A second chapter of equal size is taken up with discoveries in the plane, and a 124-page final chapter carries the subject into 3-space and higher dimensions. Up to the end of the solution of Hilbert's problem, the discussion is generally easy-going and elementary. Beyond this point, the arguments soon become longer, more complicated and sophisticated (from p. 130). While the author continues to explain everything in full detail, this part of the book demands much more drive and concentration and is clearly an object for serious study. However, the motivated reader will not go unrewarded. He will discover another instance of the unity of mathematics in the way several branches of abstract mathematics converge to solve a problem of the most concrete kind. For example, the 20-page proof of the Dehn-Sydler theorem is highly algebraic and draws not only from the now standard vector methods of modern geometry (Minkowski sums) but uses techniques and results from functional equations, group and ring theory, set theory, and linear algebra. As often observed in many quarters, it is impressive what mathematics can do when it pulls itself together.

Later topics include an extension of a few of the results to spaces of higher dimension, notably 4 dimensions, and a brief discussion of connections with the modern subject known as the algebra of polyhedra. Although the volume constitutes a self-contained account of a topic which is now essentially complete, it concludes with a short list of unresolved questions.

There are a few misprints, but few mistakes that the reader will not see through in a matter of moments.

The subject is related to a surprising discovery made recently by Robert Connelly (Cornell University), who is working in this general area at the present time. In 1813, Cauchy proved that a convex polyhedron with rigid faces is itself a rigid solid, that is, even if it were hinged along every edge, its shape could not be altered without forcibly breaking the surface. Connelly produced a nonconvex rigid-faced polyhedron which, if considered to be hinged at its edges, can be moved continuously through a small range of shapes without distortion of any face. For a description of this polyhedron and instructions for constructing a model, see Robert Connelly, *A flexible sphere*, The Mathematical Intelligencer (Springer-Verlag), volume 1, number 3, 1978.

The interested reader might also be on the lookout for a forthcoming book by Irving Kaplansky on all 23 of Hilbert's Paris problems.

ROSS A. HONSBERGER

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 1, Number 4, July 1979

0002-9904/79/0000-0307/$02.00

Mechanizing hypothesis formation. Mathematical foundations for a general theory, by P. Hájek and T. Havránek, Universitext, Springer-Verlag, Berlin-Heidelberg-New York, 1978, xv + 396 pp., $24.00.

I know of no book on statistics that has "Hypothesis formulation" in its index; nor is it in the indexes of Ralston and Meek [17], Mathematical Society of Japan [13], Polanyi [15], Winston [20], nor Boden [1]. But something had been said about the matter before Hájek began publishing his work in 1966.

Perhaps the earliest opinion about the formulation of hypotheses is that it is an exclusively human creative activity requiring imagination and inspiration. When people became interested in artificial intelligence they began to question this opinion, and some small successes were achieved. I shall attempt to give a brief account of some of these advances.

Let us begin by considering computer chess. When you analyze a chess position it is usually necessary to construct an analysis tree of variations, whether you are a human or a machine. It is seldom practicable to allow the analysis tree to grow to the end of the game; instead it is of course necessary to truncate the tree by some procedure, such as when the position reaches quiescence (suitably defined). At these quiescent positions some *evaluation* of the position must be made, and then, by working backwards we can arrive at sensible moves. This much has, I think, always been clear to introspective chess players, and part of the problem in playing good chess is to use good "evaluation functions". For example, if, for the sake of simplicity, we allow only for the values of the pieces (which is not in fact sufficient for good chess), we might score c_P, c_B, c_N, c_R, and c_Q for the pawns, bishops, knights, rooks, and queens. Then White's evaluation function would be $c_P(x_P - y_P) + \cdots + c_Q(x_Q - y_Q)$, where x_P and y_P are the numbers of white and black pawns, etc. We could now imagine a computer playing many games of chess and gradually improving on the coefficients. This would be a learning program in which a near-optimal linear evaluation function was attained by estimating its coefficients. This would normally be described as an estimation procedure. But every problem of estimation can also be regarded as one of looking for a hypothesis, namely the hypothesis that the best estimate is at or near some specified point. For the game of chess this method was first proposed in print by Shannon **[19]** and was first used effectively in a computer program for checkers by Samuel **[18]**. In statistics the corresponding method is known as "evolutionary operation" (see Box and Draper **[2]**).

But the mere optimization of coefficients, even if polynomials of higher than the first degree are used, can hardly be regarded as creative. Automatic creativity can reasonably be regarded as occurring if the program finds a good *form* for the evaluation function. Suppose we, the programmers, had included quadratic terms and had repeated the experiment. This time the experiment would take longer because there are many more quadratic than linear terms. But we might very well have discovered the principle known to chess players, for at least a century, as "the advantage of two bishops" (Good **[5]**). This means that although the values of bishops and knights are about equal, a bishop pair is usually more valuable than two knights or than a bishop and a knight. In statistical terminology there is an "interaction effect" between the two bishops. To a *human*, this might suggest the hypothesis that the control of two squares of equal value (for squares too have various values) is better than the double control of one such square; but a machine could not propose this hypothesis by the method just described.

Machine creativity becomes possible by the following polynomial search procedure. Suppose that when an "interaction" or significant quadratic terms has been discovered, this term is *redefined as a "primitive concept"*, that is, it is denoted by a single symbol x (such as x_{BB} in the above example). To quote Good [9, p. 206] (or see Good [7, p. 94]), "In this way one might then discover a concept involving three or more of the original primitive terms without ever using evaluation functions of degree higher than the second." This saving of time would be important because the number of cubic terms is enormous when the number of primitive terms is appreciable. If a machine were so fast that it could cope with all cubic terms from the start it might achieve apparent creativity by brute force, but if it used the search procedure just described it could reasonably be called "intelligently creative". Incidentally the brain compensates for the slowness of its individual neurons by having billions of neurons operating *simultaneously*, a feat that no computer can yet emulate. The ultra-parallel machine has yet to be built, which partly explains why creativity is still mainly a human accomplishment.

The question now arises whether this discussion in terms of polynomials can be extended to the domain of propositions. Everything depends on whether the propositions of interest can usually be built up hierarchically by simple operations. For, if so, an analogue could be found for the intelligently creative process. For the so-called "propositional calculus", much can presumably be done by using Boolean algebra, but when we consider ordinary science we have to use at least the predicate calculus and then the difficulties become formidable. This is clear from the attempts to express science in terms of formal logic made by Woodger [21] and by Carnap [3].

It is interesting to note how evolution proceeds by mixing and mutation of long instruction sequences known of course as chromosomes, a point discussed at some length by Holland [**10**] in the context of adaptive behavior in general. This suggests that one approach to the automatic generation of hypotheses is to use much the same method, but this is easier said than done. Some primitive experiments in somewhat this direction were described by Fogel *et al*, [**4**]. From the point of view of the survival of humanity it is worth noting that intelligent machines constructed by too closely imitating natural selection would be liable to be aggressive.

The activity of perception, including the understanding of language, depends on hypothesis formulation, although we do not usually think of it that way. We start with "elements" such as phonemes and build these up into words, phrases, and other structures. But often we recognize or assume the elements because we have already conjectured the larger structures. (This is why we often don't notice misprints.) Thus hypothesis formulation can work both up and down a hierarchical classification of linguistic or pictorial units, a point emphasized by Good [**8**].

In statistics I know of only one method for the automatic generation of hypotheses, other than the polynomial search procedure, namely the method of maximum entropy. As used by Jaynes [**11**], this method was used for generating prior distributions, but it was applied to the generation of hypotheses for multidimensional contingency tables and for Markov chains by Good [**6**]. In this application it turned out that the hypotheses generated were the same as one previously suggested by statisticians for intuitive reasons, and this suggests that the method is of value for the formal embodiment of some parts of intuition.

The book by Hájek and Havránek has little to say about the historical background, apart from pages 374 to 376 where the earliest references are to Chytil (1964, unpublished), and to Leinfellner [**12**]. Also one industrial application is given but "the most important factors occurring in output sentences are of technical character and give little information to a layman" (p. 380). Clearly there is not yet an application of scientific, as distinct from mathematical or logical, interest, and at present the whole subject is still in its infancy although the mathematics used by Hájek and Havránek is extremely complicated. The results obtained by Michalski and Negri [**14**] are much easier to understand (though they do not describe their *methods* in that paper); for example, in a chess end game with K and P versus K, one conclusion reached by their program was "Black can draw the game when the pawn is a rook ['s] pawn and the black king is ahead of the pawn on the same column or on the adjacent one." This kind of knowledge about the game is of course a powerful adjunct to the analysis of trees mentioned before.

In conclusion I again ask: *Can interesting hypotheses be built up hierarchically by simple automatic steps*? One would expect this to be possible in a manner analogous to the parsing or diagramming of sentences and to the development of chromosomes. If so, the intelligent creative process can be automated. This in my opinion is the central problem of automatic hypothesis formulation and I was surprised that, in a book of 400 pages, I was unable to find this fact stated in plain English. Perhaps it is implicit somewhere in the masses of intricate formalism. The aim of producing *interesting* hypotheses is, however, clearly stated.

REFERENCES

1. M. A. Boden, *Artificial intelligence and natural man*, Basic Books, New York, 1977.

2. G. E. P. Box and N. R. Draper, *Evolutionary operation*, Wiley, New York, 1969.

3. R. Carnap, *Introduction to symbolic logic and its applications*, Dover, New York, 1958.

4. L. J. Fogel, A. J. Owens and M. J. Walsh, *Artificial intelligence through simulated evolution*, Wiley, New York, 1966.

5. I. J. Good, *Could a machine make probability judgements*?, Computers and Automation **8** (1959), 14–16 and 24–26.

6. ______, *Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables*, Ann. Math. Statist. **34** (1963), 911–934.

7. ______, *A five-year plan for automatic chess*, Machine Intelligence 2, (E. Dale and D. Michie, eds.), Oliver and Boyd, Edinburgh, 1968, pp. 89–118.

8. ______, *Creativity and duality in perception and recall*, I.E.E./N.P.L. Conference on Pattern Recognition, July 1968 (Institution of Electrical Engineers), London, 1968, pp. 228–237.

9. ______, *Summing up of the discussion on inductive inference*, Machine Intelligence 8 (E. W. Elcock and D. Michie, eds.), Halsted Press, New York, 1977, pp. 205–206.

10. J. H. Holland, *Adaptation in natural and artificial systems*, Univ. of Michigan Press, Ann Arbor, 1975.

11. E. T. Jaynes, *Information theory and statistical mechanics*, Phys. Review **106** (1957), 620–630.

12. W. Leinfellner, *Struktur und Aufbau wissenschäftlicher Theorien*, Physica-Verlag, Vienna, 1965.

13. Mathematical Society of Japan, *Encyclopedic dictionary of mathematics* (English transl.), MIT Press, Cambridge, Mass., 1977.

14. R. S. Michalski, and P. Negri, *An experiment in inductive learning in chess end games*, Machine Intelligence 8 (E. W. Elcock and D. Michie, eds.), Halsted Press, New York, 1977, pp. 175–185.

15. M. Polanyi, *Personal knowledge*, Routledge and Kegan Paul, London, 1958.

16. K. R. Popper, *The logic of scientific discovery*, Hutchinson, London, 1959.

17. A. Ralston and C. L. Meek (eds.), *Encyclopedia of computer science*, Petrocelli and Charter, New York, 1976.

18. A. L. Samuel, *Some studies in machine learning using the game of checkers*, IBM J. Res. Develop. **3** (1959), 210–229.

19. C. E. Shannon, *Programming a computer for playing chess*, Phill. Mag. (London) 7th Ser. **41** (1950), 256–275.

20. P. H. Winston, *Artificial intelligence*, Addison-Wesley, Reading, Mass., 1977.

21. J. Woodger, *The axiomatic method in biology*, Cambridge Univ. Press, London and New York 1937.

I. J. GOOD

BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 1, Number 4, July 1979

0002-9904/79/0000-0308/$02.25

Topics in group rings, by Sudarshan K. Sehgal, Monographs and Textbooks in Pure and Applied Mathematics, vol. 50, Marcel Dekker, New York and Basel, 1978, vi + 251 pp., $24.50.

The theory of the group ring has a peculiar history. In some sense, it goes back to the 1890's, but it has emerged as a separate focus of study only in relatively recent times. We start with the early development of the theory of representations of finite groups over the complex field. Most people familiar with this think immediately of Frobenius and Burnside, who used approaches that seem unsuitable and even bizarre in the light of modern treatments. Admittedly, Frobenius' group determinant and Burnside's Lie-theoretic approach both yielded the basic properties of complex *characters*. However, they said much less about the *representations* themselves. For this reason, they have little application to the important problems of finding properties of representations over other rings–representations over fields of finite characteristic and over rings of algebraic integers have very important applications in group theory, algebraic number theory and topology. Hence, a more flexible approach was needed. In fact, the groundwork was being done by the little-known Estonian innovator Theodor Molien. Molien developed a theory of algebras over the complex field that included many of the features of the more general theory later developed by Wedderburn. He applied his theory to the case of representations of groups as follows: the Cayley representation of a finite group G as a permutation group on itself can be linearized to obtain a faithful representation of G in $\mathrm{GL}(n, \mathbf{C})$, where n is the order of G. The linear span of the image of this representation is a subalgebra of the algebra of $n \times n$ matrices. When it is analyzed by Molien's methods, the basic properties of the irreducible characters are deduced from properties of the irreducible representations.

This useful point of view lay dormant until the late 1920's, when Emmy Noether considered group representations as an illustration of her results on