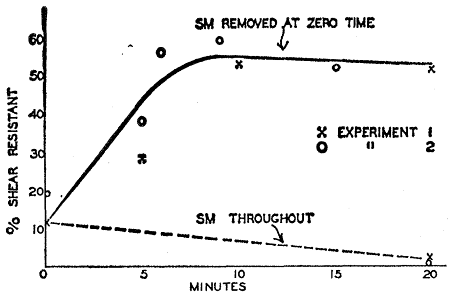

Fig. 1. Development of shear-resistance in infected cells after removal of streptomycin by dilution. Cells infected with P9 at 5 to 7 $\times$ 10^8 plaque-forming units per ml were diluted into Todd-Hewitt broth either with or without streptomycin and incubated at 37°C. SM, streptomycin.

polyamines. Since secondary amines like spermine and spermidine are more basic than the primary amines, these former compounds would probably bind to DNA more tightly and would thus be more effective than the simple diamines in reversing streptomycin activity.

Streptomycin is known to precipitate DNA, and Cohen (*9*) has shown that the streptidine portion of the molecule alone does not cause precipitation. From titration studies, Cohen concluded that the secondary amino group on the *N*-methyl-L-glucosamine portion of the molecule participated in DNA precipitation and that streptomycin probably precipitated by forming cross linkages in which the guanido group associates with one DNA molecule and the secondary amino group associates with a second molecule. Streptidine (*10*) had no inhibitory effect on phage P9 at concentrations as high as 1 $\times$ $10^{-2}M$ but competitively antagonized the inhibition induced by streptomycin. About seven streptidine molecules were required to reverse one streptomycin molecule. This difference may result from the tighter binding of the complete streptomycin molecule. The competitive relationship between these two substances suggests that at least one portion of the streptomycin molecule essential for activity is the streptidine portion.

We propose that streptomycin inhibits injection of phage P9 either by entering the phage head and combining with the coiled DNA in such a way that various portions of the molecule are cross-linked, so that the uncoiling which is essential for injection cannot occur, or by linking the DNA to the protein coat. Divalent ions, polyamines, and streptidine would then reverse streptomycin inhibition by preventing the binding of the antibiotics or by displacing the streptomycin from the nucleic acid.

As yet we have no idea why two of the streptococcus phages are resistant to the antibiotic. Permeability of the phage head to the antibiotic does not seem to be an explanation, since both resistant and sensitive phages of streptococcus undergo photoinactivation by methylene blue without a lag and are resistant to osmotic shock, implying free permeability. Also, a permeable mutant of phage T4 (*11*) of *E. coli* is completely resistant to streptomycin as is wild type T4.

Since it is possible that the simple system used here has isolated a single molecular aspect of streptomycin action, it is pertinent to ask whether our results will provide any insight into the mode of action of streptomycin as an antibacterial agent. Streptomycin inhibits protein synthesis in a cell-free ribosomal system (*12*) and alters the integrity of the cell membrane (*13*). Both ribosomes and cell membranes are stabilized by magnesium ions or by polyamines (*14*). Mager *et al.* (*15*) have suggested a common site for the action of polyamines, magnesium ions, and streptomycin in *E. coli* ribosomes. Cations reverse the antibacterial action of streptomycin and they are also able to inhibit the precipitation of DNA by the antibiotic (*16*). It thus seems possible that streptomycin acts on all these bacterial structures by competing for magnesium ions or polyamines and precipitating or altering the structure of the sensitive sites. This "unitary" hypothesis has the dual virtues both of explaining a confusing variety of antibacterial effects of streptomycin, and at the same time suggesting a number of possible experiments.

Recent work has shown that the DNA of purified phage P3 (streptomycin-resistant) and phage P9 (streptomycin-sensitive) do not contain any unusual bases, so that streptomycin sensitivity cannot be due to a peculiarity of base composition. The importance of these observations for an understanding of approaches to antiviral chemotherapy seems clear. Curiously, study of the injection process has been neglected in recent years, perhaps because of the greater fascination with how the nucleic acid works after it has entered the cell (*17*).

THOMAS D. BROCK
SUSAN O. WOOLEY
Department of Bacteriology, Indiana University, Bloomington

References and Notes

1. T. D. Brock, *Biochem. Biophys. Res. Commun.* **9**, 184 (1962).
2. ———, J. Mosser, B. Peacher, *J. Gen. Microbiol.*, in press.
3. A. D. Hershey and M. Chase, *J. Gen. Physiol.* **36**, 39 (1952).
4. S. E. Luria and D. L. Steiner, *J. Bacteriol.* **67**, 635 (1954).
5. McDonald Engineering Co., 27105 Knickerbocker Road, Bay Village, Ohio.
6. Pavement Marking Beads, Minnesota Mining and Manufacturing Co., Type 110.
7. D. Fraser and H. R. Mahler, *J. Am. Chem. Soc.* **80**, 6456 (1958).
8. H. R. Mahler, B. D. Mehrotra, C. W. Sharp, *Biochem. Biophys. Res. Commun.* **4**, 79 (1961).
9. S. S. Cohen, *J. Biol. Chem.* **168**, 511 (1947).
10. Streptidine was provided by Merck, Sharp and Dohme.
11. Bacteriophage T4 $r^{\pm}o_5$ was provided by Dr. S. Champe.
12. J. G. Flaks, E. C. Cox, J. R. White, *Biochem. Biophys. Res. Commun.* **7**, 385 (1962).
13. N. Anand, B. D. Davis, A. K. Armitage, *Nature* **185**, 23 (1960).
14. H. Tabor, C. W. Tabor, S. M. Rosenthal, *Ann. Rev. Biochem.* **30**, 579 (1961).
15. J. Mager, M. Benedict, M. Artman, *Biochim. Biophys. Acta* **62**, 202 (1962).
16. R. Donovick, A. P. Bayan, P. Canales, F. Pansy, *J. Bacteriol.* **56**, 125 (1948).
17. One of us (T.D.B.) is a research career development awardee of the U.S. Public Health Service (grant No. AI-K3-18,403). This work was supported in part by grants from the U.S. Public Health Service (AM-04630-03) and U.S. Army (DA-CML-18-064-61-G13). We thank W. B. Deville for help with certain experiments.

22 May 1963

Information Storage Requirements for the Contents of the World's Libraries

Abstract. *Three calculations of the number of different things stored in the world's libraries yield estimates from* 7.5×10^7 *to* 7.7×10^8. *At* 10^5 *words per volume, five letters per word, and 12 bits per letter, the information capacity used for storage is between* 4.6×10^{14} *and* 4.6×10^{15} *bits, and is increasing at about* 2×10^6 *bits per second.*

In the course of a project aimed at understanding the problems and possibilities of the library (or the successor of libraries) of the next century, we have collected and reorganized information about the "size" of the world's literature. The information is useful as a basis for estimating the magnitude of the storage problem. Particularly if we are to consider the use of computers and computer memory systems as replacements for many present functions of a library, we wish to have estimates of the amount of storage that is needed and the rate at which it is growing. With that information we can judge whether presently available computers can serve as a major library, and, if not, what kinds of developments are needed.

The problem was to estimate how many differently arranged printed letters (or equivalent letters) there are in the world's libraries; how fast the number is growing; and how many bits of information storage capacity these represent. I have obtained estimates of the total number of different letters by three different routes. These are presented below.

First, using the United States total library holdings as a base, I estimate the replication ratio, that is, the average number of each title held in libraries.

In the United States in 1958 there were 1.1×10^7 volumes in the Library of Congress, 1.69×10^8 in university libraries, and about 3.26×10^8 in school and public libraries combined (*1*, *2*) making a total holding of about 5.1×10^8. I assume that the entire contents of the public and school libraries replicate things in the Library of Congress (*3*). A more difficult question is that of the university libraries. I estimate that it is generous to add 1 percent of their total holdings to those of the Library of Congress. I assume that the Library of Congress holds no duplicates (or at least an insignificant number in view of the rounding to be done). We add, then, 1.7×10^6 to the holdings of the Library of Congress and round to 1.3×10^7 as our estimate of the number of *different* volumes in the United States. Since the total holdings of the libraries of the United States were 5.1×10^8, the apparent replication rate is 40 of each volume, on the average. If the same ratio holds for the rest of the world, then our estimate of the number of volumes in the world's literature is equal to the world's total reported library holdings divided by 40. The estimated total is 3×10^9 (*1*); the estimate of "different volumes" is 7.5×10^7 (estimate 1a). However, the Library of Congress lists *all* its items as 3.5 times its number of volumes (*2*). If this ratio holds for all other library systems, the estimate of "different items" (equivalent volumes) is $3.5 \times 7.5 \times 10^7$, or 2.6×10^8 (estimate 1b) (*4*).

As the second route, using as an approximation the notion that each national library listed (*1*) has all the different books published in its country and none of the books published outside that country, I merely add all the national library holdings and apply the same two factors used above: 1 and 3.5.

The total holdings of all the national libraries reported (*1*) is 1.4×10^8 volumes. If "volumes" includes everything in the reporting libraries, then that figure is the estimated number of different items (estimate 2a); the factor of 3.5 gives the number of different items as $3.5 \times 1.4 \times 10^8$ or 4.9×10^8 (estimate 2b).

Table 1. Estimates of the number of "different volumes and pamphlets" in the world's libraries, arrived at by three different routes.

Factor	Route 1	Route 2	Route 3
1.0	7.5×10^7	1.4×10^8	2.2×10^8
3.5	2.6×10^8	4.9×10^8	7.7×10^8

As the third route, instead of using the reported size of the store and estimating replication rate, one can estimate the rate of addition to the store and the rate of growth of the store, and compute the size of the store such that the rate of growth times the size of the store is equal to the rate of addition.

The rate of growth of libraries of the United States is fairly well documented. The 42 major university libraries added 3.7 percent to their holdings in 1959–60 (*5*). The public libraries added 6 to 6.5 percent per year from 1939 through 1956 (*5*). The Library of Congress added both "volumes" and "other items" at a slightly smaller rate than the university libraries —3.1 percent per year. [From 1951 through 1954 the two figures were 9.9 and 9.5 percent; and from 1954 through 1960, 19.6 and 20.3 percent. These figures give a fairly constant growth rate of 3.1 percent per year (*2*).] Since the Library of Congress is more likely to grow by the acquisition of *new* books than by duplication, we may take its growth rate as more indicative of the rate at which *different* items are added to the holdings of the libraries of the United States. The book trade statistics give a new "book title and editions" publication rate of about 15,000 for 1959 (*5*). However, the growth rate of the Library of Congress is about 3.5×10^5 volumes and pamphlets per year. If this ratio of 23.3 to 1.0 [$(3.5 \times 10^5)/(1.5 \times 10^4)$] holds for the rest of the world between "books" and "volumes and pamphlets," and the total number of "new book titles and editions" for the world is about 3×10^5 (*5*), then the total number of different "volumes and pamphlets" added to the world's store for the year is $23.3 \times 3 \times 10^5$ or 7×10^6 volumes and pamphlets. If this number is 3.1 percent of the holdings, then the number of different volumes and pamphlets is $(7 \times 10^6)/(3.1 \times 10^{-2})$ or 2.2×10^8 "different volumes and pamphlets" (estimate 3a). Using the previously used ratio of 3.5, the estimate is 7.7×10^8 (estimate 3b) "equivalent volumes" (*6*). Table 1 shows the values arrived at by the different routes.

The total range of the various estimates is only slightly more than one order of magnitude: 7.5×10^7 to 7.7×10^8. At 10^5 words per book, and five letters per word, the number of letters to be stored is between 3.8×10^{13} and 3.8×10^{14}. At 50 characters per alphabet and with 50 alphabets, there are 2500 characters to be identified, or about 12 bits per character. The information storage measure of the "world's literature" then is between 4.6×10^{14} and 4.6×10^{15} bits; 2×10^{15} is conveniently close to the mean. Naturally we cannot store "efficiently"; every literary work must be stored verbatim. A nonredundant reconstruction of a detective novel would be dull reading.

As indicated above, the growth rate is estimated to be about 3.1 percent per year, doubling in 22 years. Thus, the current addition rate is about 6.2×10^{13} bits per year, or 2×10^6 bits per second—a sobering thought (*7*).

JOHN W. SENDERS

Bolt Beranek and Newman Inc.,
50 Moulton Street,
Cambridge 38, Massachusetts

References and Notes

1. *Basic Facts and Figures: International Statistics Relating to Education, Culture, and Mass Communication* (UNESCO Publication, New York, International Document Service, Columbia Univ., 1952–1960).
2. *American Library Directory* (Bowker, New York, eds. 19, 20, and 21, 1951, 1954, and 1960).
3. Of course, every library has some things which are unique to it, but they are as nothing beside the enormous numbers involved.
4. I equate "item" to "volume" with many qualms, but it is difficult to estimate the information storage needs for a map or a photograph of unspecified nature, and easy to do so for a printed volume.
5. *The Publishers Trade List Annual, 1962* (Bowker, New York, 1962).
6. As an aside it must be noted that the 2×10^6 scientific articles per year, as estimated by Bourne [C. P. Bourne, *The World's Technical Journal Literature: An Estimate of Volume Origin, Language, Field, Indexing, and Abstracting*, (Stanford Research Inst., Menlo Park, Calif., 1961)], contribute the equivalent at ten pages per article, of about 10^5 books per year, and this does not significantly alter the estimate. In a sense the figure is included in the expansion of "books" to "volumes and pamphlets."
7. The preparation of this paper was supported by a grant from the Council on Library Resources, Inc.

3 July 1963