Anne M. Wilson has connected the chronic underrepresentation of women in organic synthesis compared with other fields of chemistry to a culture of "'macho' syntheses of exceedingly complicated natural products" (quoting chemist Carl Djerassi); Anne M. Wilson, "Harry S. Mosher and Arthur C. Cope, Early Organic Chemists Who Mentored Women," *Bulletin for the History of Chemistry* 34, no. 1 (2009): 21–29, on 22. The LHASA project, itself marginal within the Harvard chemistry department, occupied space left open by this underrepresentation; a women's bathroom was commandeered as temporary office space for two (male) group members. Author interviews with David Pensak, 7 Mar and 17 Mar 2013.

[39] Woodward, "Art and Science" (ref. 37), 28. Visual imagination and creative manipulation of structural formulas had long played an important role in organic chemistry; see Klein, *Experiments, Models, Paper Tools* (ref. 19); Rocke, *Image and Reality* (ref. 21).

incorporated into the final product; the meticulous effort to build a molecular assemblage that would spontaneously rearrange itself to form a desired structure; the invention of an entirely new form of matter as an key intermediate.[40] In such situations, Woodward explained:

> An apparently rather dull grouping of atoms suddenly, under the impact of especially chosen reactants, undergoes unusual transformations which are of great utility in progress toward the objective. The impact on an observer may perhaps be compared with that on the traveller down an uninteresting street, who turns through a small hidden doorway into a delightful and charming garden.[41]

According to Woodward, such leaps of creative imagination were the wellspring of "a rich dividend of new and often unexpected fact which can be of enormous value in extending the frontiers of the science."[42] Woodward tied such claims about the productivity of synthetic planning to his insistence that it could not be reduced to method. "Although the experimental aspects of some kinds of synthetic activity may be susceptible of mechanization," Woodward wrote in 1963, "the creative aspects of synthetic design will not!"[43]

A corollary to Woodward's account of synthetic planning was that the subject was teachable only by example. Woodward flaunted the difficulty of his courses on natural products chemistry – in one course announcement proof, he replaced a reference to students who "plan to register" with "*dare* to register" – and celebrated synthetic design as a matter of taste rather than method, "suited only for the special delectation of the

[40] Woodward, "Art and Science" (ref. 37), 31–36.

[41] Ibid., 36.

[42] Ibid., 38.

[43] Ibid., 41.

initiated."[44] In 1960, Woodward was named to a professorship that relieved him of classroom instruction obligations. The responsibility for teaching Harvard students how to plan organic syntheses fell to E. J. Corey.

Corey received his PhD from MIT (also Woodward's alma mater), where he contributed to John Sheehan's longstanding project to synthesize penicillin.[45] At MIT, he also studied with John Roberts, the pioneer of nuclear magnetic resonance spectroscopy; Corey made use of physical analytical methods from the beginning of his career. In his nine years at the University of Illinois, Corey's teaching and most of his research focused on theoretical and physical organic chemistry, not synthesis. In 1960, Corey joined Harvard's chemistry department, where he was charged with teaching organic synthesis for the first time. He covered the subject in a graduate course on special topics in organic chemistry, and he also had to figure out how to teach synthetic planning to undergraduates with little background in chemistry, as part of a new integrated science course for a select group of freshmen and sophomores. Corey observed that most instruction in synthetic planning proceeded case-by-case through landmark syntheses, each an illustration of the distinctive creative thinking that a chemist brought to bear on the challenge presented by a unique molecule.[46] Synthesis textbooks consisted of catalogs of useful reactions without any

[44] R. B. Woodward, Papers of Robert Burns Woodward, Course announcement, 24 Sept 1957, HUGFP 68.8, Box 7, Harvard University Archives; Woodward, "Art and Science" (ref. 37), 36.

[45] Biographical sources include Charles R. Allerson, "Elias J. Corey," in *Nobel Laureates in Chemistry, 1901-1992* (Washington: American Chemical Society, 1993), 750–58; A. Maureen Rouhi, "Above and Beyond Organic Synthesis," *Chemical & Engineering News* 82, no. 13 (2004): 37–41; E. J. Corey, "Impossible Dreams," *Journal of Organic Chemistry* 69, no. 9 (2004): 2917–19.

[46] Author interview with E. J. Corey, August 6, 2013 (recording in author's possession).

instruction on how to assemble them into a synthetic plan.[47] This was a pedagogy in keeping with Woodward's account of the field.

Corey felt that such an approach would plainly not do for his undergraduates, and that the new course presented an opportunity to experiment with a different approach. In order to teach students as much synthetic planning as possible in a brief period of time, he sought to organize his course in a general, systematic manner, drawing together insights from various different landmark syntheses. Rather than treating each synthetic plan as distinctive and original, Corey looked for recurring patterns, aiming to distill, condense, and organize this material to pack it into the few classroom hours at his disposal. The novelty of Corey's approach to synthetic design lay in his decision, born of necessity, not to treat each synthetic plan as *prima facie* novel.

New approaches to teaching may yield new ideas about what is taught.[48] Pedagogical exigencies have played a central role in the development of some of the best known conceptual developments in modern chemistry. The periodic system had its origins in the challenge of classifying diverse material in a pedagogically useful fashion.[49] So did retrosynthetic analysis. For Mendeleev, the materials in question were the properties of the diverse chemical elements; for Corey, they were the rapidly accumulating achievements in the synthesis of natural products. Pedagogy has also provided fertile ground for the development of novel methods marginalized by established research traditions – structural

[47] E.g. Romeo B. Wagner and Harry D. Zook, *Synthetic Organic Chemistry* (New York: Wiley, 1953); Vartkes Migrdichian, *Organic Synthesis* (New York: Reinhold, 1957).

[48] David Kaiser, ed., *Pedagogy and the Practice of Science: Historical and Contemporary Perspectives* (Cambridge, MA: MIT Press, 2005).

[49] Michael D. Gordin, *A Well-Ordered Thing: Dmitrii Mendeleev and the Shadow of the Periodic Table* (New York: Basic Books, 2004), 15–45.

organic chemistry in mid-nineteenth century France, quantum interpretations of steric hindrance in Nazi Germany, or molecular-orbital theory in mid-twentieth century Britain and America, for example.[50] Corey's pedagogical device soon became a methodological project of this sort.

General methods

In a 1964 article, Corey brought his approach to teaching synthetic planning into an account of his research. The article detailed the total synthesis of longifolene, a constituent of pine resin whose bridged polycyclic structure (multiple overlapping carbon rings) made it an especially interesting synthetic target. In this article, submitted the summer after Corey first taught his undergraduate synthesis course, Corey and his co-authors highlighted not the unique features of their synthetic plan – the typical point of emphasis in such discussions – but the general considerations that governed it. They explained that the synthesis of a bridged polycyclic compound such as longifolene should begin with "an exhaustive analysis of the topological properties of the carbon network." They described such an analysis in general terms: identify atoms contained in more than one ring, then develop sets of precursors formed by disconnecting the ring-forming bonds. Corey and his collaborators then presented the plan for their synthesis of longifolene – including several directions they chose not to pursue in the laboratory – as an application of such an analysis.

This account was exceptional in three ways. First, Corey did not just present his synthetic plan but described his process of synthetic planning, thus presenting synthetic

[50] José Ramon Bertomeu-Sánchez, Antonio Garcia-Belmar, and Bernadette Bensaude-Vincent, "Looking for an Order of Things: Textbooks and Chemical Classifications in Nineteenth Century France," *Ambix* 49, no. 3 (2002): 227–50; Johnson, "German Quantum Chemists" (ref. 28)**;** Buhm Soon Park, "In the Context of Pedagogy: Teaching Strategy and Theory Change in Quantum Chemistry," in Kaiser, *Pedagogy and the Practice of Science* (ref. 48), 287–322.

design as an activity worthy of description and analysis. Second, he presented synthetic planning as an application of general methods to a particular case. Third, he accounted for synthetic planning in its own terms, referring to objects (atoms contained in more than one ring, ring-forming bonds) and processes (disconnection) defined by their significance for synthetic design, not for laboratory manipulations involving material substances.[51]

Corey fleshed out this approach to synthetic design in a 1966 keynote address, in which he presented synthetic planning as a research topic in its own right.[52] By singling out two of Woodward's best-known reviews as "superb general accounts of Synthesis,"[53] Corey framed his lecture as a successor to Woodward's treatments of the topic. For Corey, as for Woodward, synthetic planning was the intellectual center of synthesis, and synthesis was a wellspring of novel chemistry.[54] However, Corey's title for the lecture, "General Methods for the Construction of Complex Molecules," signaled a break with Woodward's approach.

In Corey's view, synthetic design was not irreducibly idiosyncratic, but the expression of general methods. Against the prevailing view that synthetic planning was "tenuously hypothetical and is mainly a function of the unique circumstances in each particular case," Corey contended that the diversity of recent achievements in synthesis provided ample evidence for putting the thinking involved in synthetic design into a general form. Such an account was "prerequisite to a deeper comprehension of Synthesis and the methodologies which are fundamental to it," Corey asserted, "and it is likely to be a

[51] E. J. Corey et al., "Total Synthesis of Longifolene," *Journal of the American Chemical Society* 86, no. 3 (1964): 478–85, on 479, 482.

[52] The lecture was published the following year as E. J. Corey, "General Methods for the Construction of Complex Molecules," *Pure and Applied Chemistry* 14, no. 1 (1967): 19–38.

[53] Ibid., 36.

[54] Ibid., 31.

keystone in the rational development of Synthesis to still higher forms."[55] As an example of such "higher forms" of synthesis, Corey mentioned computing, with which he had little experience but whose scientific potential intrigued him.[56] A general definition of the intricacies of synthetic planning, he argued, would be a precondition for "any technique for the automatic generation of synthetic schemes by a computer."

Corey thus began his method-making project from the premise that synthetic design *already had* "fundamental methodologies." As Corey presented it, his aim was not to devise but to describe general methods for synthetic planning, working inductively from successful synthetic plans for complex natural products.[57] A general account of how chemists behaved in the course of synthetic design, Corey suggested, would provide the basis for formulating novel ways to make chemicals behave in the course of syntheses. The prospect of computerizing this process provided further justification that his method mattered.

What was this method? Corey described synthetic design as a process made up of twelve distinct steps (*Fig. 1*). This segmentation, Corey asserted, provided a basis for assembling a general account of synthetic design, including "many more steps of analysis… than have customarily been used" in the planning of any given synthesis.[58] In order to capture the serpentine path typical of skillful synthetic planning, Corey noted that his schema had to provide for interactions among all of its steps.[59] He did so by including an

[55] Ibid., 19.

[56] Corey interview (ref. 46).

[57] Corey, "General Methods" (ref. 52), 36.

[58] Ibid., 29.

[59] Woodward, "Art and Science" (ref. 37), 41; Corey, "General Methods" (ref. 52), 20.

"operational loop" within the planning process, accounting for the return to earlier stages as new ideas arose (*Fig. 1, steps 7 and 11*).[60]

1. Simplification of problem.
2. Systematic recognition of synthons.
3. Generation of equivalent and modified synthons.
4. Addition of control synthons.
5. Systematic disconnection of synthons.
6. Formulation of the possible synthetic transformations which reform the starting structure from the derived intermediate(s).
7. Repetition of items 1–6 for each intermediate and each sequence (parallel sequences may be generated), including previously generated intermediates.
8. Generation of intermediates until the required starting point is reached.
9. Removal of inconsistencies.
10. Identification of unresolved problems.
11. Repetition of items 1–10 to generate alternative schemes.
12. Assignment of merit.

Figure 1. Corey's general schema for synthetic planning.[61]

Here, Corey invoked computing again, this time as a model for articulating the thinking involved in synthetic planning as an iterative schema of discrete operations. His twelve steps bore "a vague resemblance to a computer programme," he wrote.[62] It was a passing reference, but it suggested that computing was beginning to play a role in the method-making project. Corey's understanding of computer programs, and of their family resemblance to human thinking, served as an analytical resource for working out his method and a rhetorical resource for promoting it.

A crucial aspect of synthetic planning, Corey emphasized, was the distinctive kind of chemical object that the experienced chemist reckoned with during this process. Organic chemists of the mid-twentieth century predicted and accounted for the behavior of organic

[60] Ibid., 27.

[61] Ibid., 28–29.

[62] Ibid., 29.

compounds in terms of several well-established kinds of chemical entities: atoms, chemically-individual building blocks; bonds, direct links between atoms; electrons, particles whose behavior accounted for the formation, stability, and breaking of bonds; and functional groups, patterns of atoms and bonds associated with characteristic properties and reactions. However, Corey was out to account for the behavior of synthetic chemists, not the behavior of chemical compounds. "The consideration of a molecule as a collection of the constituent atoms is perfectly definite," Corey observed, "but hardly useful in the design of a synthesis." Instead, "the synthetic chemist has learned by experience to recognize within a target molecule certain units which can be synthesized, modified, or joined by known or conceivable synthetic operations."[63] Corey coined the term *synthon* to refer to these units of synthetic planning. (*Fig. 2*)

[63] Ibid., 22.

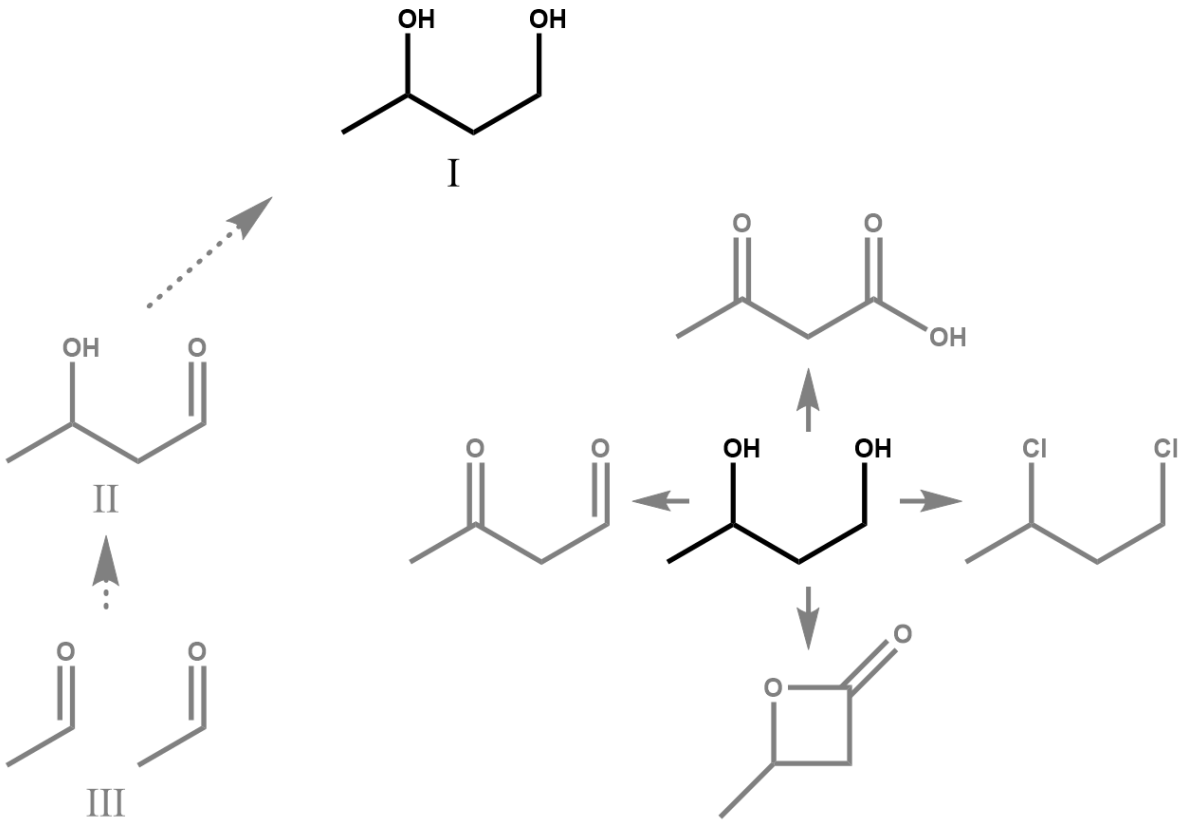

Figure 2. Considered as functional groups, the two –OH subunits in **I** indicate that this compound will display the characteristic chemical properties of alcohols. Under certain conditions, the compound *would* react to form the various products shown at bottom right. Considered as a synthon, the two –OH subunits in **I** were a structural signature suggesting that the compound *could* be produced from **III** via **II**, as shown at bottom left, by means of the well-known aldol reaction. Corey cited this example as an illustration of the synthon concept.[64]

Methodology begot ontology. Corey went looking for method, and found that he had to anatomize the molecular world in a new way, in terms of entities linked to synthetic possibilities rather than material properties.[65] Woodward had described the recognition of such possibilities as a creative act of inspiration that eluded generalization. Corey, in

[64] Ibid., 22–23. Unlabeled vertices in structural formulas represent carbon atoms, and hydrogen atoms bound to carbon are not shown.

[65] On the history of such paper chemical entities vis-à-vis computing, see Evan Hepler-Smith, "Paper Chemistry: François Dagognet and the Chemical Graph" forthcoming in *Ambix* (2018).

contrast, argued that "intellectual processes such as the recognition and use of synthons require considerable ability and knowledge; here, too, genius and originality find ample opportunity for expression."[66] As Corey presented it, skillful synthetic planning was an intellectual achievement of the highest order, *and* it could be studied, generalized, and made subject to method.

Method and the machine

At the end of his 1966 lecture, Corey expressed the hope that "a much more systematic, rigorous and complete account of the logic of Synthesis" was soon to come.[67] As the lecture's references to computing suggested, Corey had already conceived a plan to refine and extend his method for synthetic design: developing a computer program for carrying it out. Three years later, Corey and postdoctoral fellow Todd Wipke introduced readers of *Science* to OCSS, "Organic Chemical Simulation of Synthesis," subsequently renamed LHASA, "Logic and Heuristics Applied to Synthetic Analysis."[68] Corey and Wipke framed the program as an outgrowth of Corey's stepwise definition of synthetic planning, which both suggested the possibility of applying computers to the problem and, if the method was to be followed exhaustively, appeared to require it.[69] But LHASA was not just an application of "retrosynthetic analysis," as Corey christened his method in a 1971 publication on the program. LHASA group members also developed central aspects of this method as they built their computer-based tool. In navigating the constraints and

[66] Corey, "General Methods" (ref. 52), 30.

[67] Ibid., 36.

[68] E. J. Corey and W. Todd Wipke, "Computer-Assisted Design of Complex Organic Syntheses," *Science* 166, no. 3902 (1969): 178–92. During the 1940s, the Hungarian-American mathematician George Pólya popularized "heuristics" as a term for problem-solving methods.

[69] Ibid., 181.

affordances of machines, Corey and his collaborators worked out a distinctive chemistry of synthetic design, in which tractability to computing and legibility to human chemists went hand in hand.[70]

By the late 1960s, members of interdisciplinary "knowledge communities" and chemical subfields intersecting organic chemistry were applying computers to storing and retrieving chemical information, performing complex quantum chemical calculations, and processing data from analytical instruments.[71] While such projects aimed to generate results legible to structural organic chemists, they relied on method-makers trained in physics, applied mathematics, and engineering. The Stanford-based DENDRAL project exemplified this relationship between chemical computing and the problems and methods of other fields.[72] DENDRAL, a collaboration led by the molecular biologist Joshua Lederberg, the computer scientist Ed Feigenbaum, and the chemist Carl Djerassi, brought together chemical structure determination and the emerging field of artificial intelligence. Lederberg wanted to build a computer program capable of inferring the structure of a compound from mass spectrometry data. Djerassi was a leading expert in the

[70] Researchers at the Soviet Institute of Scientific and Technical Information (VINITI) had previously explored the prospect of applying computers to synthetic planning in the context of work on machine languages and information retrieval; L. L. Gutenmakher and G. E. Vleduts, "The Prospects for the Utilization of Informational-Logical Machines in Chemistry (USSR)," *Journal of the Association for Computing Machinery* 8, no. 2 (April 1961): 240–51; G. É. Vléduts and V. K. Finn, "Creating a Machine Language for Organic Chemistry," *Information Storage and Retrieval* 1, no. 2–3 (July 1963): 101–16; G. É. Vléduts, "Concerning One System of Classification and Codification of Organic Reactions," *Information Storage and Retrieval* 1, no. 2–3 (July 1963): 117–46.

[71] On information retrieval, see Leah Rae McEwen and Robert E. Buntrock, eds., *The Future of the History of Chemical Information* (New York: Oxford University Press, 2015). On calculation: Ann Johnson, "Modeling Molecules: Computational Nanotechnology as a Knowledge Community," *Perspectives on Science* 17, no. 2 (2009): 144–73. On analytical instruments: Reinhardt, *Shifting and Rearranging* (ref. 12).

[72] Robert K. Lindsay, *Applications of Artificial Intelligence for Organic Chemistry: The DENDRAL Project* (New York: McGraw-Hill, 1980); Joseph Adam November, *Biomedical Computing: Digitizing Life in the United States* (Baltimore: Johns Hopkins University Press, 2012), 237–68; Reinhardt, *Shifting and Rearranging* (ref. 12), 289–302.

interpretation of mass spectra; Feigenbaum saw structure determination as an ideal domain for developing his expert systems approach to artificial intelligence. For Lederberg and Feigenbaum, DENDRAL amounted to a series of experiments and models probing the possibilities of artificial intelligence and its relationship to human cognition; from Djerassi's perspective, the program was a tool for exhaustive data analysis beyond the capacity of human chemists.[73] Insofar as it dealt with human cognition, the project addressed reasoning in general, not chemical reasoning. Insofar as it addressed chemistry, DENDRAL drew inferences in a way that human chemists supposedly couldn't.

In contrast to such interdisciplinary projects to use computers to apply methods from other fields to chemical problems, Corey situated the LHASA project squarely within natural products synthesis. While Corey consulted computing experts, including computer graphics pioneer Ivan Sutherland, and made use of resources at Boston-area firms and in other Harvard departments, including Sutherland's ARPA-funded PDP-1 computer, the core members of the project were synthetic organic chemists who knew something about computers (or were willing to learn).[74] Building such a team was easier said than done. Corey was unable to persuade current graduate students to take time away from the prestigious, demanding work of natural products synthesis to build a computer program; the quixotic project seemed to offer little opportunity for professional advancement within

[73] Carl Djerassi, *The Pill, Pygmy Chimps, and Degas' Horse: The Autobiography of Carl Djerassi* (New York: Basic Books, 1992), 141–42.

[74] Members of the LHASA group consulted with Sutherland and Thomas Cheatham (an expert on data structures and programming languages) at Harvard, with the staff of the Boston-area firms Digital Equipment Corp. and Bolt, Beranek & Newman, and with specialists on computer representation of chemical structure at Chemical Abstracts Service. Pensak interview (ref. 38); Corey and Wipke, "Computer-Assisted Design" (ref. 68), 191n33.

academic chemistry.[75] Corey's first collaborator on the project, Wipke, came to Harvard as a postdoctoral fellow in 1967 after completing a PhD in organic chemistry and a stint in automatic data processing and analysis for the US Army.[76] By the time Wipke departed in 1969 for an assistant professorship at Princeton, Corey had begun recruiting a small group of graduate and postdoctoral synthetic organic chemists whose career ambitions lay outside total synthesis.[77]

The LHASA group organized their program around the general steps of synthetic planning that Corey had outlined in his 1966 lecture. They did not try to automate the entire process. Instead, they divided synthetic design into "logic centered" steps, each of which was the subject of a program module, and "information centered" steps, which were left to the judgment of the chemist-user. Corey and his collaborators reassembled these steps to form an "'interactive' system" in which some stages were carried out by machine processes and others by human guidance.[78]

The chemist-user began by drawing the structural formula for the molecule to be synthesized using a RAND tablet and a CRT display (*Fig. 3*).[79] The program stored a

[75] Corey interview (ref. 46).

[76] William Todd Wipke, "Biographical Outline," 1978, Stanford University Libraries, Edward A. Feigenbaum Papers, Accession 1986-052, Box 43, Folder 39, <https://saltworks.stanford.edu/catalog/druid:fq785yn3952>. On the role of the US military as a sponsor of information technology research and development during the Cold War, see James W. Cortada, *All the Facts: A History of Information in the United States since 1870* (New York: Oxford University Press, 2016), 243–54.

[77] Author interview with Harry Orf, 7 Aug 2013 (recording in author's possession); E. J. Corey and David A. Pensak, "LHASA – Logic and Heuristics Applied to Synthetic Analysis," in *Computer-Assisted Organic Synthesis*, ed. W. Todd Wipke and William Jeffrey Howe (Washington: American Chemical Society, 1977), 1–32, on 31.

[78] Corey and Wipke, "Computer-Assisted Design" (ref. 68), 179–82. On the rhetoric of "interactivity," see Andrew Utterson, "Early Visions of Interactivity: The In(put)s and Out(put)s of Real-Time Computing," *Leonardo* 46, no. 1 (2012): 67–72.

[79] On contemporaneous molecular graphics research at MIT, see Eric Francoeur, "Cyrus Levinthal, the Kluge and the Origins of Interactive Molecular Graphics," *Endeavour* 26, no. 4 (2002): 127–31.

representation of this drawing in the form of tables of atoms and bonds. Next, the program analyzed these tables to generate lists of synthetically significant structural features. LHASA then prompted the chemist-user to select one of several heuristic "strategies," which corresponded to the general approaches to synthetic design that Corey taught in his synthesis courses. Based on the selected strategy, the program generated a goal: a criterion for the incremental simplification of the structure of the target molecule.[80] The program then searched through a knowledge base of synthetic transformations for every possible way to work backwards one step in a manner that satisfied that goal. LHASA ranked and displayed the resulting set of precursor compounds one synthetic step removed from the target, eliminating structures that violated rules of chemical valence.[81] The program then prompted the chemist to choose one or more of these precursors and a strategy to apply to *them*, generating a set of second-order precursors, and so on. In this manner, the program and its chemist-user worked together to generate what the LHASA group called a "synthetic tree": a branching collection of paths extending backwards from the target molecule toward simpler compounds that could serve as starting materials (*Fig. 4*). By channeling the production of this tree through heuristics and along paths selected by the chemist-user, the program avoided a combinatorial explosion of conceivable synthetic routes and took advantage of the chemist's capacity to recognize promising synthetic possibilities.

[80] Corey and Wipke defined a "heuristic" as "a 'rule-of-thumb' which may lead by a shortcut to the solution of a problem or may lead to a blind alley," citing the research of Feigenbaum and his mentor Allan Newell; Corey and Wipke, "Computer-Assisted Design" (ref. 68), 191n8, 192n27; E. J. Corey, "Computer-Assisted Analysis of Complex Synthetic Problems," *Quarterly Review of the Chemical Society* 25, no. 4 (1971): 455–82, on 456n1.

[81] Carbon formed four bonds, hydrogen formed one, oxygen formed two bonds and had two "lone pairs" of unbonded electrons, and so on.

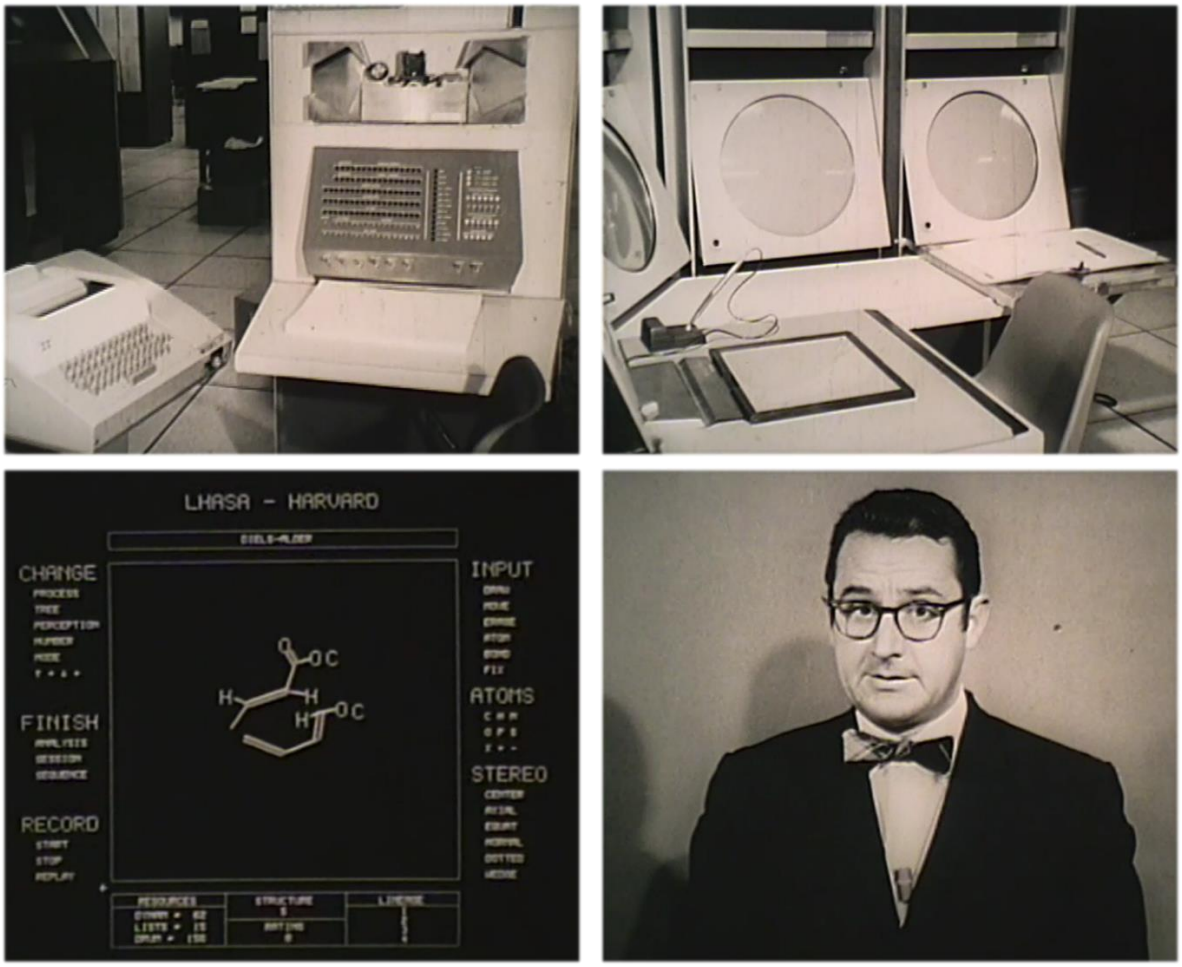

Figure 3. Four stills from a 1973 film documenting the LHASA project. Top left: PDP-1 computer and teletype input keyboard. Top right: CRT displays and RAND tablet. Bottom left: display during program use. Bottom right: Corey.[82]

[82] Jeff Howe, "LHASA: A computer program to assist in the design of synthetic routes to complex organic molecules" (1973), 16mm film in the personal collection of Gerald Lotto. (Digital copy in author's possession.)

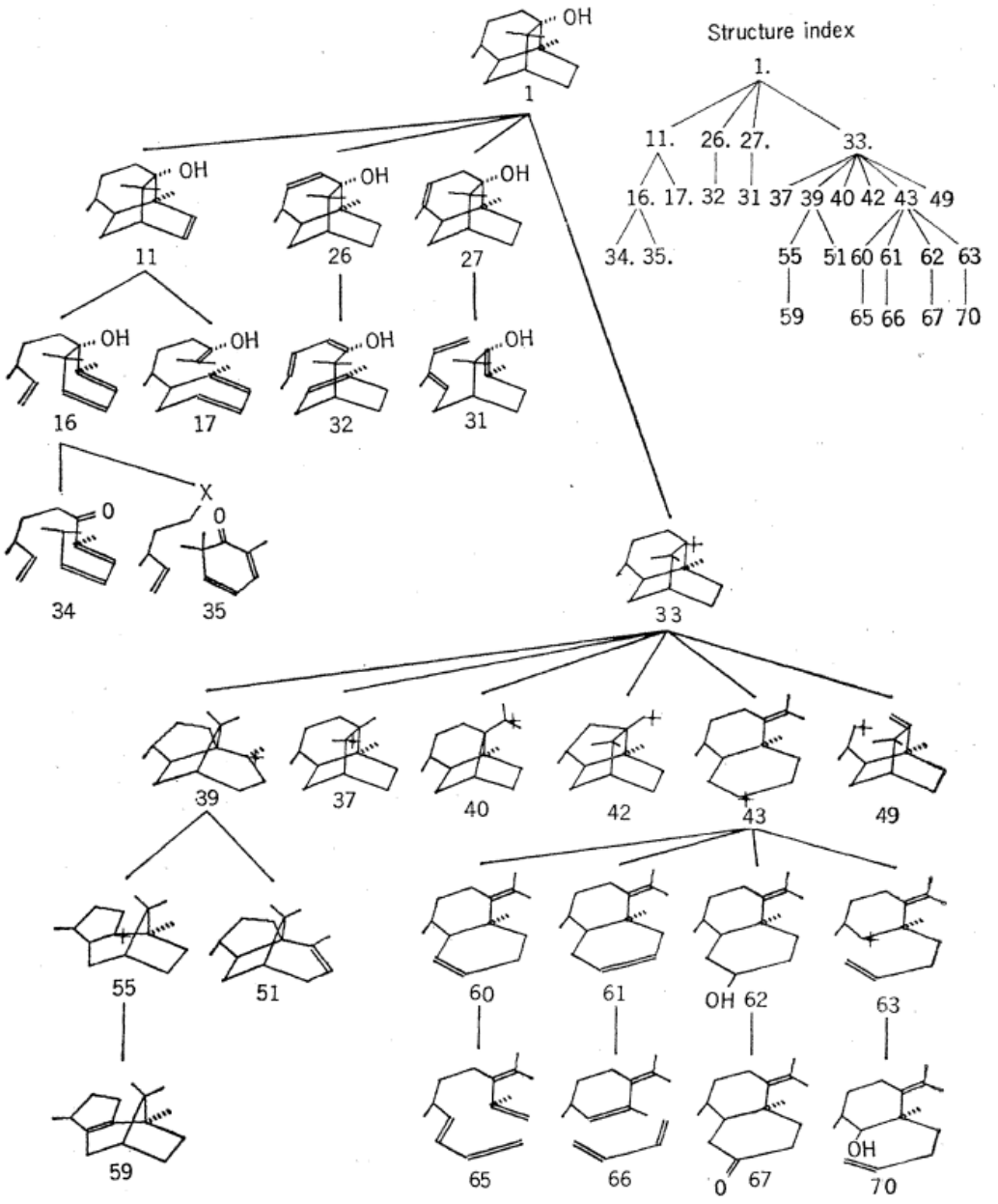

Figure 4. A "synthetic tree" (also referred to as a "synthesis tree") for the compound patchouli alcohol (1), generated by LHASA.[83]

Corey referred to this iterative process of working backwards as a *retrosynthetic* approach to synthetic design. He made clear that while his articulation of the method was novel, the approach itself was not. Corey identified retrosynthetic analysis as the starting

[83] Corey and Wipke, "Computer-Assisted Design" (ref. 68), 191.

point of "the most common approach of the chemist to problem solving" in the planning of complex syntheses.[84] Indeed, a widely-used 1969 textbook of organic synthesis described working backwards as "the cardinal rule of synthesis."[85] However, other chemists had not attached any special chemical significance to this approach. Corey cast working backwards in synthetic design as a distinctive way of thinking about chemicals and chemistry, in a manner intimately tied to its computer implementation.

He did so in three ways. First, Corey defined retrosynthetic analysis in a way that entailed the use of a computer. The strength of the method, Corey noted, was also a practical weakness: "a complete, logic-centered synthetic analysis of a complex organic structure often requires so much time, even of the most skilled chemist, as to endanger or remove the feasibility of this approach."[86] Corey presented the computer, with its brute-force capacity to catalog structural features and test out synthetic steps, as a tool that could make his method feasible. Corey did not mention this aspect of his approach to synthetic design during his 1966 lecture; such a demanding method may only have been conceivable in light of its automation.

Second, the LHASA group redefined synthetic planning in terms of novel chemical objects that emerged at the interfaces of their computer program. Retrosynthetic analysis did not operate on the molecules, functional groups, and reactions of organic chemistry, but on these new kinds of fundamental entities and manipulations. Discussing how they set up

[84] Corey, "Computer-Assisted Analysis" (ref. 80), 459.

[85] Robert E. Ireland, *Organic Synthesis* (Englewood Cliffs, NJ: Prentice-Hall, 1969), 17.

[86] Corey and Wipke, "Computer-Assisted Design" (ref. 68), 181.

the program's knowledge base of synthetic transformations, Corey, postdoctoral fellow Richard Cramer, and graduate student Jeff Howe wrote,

> A variety of rational schemes for creating families of synthetic reactions already exists. However, most of these depend on properties of the *reactants*, and as such they are irrelevant to a computer program which analyzes the features of a *target or product molecule* in order to generate appropriate starting materials.[87]

The LHASA group members instead took synthons, the category that Corey first described in his 1966 lecture, as the basis for their classification. In incorporating synthons into the computer program, they defined the term more precisely: "key structural features of target molecules [associated] with specific structural features that might be created by the operation of particular synthetic reactions."[88]

Such definitions had a cascading effect. In order to follow the general strategies that were the centerpiece of Corey's approach to synthetic design, the program needed to be able to pass data back and forth between different modules and subroutines. Corey and his colleagues accordingly extended their synthon-based classification throughout the program, redefining *all* of the elements of synthetic design. They explained:

> The resulting organization of synthetic chemistry is based on the reverse formulation of synthetic reactions. This organization, with its reverse-synthetic or retrosynthetic focus, must be expressed explicitly and consistently in order to avoid intolerable confusion, and this has necessitated the introduction of some new nomenclature.[89]

[87] E. J. Corey, Richard D. Cramer, and W. Jeffrey Howe, "Computer-Assisted Synthetic Analysis for Complex Molecules. Methods and Procedures for Machine Generation of Synthetic Intermediates," *Journal of the American Chemical Society* 94, no. 2 (1972): 440–59, on 440. My emphasis.

[88] Ibid., 441.

[89] Ibid.

In this new terminology, the overall direction of analysis was *antithetic* or *retrosynthetic*. The structural change constituting each retrosynthetic step was a *transform*. The symbol representing a retrosynthetic step was a two-lined arrow. The LHASA group provided a point-by-point comparison of their vocabulary with standard chemical terminology, articulating what distinguished thinking about synthetic design from thinking about chemical reactions (*Fig. 5*). The adoption of the two-lined arrow (pointing forward) rather than simply a backwards arrow most clearly expressed this point. Retrosynthetic analysis was not just chemistry in reverse, but a distinctive way of thinking backwards.

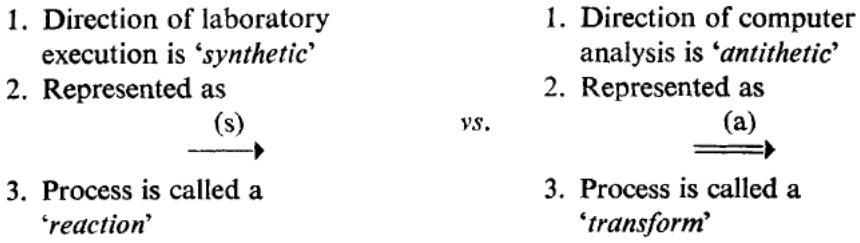

Figure 5. The vocabulary of retrosynthetic analysis (right) compared with the vocabulary of synthetic organic chemistry (left).[90]

Third, the LHASA group framed the development of LHASA's strategy module, the heart of the computer program, as a sophisticated form of creative chemical thinking. The work involved in encoding these heuristic strategies – designing intricate flow charts and writing machine-executable statements and queries – did not immediately look like an exercise in chemical problem-solving. However, what the singular imaginative insight was to Woodward's account of synthesis, the articulation of a heuristic strategy as a machine algorithm was to Corey's: a way of thinking central to effective synthetic planning that advanced chemistry in general.

[90] Corey, "Computer-Assisted Analysis" (ref. 80), 457.

A particularly well-documented example of this process is the Diels-Alder search strategy, developed by Corey and graduate students Howe and David Pensak. The Diels-Alder reaction was (and is) a versatile and precise means of creating rings of six carbon atoms, one of the most common structural units of complex natural products (*Fig. 6*). The guiding heuristic of the Diels-Alder search strategy was the goal of working backwards from a synthetic target to a precursor that could be the product of a Diels-Alder reaction. The program module consisted of a collection of subroutines that searched the program's chemical knowledge base for a series of transforms (retrosynthetic steps) that could accomplish this goal. Thus, when the chemist entered the target compound **A** and selected this strategy, the program generated four distinct series of precursor compounds, culminating in the application of the Diels-Alder transform to compounds **H, K, O,** and **R** (*Fig. 7*). While compound **A** was a hypothetical synthetic target, the LHASA group often illustrated the application of new strategy modules with novel and promising LHASA-generated routes to complex natural products such as prostaglandin $F_{2\alpha}$, the subject of one of Corey's most celebrated syntheses.[91]

[91] E. J. Corey et al., "Computer-Assisted Synthetic Analysis. Long-Range Search Procedures for Antithetic Simplification of Complex Targets by Application of the Halolactonization Transform," *Journal of Chemical Information and Computer Sciences* 20, no. 4 (1980): 221–30, on 226.

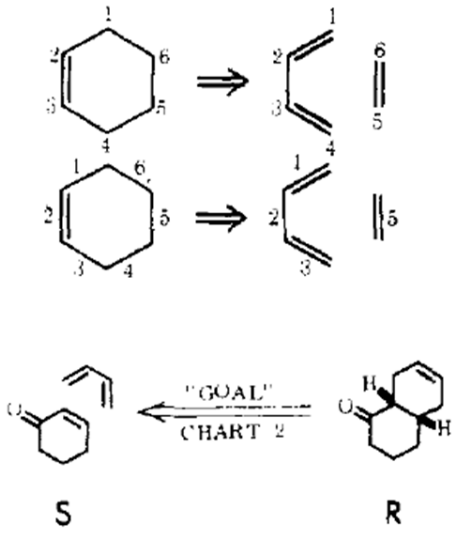

Figure 6. At top, the general schema for the Diels-Alder transform, with atoms and bonds labeled for reference. At bottom, an application of the transform to a target **R** containing the necessary six-membered ring with one double bond, generating the pair of precursors **S**.[92]

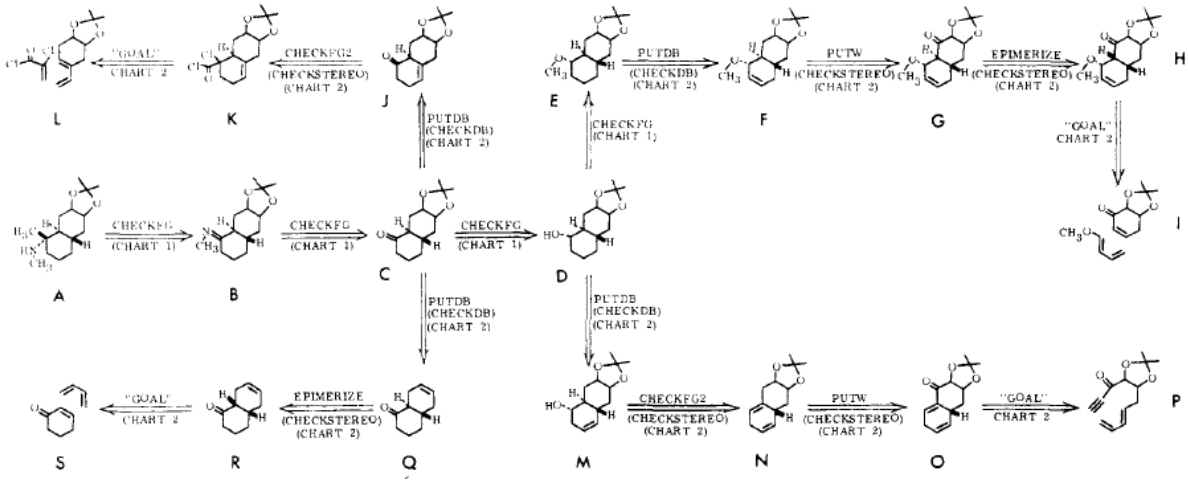

Figure 7. A synthetic tree generated by the application of the Diels-Alder search strategy to target **A,** leading to the application of Diels-Alder transform to arrive at four possible pairs of simpler precursors (**I**, **L**, **P**, and **S**).[93]

[92] E. J. Corey, W. Jeffrey Howe, and David A. Pensak, "Computer-Assisted Synthetic Analysis. Methods for Machine Generation of Synthetic Intermediates Involving Multistep Look-Ahead," *Journal of the American Chemical Society* 96, no. 25 (1974): 7724–37, on 7727, 7736.

[93] Ibid., 7736.

To add this strategy to the program's repertoire, the LHASA chemists learned all that they could about how Corey and his group members used the Diels-Alder reaction in their synthesis projects. They then organized these observations into a set of chemical flow charts that directed a search for retrosynthetic pathways to a Diels-Alder product.[94] After the eleven intricate flow charts that constituted the Diels-Alder strategy were complete (*Fig. 8*), they had to be translated into instructions that the computer could execute. The LHASA group did so using CHMTRN, "a language based on 'chemical English' which has been designed to be computer-readable and intelligible to a chemist with little or no programming experience."[95] Along with the legibility of CHMTRN, Corey, Pensak and Howe emphasized that the LHASA program treated algorithms expressed in this language as data, not code. The distinction between these categories was far from settled, but the LHASA authors invoked it to argue that "[n]o 'programming' is involved in the addition of new search patterns." Rather, the "time-consuming factor in the addition of new ring transforms to LHASA is the detailed study and analysis that is necessary for the development of an effective search pattern for a very broad range of organic structures."[96] In this way, Corey and his co-authors located the work of adding new strategies to LHASA within the domain of synthetic organic chemistry, not computer programming.

[94] As one group member put it, "our job was to think about how chemists thought about these things, and see if we could kind of codify that"; Orf interview (ref. 77).

[95] Corey, Howe, and Pensak, "Computer-Assisted Synthetic Analysis" (ref. 92), 7727, 7730.

[96] Ibid., 7733–34.

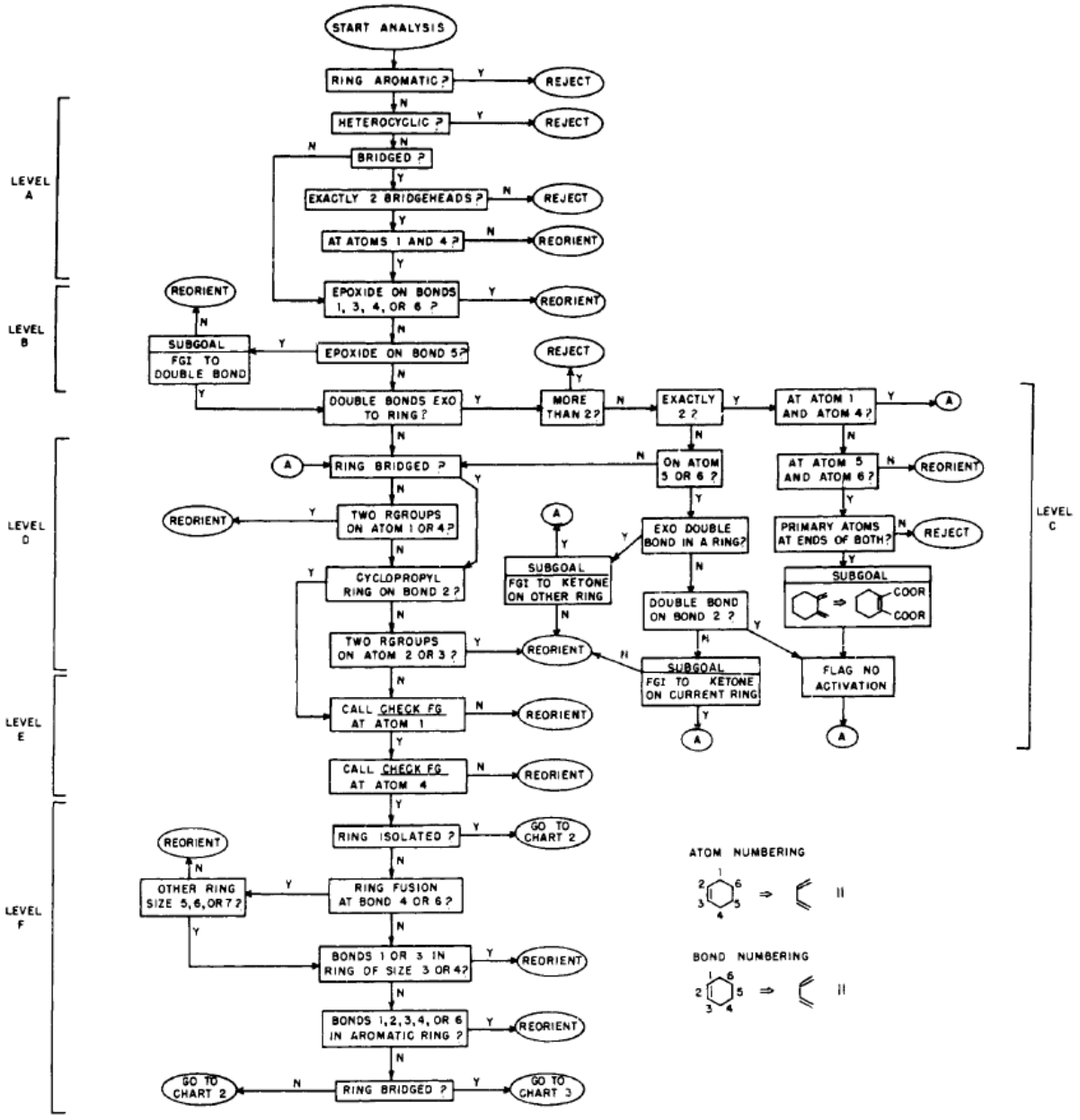

Figure 8. A simplified version of the first of eleven flow charts representing the Diels-Alder search strategy. Each of the six levels of this chart addressed an aspect of the target compound's structure that could enable or impede the retrosynthetic application of the Diels-Alder transform (that is, a structural feature that made the compound a more or a less favorable *product* of a Diels-Alder *reaction*). Once the LHASA-user entered a target structure and selected the Diels-Alder strategy, all the operations represented in this and the other Diels-Alder flow charts were carried out by the computer.[97]

[97] Ibid., 7727–29.

Building up LHASA's strategy module was no small task. Corey, Howe, and Pensak estimated that it would take up to a half-dozen chemists several weeks to design flow charts for each strategy.[98] Just in the category of ring-forming reactions like Diels-Alder, the LHASA authors mentioned fifteen more heuristics that could profitably be added to their program; they also noted a half-dozen other classes of synthetic reactions that could be likewise be the basis for useful strategies. The process of formalizing such strategies, in turn, could yield further generalizations about synthetic design, such as Corey's and Wipke's recognition that many strategies hinged on the relationship between a pair of structural features.[99] Corey presented all of this work not as an obstacle to the development of LHASA, but as an indication of its promise as an ongoing research project. After all, natural products synthesis itself demanded an enormous amount of time and effort. Corey and his co-authors placed their method-making on par with molecule-making, contending that developing flow charts and translating them into CHMTRN was "not less challenging and rewarding than the conception and execution of a specific synthesis."[100] Corey imagined teams of chemists at twenty to thirty top universities joining in this effort to build up LHASA's bag of heuristic tricks. Expanding the program in this way could provide an organizing logic for the entire field of synthetic organic chemistry—a crucial contribution, in Corey's view, because "messy disciplines die."[101]

The LHASA group could have presented their project differently, as an application or case study of artificial intelligence. This was how Lederberg and his collaborators

[98] Ibid., 7737n34.

[99] Corey and Wipke, "Computer-Assisted Design" (ref. 68), 187.

[100] Corey, Howe, and Pensak, "Computer-Assisted Synthetic Analysis" (ref. 92), 7736.

[101] Quoted in Irene Kiefer, "Choosing Chemical Routes," *Mosaic* 5, no. 4 (1974): 21–25, on 25.

conceived of DENDRAL, and indeed, the LHASA group graduate student Pensak dedicated the final pages of his dissertation to framing LHASA in this way.[102] But this was not the direction in which Corey guided the project. He did not aim to use his expertise in chemical reasoning to colonize artificial intelligence. Rather, he saw the computer as a means of advancing a program of chemical research dedicated to the "rational development of Synthesis to still higher forms."[103]

Method for all

Corey's vision of the broad-based adoption of LHASA as both tool and research program did not come to pass. This was not for lack of interest in computer-assisted synthetic design. During the 1970s and 1980s, numerous chemists, engineers, and computer scientists took an interest in the subject, contributing to a "flourishing science of the algorithms and heuristics of optimal synthesis design … from the foundation laid by Corey," as one review put it.[104] In 1976, a symposium on the topic showcased research at Carnegie Mellon, Stanford, Technische Universität München (Munich), Rutgers, UC Santa Cruz, the University of Toronto, and the chemical firms Merck, ICI, and Du Pont.[105] However, most new entrants (and some former LHASA group members, including Wipke) launched their own programs rather than contributing to Corey's. Despite a few promising demonstrations, few synthetic organic chemists ever used LHASA in their own synthetic planning. According to one chemist who did, most others found the learning curve too

[102] David Alan Pensak, "Computer-Aided Design of Complex Organic Syntheses" (PhD dissertation, Harvard University, 1973), 225–32.

[103] Corey, "General Methods" (ref. 52), 19.

[104] Malcolm Bersohn and Ashmeed Esack, "Computers and Organic Synthesis," *Chemical Reviews* 76, no. 2 (1976): 269–82, on 281.

[105] Wipke and Howe, *Computer-Assisted Organic Synthesis* (ref. 77).

steep, the program's knowledge base too limited, and the idea of sharing the treasured puzzle of synthetic planning with a machine anathema.[106]

By the 1980s, industrial and academic interest in computational chemistry shifted toward tasks for which machine methods offered more frequent and decisive advantages, especially chemical databases and search tools.[107] Programs spun off from LHASA filled new niches in chemical drawing and predictive toxicology.[108] When a new generation of computational chemists took up synthetic planning in the 2000s, they paid their respects to LHASA but consigned it to prematurity. Some even blamed the program for inadvertently discouraging research in the field. "After all," write the authors of a recent study, summing up the attitude of their predecessors, **"**if Corey failed, who should even try?"[109]

Evan as interest in LHASA was declining, an increasing number of chemists were describing their synthetic planning in terms of retrosynthetic analysis. Corey regularly pointed out the crucial role that retrosynthetic strategies played in the synthetic planning behind his group's syntheses, sometimes including a detailed discussion of a retrosynthetic analysis as a first stage in accounts of his syntheses.[110] (K. C. Nicolaou, a postdoctoral fellow with Corey during the 1970s who became a leading synthetic organic chemist, published

[106] Judson, *Knowledge-Based Expert Systems* (ref. 10), 29–31.

[107] Ibid.

[108] On chemical drawing, see David A. Evans, "History of the Harvard ChemDraw Project," *Angewandte Chemie International Edition* 53, no. 42 (2014): 11140–45. On predictive toxicology, see "About Us," *Lhasa Limited*, accessed March 2, 2017, <http://www.lhasalimited.org/about-us.htm>.

[109] Szymkuć et al., "Computer-Assisted Synthetic Planning" (ref. 9), 5919–20.

[110] E. J. Corey and Richard D. Balanson, "Total Synthesis of (+-)-Porantherine," *Journal of the American Chemical Society* 96, no. 20 (1974): 6516–17.

entire articles on his retrosynthetic analysis of synthetic targets.)[111] Corey also often cited particular synthetic plans as the basis for his formalization of general retrosynthetic strategies.[112] For Corey, this was the point: retrosynthetic analysis provided a means for making specific aspects of synthetic design general and accessible. For others, such as Corey's contemporary Samuel Danishefsky, it was a sign that retrosynthetic analysis had always been implicit in synthetic planning. But even as he called the method's novelty into question, Danishefsky devoted his prestigious Roger Adams Award essay to a discussion of retrosynthetic analysis.[113] Indeed, more than a quarter of the articles addressing total synthesis published in the *Journal of the American Chemical Society* since 1990 (the year of Corey's Nobel Prize) discuss retrosynthetic analysis.[114]

Retrosynthetic analysis became even more prominent in teaching. From early on, Corey and his colleagues developed LHASA with pedagogical applications in mind, and group members felt that their work on the program made them much better teachers of synthetic planning.[115] Clark Still, a postdoctoral fellow working with Wipke at Princeton, developed a paper template for teaching students to generate a LHASA-style synthetic tree

[111] E.g., K. C. Nicolaou, S. P. Seitz, and M. R. Pavia, "Synthesis of 16-Membered-Ring Macrolide Antibiotics. 3. Carbomycin B and Leucomycin A3: Retrosynthetic Studies," *Journal of the American Chemical Society* 103, no. 5 (1981): 1222–24.

[112] E. J. Corey, Alan K. Long, and Stewart D. Rubenstein, "Computer-Assisted Analysis in Organic Synthesis," *Science* 228, no. 4698 (1985): 408–18, on 418n9,12–18.

[113] Rebecca M. Wilson and Samuel J. Danishefsky, "Pattern Recognition in Retrosynthetic Analysis: Snapshots in Total Synthesis," *Journal of Organic Chemistry* 72, no. 12 (2007): 4293–4305.

[114] A September 4, 2017 search of pubs.acs.org for articles published in the *Journal of the American Chemical Society* since 1990 whose title includes "total synthesis" shows that 363 of 1303 (28%) include "retrosynthetic analysis" in their full text.

[115] Corey, Howe, and Pensak, "Computer-Assisted Synthetic Analysis" (ref. 92), 7737n34.

by hand.[116] Corey's graduate course in synthetic design, "Chem 115," was closely associated with the LHASA project. Corey selected heuristics to add to the program from the pool of strategies for synthetic design covered in Chem 115, and he reciprocally incorporated insights that emerged from the LHASA project into the course.[117] Students in the course encountered synthetic design through the distinctive principles, terminology, and notation developed to make retrosynthetic planning tractable to automation: transforms, strategies, the retrosynthetic arrow, and so on. Chem 115 took on legendary status among generations of Harvard chemists, some of whom sat in year after year; it became a major avenue for spreading the method of retrosynthetic analysis.[118]

The LHASA group also developed a simplified version of their program as an instructional tool. Undergraduate courses at Harvard and Tufts tasked students with following the program's workings as an interactive exemplar of the thinking involved in synthetic design.[119] Building on these efforts, in 1980, two instructors at Franklin and Marshall College organized an undergraduate seminar around the question of how to "teach a computer to do organic synthesis problems" according to Corey's retrosynthetic

[116] W. Clark Still, "Synthesis Sheets: An Aid to Synthetic Analysis," *Journal of Chemical Education* 50, no. 6 (1973): 378–79.

[117] For example, this was true of the strategies that Corey, Howe, and Pensak proposed for addition to LHASA in 1974. Corey, Howe, and Pensak, "Computer-Assisted Synthetic Analysis" (ref. 92), 7735; Rick Danheiser, Chem 115 course notes, 31 Oct – 14 Dec 1972 and 29 Sept 1973, personal papers of Rick Danheiser.

[118] Martin D. Burke and Gojko Lalic, "Teaching Target-Oriented and Diversity-Oriented Organic Synthesis at Harvard University," *Chemistry & Biology* 9, no. 5 (2002): 535–541, on 535–536; Rouhi, "Above and Beyond Organic Synthesis," 41. Corey students taught a version of the course at a National Science Foundation-sponsored summer school, further disseminating retrosynthetic analysis; Corey interview (ref. 46).

[119] H. W. Orf, "Computer-Assisted Instruction in Organic Synthesis," *Journal of Chemical Education* 52, no. 7 (1975): 464–67; Robert D. Stolow and Leo J. Joncas, "Computer-Assisted Teaching of Organic Synthesis," *Journal of Chemical Education* 57, no. 12 (1980): 868–73.

approach. In a report on the course, they concluded that "teaching a computer to synthesize molecules is a good way to teach some organic chemistry."[120]

Other educators felt that the association with computing had to be broken in order to make Corey's method accessible to a broader audience of chemists. Stephen Turner, a British pharmaceutical chemist, believed that LHASA heralded a momentous transformation of how synthetic planning could and would be done.[121] However, he felt that as it was described in the LHASA group's articles, retrosynthetic analysis was not suitable for "the 'everyday' chemist," who did not have access to a computer and, in Turner's view, did not think like one.[122] In a 1971 pamphlet and a 1976 textbook, Turner took up the distinctive terminology of retrosynthetic analysis but employed it in less precise ways than the LHASA group did. For example, Turner cast a synthon simply as "a portion of a molecule which is recognizably related to a simpler molecule." He treated disconnections not as a process guided by well-defined strategy but as an exploratory procedure to be applied to all bonds in a molecule.[123] To "enable a chemist, using only pencil and paper, to consider a specific synthesis in a general way," Turner elided the aspects of retrosynthetic analysis that struck him as necessary or feasible only for computer-based synthetic planning.[124]

---

[120] Bonnie Burns Sandel and Robert W. Solomon, "Integration of Major Computer Program Packages into Experimental Courses. 2: Organic Synthesis Design and the Computer," *Journal of Chemical Education* 58, no. 10 (1981): 798–800.

[121] Stephen Turner, *The Design of Organic Syntheses* (New York: Elsevier, 1976), 9.

[122] Ibid., vii, 48.

[123] Ibid., 53; Stephen Turner, *An Introduction to the Design of Organic Synthesis* (England: Koch-Light Laboratories Ltd., 1971), 9.

[124] Ibid., 2.

While Turner's pamphlet and textbook were pitched to professional chemists, University of Cambridge chemist Stuart Warren's *Organic Synthesis: The Disconnection Approach* brought retrosynthetic analysis into the undergraduate classroom.[125] Warren had encountered retrosynthetic analysis in Chem 115, which he audited during a postdoctoral fellowship at Harvard.[126] As an instructor at Cambridge, he adapted Corey's rigorous schema into a form he considered accessible to undergraduates. In 1978, Warren published a workbook based on this course, and in 1982, he published his textbook.[127] *The Disconnection Approach* was taken up broadly in undergraduate teaching and became a model for subsequent textbooks on synthetic planning.[128]

In *The Disconnection Approach*, Warren eschewed the conventional presentation of great achievements in synthesis. "This book is to help you design your own syntheses rather than tell you about those devised by others," it began. To this end, Warren elected to present "how this planning is done: to help you learn the *disconnection* or *synthon* approach to organic synthesis."[129] Warren said nothing about the origin of the method that he outlined, but his disconnection approach looked a lot like retrosynthetic analysis. He employed the two-lined retrosynthetic arrow and other distinctive terms coined by the LHASA group.[130] As the LHASA group distinguished transforms from strategy, so Warren addressed various classes of "disconnection" and various strategies in alternating

[125] Stuart G. Warren, *Organic Synthesis: The Disconnection Approach* (New York: Wiley, 1982).

[126] Corey interview (ref. 46).

[127] Stuart G. Warren, *Designing Organic Syntheses: A Programmed Introduction to the Synthon Approach* (New York: Wiley, 1978).

[128] Patrick T. Flaherty, "Book Review of Organic Synthesis: The Disconnection Approach. Second Edition," *Journal of Medicinal Chemistry* 52, no. 18 (2009): 5768–69.

[129] Warren, *Organic Synthesis: The Disconnection Approach* (ref. 125), xi, 1.

[130] Ibid., 15.

chapters.[131] Corey had always insisted that retrosynthetic analysis was a generalization of patterns of synthetic planning found in prominent syntheses of the recent past. Warren took this one step further. In his silence about the role of Corey and computers in the development of the "disconnection approach," Warren universalized retrosynthetic analysis as timeless chemical common sense.

In 1989, Corey published *The Logic of Chemical Synthesis*, co-written with postdoctoral fellow Xue-Min Cheng. The textbook presented retrosynthetic analysis as a fundamental chemical method, "prerequisite to expertise in synthetic design."[132] The influence of the computational context of the LHASA project upon this method was as evident as ever. Corey and Cheng described how chemists approached synthetic planning as if they were machines: "cycles of perception and logical analysis applied reiteratively to a target structure and to the 'data field' of chemistry lead to the development of concepts and ideas for solving a synthetic problem."[133] The book hewed faithfully to the (by then extensive) technical vocabulary of the LHASA project.[134] Readers searching for in-depth discussion of retrosynthetic strategies outlined in the textbook were referred to the publications of the LHASA group.[135] But whereas earlier publications framed the LHASA program as a chemical research project, Corey and Cheng described computer-assisted

[131] Ibid., vii–viii.

[132] E. J. Corey and Xue-Min Cheng, *The Logic of Chemical Synthesis* (New York: Wiley, 1989), preface. Reprinted in paperback in 1995, the text remains a standard reference.

[133] Ibid., 3.

[134] Ibid., 96–98.

[135] This included strategies that the LHASA group had studied but had not been able to add to their program. As the historian Michael Mahoney noted, "the enduring experience of the communities of computing has been the huge gap between what we can imagine computers doing and what we can actually make them do." Michael S. Mahoney, "The Histories of Computing(s)," *Interdisciplinary Science Reviews* 30, no. 2 (2005): 119–35, on 128.

synthetic planning as "a major challenge in the field of machine intelligence."[136] *The Logic of Chemical Synthesis* abstracted the method of retrosynthetic analysis from the project of computer-assisted synthetic planning, even as the method retained the shape of its machine embodiment.

Between Chem 115, the network of Corey group alumni, Warren's and Corey's textbooks, and Corey's 1990 Nobel Prize, by the 1990s retrosynthetic analysis was firmly established around the world as the dominant method for teaching synthetic design.[137] "No chemists today finish their university education without learning the step-by-step planning of simple model syntheses with the aid of Corey's synthon concepts," wrote a pair of German authors in 1995.[138] The American Chemical Society selected retrosynthetic analysis as one of the eleven conceptual topics that every introductory organic chemistry course should address.[139] A group of educators even applied "retrosynthetic analysis" analogically to the problem of chemical instruction itself, a half-joke that suggests how deeply ingrained the method has become.[140] In the 2008 second edition of his textbook, Warren equated his "disconnection approach" with retrosynthetic analysis and credited

[136] Corey and Cheng, *The Logic of Chemical Synthesis* (ref. 132), 25.

[137] Retrosynthetic analysis is not the only method of synthetic design presently in use. "Diversity-oriented synthesis," also known as "combinatorial chemistry," applies the systematic approach of retrosynthetic analysis to the production of a wide variety of related molecules. Stuart L. Schreiber, "Target-Oriented and Diversity-Oriented Organic Synthesis in Drug Discovery," *Science* 287, no. 5460 (2000): 1964–69. "Organocascade catalysis" replaces the many discrete steps of a retrosynthetic plan with transformations that follow one another spontaneously in the presence of a suitable catalyst. Eric N. Jacobsen and David W. C. MacMillan, "Organocatalysis," *Proceedings of the National Academy of Sciences* 107, no. 48 (2010): 20618–19.

[138] W. D. Ihlenfeldt and J. Gasteiger, "Computer-Assisted Planning of Organic Syntheses: The Second Generation of Programs," *Angewandte Chemie International Edition* 34, no. 23–24 (1995): 2613–33, on 2613.

[139] ACS Committee on Professional Training, "Organic Chemistry Supplement," March 2015, accessed March 2, 2017, <http://www.acs.org/content/acs/en/about/governance/committees/training/acs-guidelines-supplements.html>.

[140] Gregory T. Rushton et al., "Setting a Standard for Chemistry Education in the Next Generation: A Retrosynthetic Analysis," *ACS Central Science* 2, no. 11 (2016): 825–33.

Corey with the development of the technique. "All that is in this book owes its origin to his [Corey's] work," Warren wrote.[141]

In one way, Woodward's prediction that the "creative aspects of synthetic design" would not be mechanized remained true: neither LHASA nor other computer programs for synthetic planning found widespread use. But the method developed to make synthetic design accessible to computer automation did take hold. Tamed and naturalized through its pedagogical adaptations, retrosynthetic analysis came to be seen as a fundamental methodological innovation embodying timeless chemical common sense, developed for computers but more useful for humans.

Conclusion

In 1988, on his acceptance of the Robert Robinson Award of Britain's Royal Society of Chemistry – another honor whose first recipient had been Woodward – E. J. Corey delivered a lecture on "retrosynthetic thinking." Corey argued that the "intuitive search for clues to the solution of the problem at hand" (the approach to synthetic design celebrated by Woodward) had given way to the "effective and consciously applied general problem-solving techniques" of retrosynthetic analysis. Of these "fundamentals of retrosynthetic thinking," Corey explained:

[141] Warren and Wyatt, *Organic Synthesis: The Disconnection Approach*, 2nd ed. (ref. 11), 2.

> These strategies have been described previously in connection with the computer-assisted analysis of synthetic problems and the interactive program, LHASA, which is designed to emulate the problem solving techniques used by chemists. In turn, the LHASA project has been of great value in the development of new and general ways of thinking about synthesis.[142]

Returning to the questions raised at the outset about idiosyncrasy and generality, novelty and continuity, and machine and human orientation in the development of retrosynthetic analysis: Corey and his collaborators developed a general method for human thinking by taking particular intuitions embodied in past synthetic plans and bringing them together within a computer program. Retrosynthetic analysis did not enable any otherwise unthinkable synthesis or synthetic step. Rather, the novelty of the method lay in its deliberateness and generality, and in the new entities of synthetic planning such as synthons that it brought into focus. The computer program became an analog for the synthetic chemist: Corey taught chemists to think like he and the LHASA group thought a computer thinking like a chemist would think. Through a process of dispersion analogous to the spread of Feynman diagrams through postwar physics documented by David Kaiser, adaptations of retrosynthetic analysis by Warren and others helped to fix the method as a mainstay of chemical education and practice.[143]

A few years before his death in 1979, Woodward proposed Corey for the National Medal of Science. In his letter of nomination, Woodward offered an elegant summation of his colleague's achievement in method-making.[144] By means of "the explicit formulation of

---

[142] E. J. Corey, "Robert Robinson Lecture. Retrosynthetic Thinking—Essentials and Examples," *Chemical Society Reviews* 17 (1988): 111–33, on 111, 113, 115.

[143] David Kaiser, *Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics* (Chicago: University of Chicago Press, 2005).

[144] Corey was awarded the medal in 1988.

strategies and problem solving methods of organic synthesis," Woodward wrote, Corey had "contributed to a major change in thinking about synthetic problems."[145] In earlier accounts, Woodward had rendered synthetic planning as a matter of intuitive creativity rather than method. Corey made it explicit.

The demands of computing, as Corey and his collaborators conceived them, provided a framework for dividing synthetic planning into well-defined objects and processes and recombining them as retrosynthetic analysis. There are numerous other fields in which automation projects, the empirical study of methods, and method-making have gone hand in hand. Such enterprises are analogous to philosophical and literary methods of *explication*: dividing a whole into parts and thereby transforming it, by means of operations made possible by a formal framework of analysis. Explication may prove generally useful for making sense of such stories of scientific computing and "computational thinking," particularly during the period in the mid-to-late twentieth century when computers, available but not yet ubiquitous, were especially liable to be mobilized in this way.[146] The LHASA group's retrosynthetic categories, flow charts, and machine-readable chemical English were not just blueprints for a computer-based tool. They were productive constraints within which Corey and the LHASA group put ways of thinking mined from synthetic plans of the past into the form of retrosynthetic analysis. They were themselves tools by which humans made methods for humans.

---

[145] R. B. Woodward, "An Account of Professor Elias J. Corey's Contributions to Chemistry" (Medal of Science Nomination), n.d. [1976], Woodward papers (ref. 44), Box 4, Folder "Corey [1976-78]."

[146] Jeannette M. Wing, "Computational Thinking," *Communications of the ACM* 49, no. 3 (2006): 33–35. For examples from literary studies and from the history of science, respectively, see Jonathan D. Culler, *Structuralist Poetics: Structuralism, Linguistics, and the Study of Literature* (Ithaca, NY: Cornell University Press, 1975), 76, 258–259; Thomas S. Kuhn, "Second Thoughts on Paradigms," in *The Essential Tension: Selected Studies in Scientific Tradition and Change* (Chicago: University of Chicago Press, 1977), 293–319.

Acknowledgements

I thank Robert Cook-Deegan, E.J. Corey, Rick Danheiser, Alan Long, Gerald Lotto, Harry Orf, David Pensak, and Cheryl Ryder for speaking with me, sharing materials from their personal collections, and helping arrange interviews. I thank William Deringer, Stephanie Dick, Michael Gordin, Roald Hoffmann, Ursula Klein, Carsten Reinhardt, two anonymous reviewers, and participants in the Program Seminar in History of Science at Princeton University for thoughtful comments on versions of this paper.