Punched Cards: A Brief Tutorial

Robert V. Williams

Because of space considerations in the April-June 2002 issue, we were unable to feature the following items in the Robert V. Williams article, "The Use of Punched Cards in US Libraries and Documentation Centers, 1936-1972."

A taxonomy of punched cards can take many forms: by size, by number of recording punches, by the method used to sort or retrieve the information, by whether the card is punched internally or only on the edges, and other methods, such as whether the card contains other information-handwritten or typed, or recorded via an encased microfilm image. The usual approach classifies them as machine sorted or hand sorted and then subclassifies them with one or more of the noted criteria. Such a system is used in this brief tutorial, with the caveat that some punched-card systems, such as optical coincidence punched cards and large files of edge-notched cards (which are usually hand sorted) may employ machinery for sorting purposes.1 In all cases, sorting refers to the process of arranging cards in a predetermined order.

Machine-sorted punched-card systems

The Hollerith/IBM card was the original data recording card and the only one available for many years, until James Powers produced cards and sorting machines under his own name, the Powers Company. Later, Powers became part of Remington Rand (and, in the UK and Europe, Powers-Samas). Other companies would try to take advantage of the success of IBM and Remington Rand in selling cards and equipment, but none came close. Thousands of card varieties were developed for both general and specialized uses between the original machine-sorted punched-card format and the "standard" punched cards of the late 1960s, with 80 (IBM) or 90 (Remington Rand) punch positions. Until about 1930, the 45-column card was standard, but demands for increased data capacity made it almost obsolete. After this, except for specialized operations, the standard Hollerith/IBM card was 80 columns, 3 1/4 × 7 3/8 inches in size and exactly 0.0067 inches in thickness (with permitted variations of only 0.0005"), and made to exacting specifications for paper content, durability, and flexibility. Punches in the card were standardized to a rectangular shape, and the card had a left-hand diagonal cut on the top side to detect misfiled cards. The 80-column card had 12 punch rows. Three punch rows were called the zone, and 10 punch rows the field (overlapping at the zero row). Numbers were encoded only in the field although letters and punctuation used both zone and field.

The standard Remington Rand card had 45 hole-punch columns and 12 punch rows divided into two sets, each of six rows, producing a 90-column card. Data were encoded by combinations of one to three punches among the six punch rows assigned to each set. The cards measured 3 1/4 × 7 3/8 inches, no more than 0.007 inches or less than 0.00625 inches in thickness, and the punches were oval. These cards were also made to exacting specifications for paper stock and had a left-hand diagonal cut.2

Although "standard" cards were certainly the most commonly used (and the least expensive) cards in machine sorting operations, there were many variations for particular uses. Cards could be plain or preprinted to customer specifications or uses; they could be smaller or larger than the standard size, depending on customer needs; they could be perforated so that parts were detachable for separate processing later; they could be designed for hand marking (called mark sense) so that a card reader/puncher could then punch the marked spots (see Figure A). Certain columns could be designated for specific information (fixed field); or some columns fixed and some left entirely blank, or all column designations left up to the user (free field).

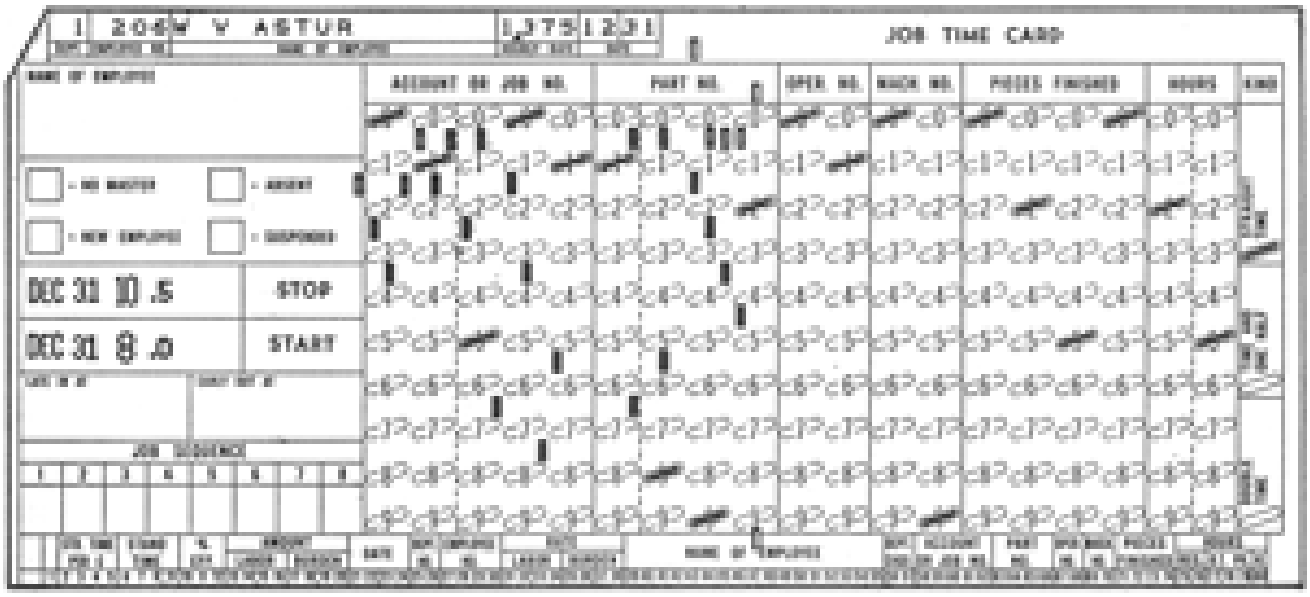

Figure A.
IBM "mark-sensed" card, preprinted for specific job application. (Source: R.S. Casey and J.W. Perry, eds., Punched Cards: Their Applications to Science and Industry, 1951. Courtesy of John Wiley & Sons.)

Machine-sorted cards obviously required machinery. A wide variety of equipment was available from IBM and Remington Rand, but the following were essential to most operations: key punches for punching the holes, with or without a device for printing the punched characters on the card; verifiers to double-check for punching and transcription errors; sorters for placing cards into the appropriate categories needed for later analysis or storage; collators for more rapid sorting, merging, and comparing of large groups of cards; and accounting and printing units where arithmetical or other calculations take place prior to printing the results specified by the control panel or program. Other useful equipment for machine-sorted cards were card reproducers (or "reproducing punches"), calculating punches (performing arithmetical operations on some columns in the card and punching the specified results in other columns or separate cards), interpreters (which read the punches and printed results), and combination machines that performed several of the operations ordinarily done by separate machines (for example, the IBM 101 "electronic statistical machine" combined sorting, counting, accumulating, balancing, editing, and printing).

Punched cards and the machines both had to be made to exacting specifications because electricity was used to sense the card data and the nature of the processing operations. Holes in the punched card transmitted an electric current recorded by the machinery (originally, electromechanical counters), which was wired differently for each specified operation. These circuit-closing processes changed considerably over time, from pins and mercury cups in the original Hollerith designs3 to sophisticated sensing devices built into the equipment by the early 1950s.4 A card that was too thin or poorly made might result in a false signal, and a card that was too thick or not perfectly proportioned could easily jam the equipment and slow the processing operations. Cards were expensive and so was equipment rental (originally, the only way to use a Hollerith/IBM machine).

Arithmetical calculations and the printing of results were the machines' primary applications, but clearly sorting could be done on any specified characters, including the alphabet, although it was an expensive use of cards, which could contain only a few characters. Coding of one or more columns-

always the key to efficient processing and intelligible results-was a constant source of concern, particularly for operations requiring extensive use of the alphabet for sorting. Most early manuals or other books that dealt with establishing a punched-card operation contained considerable detail on recommended procedures for developing a coding system.5 Coding essentially took two forms. First, coding required setting up the machine to perform the operations required-initially with wiring that needed to be changed with each different operation, then to removable preset panels (called plugboards), and later to a set of cards that contained a job-control language for computer processing. Second, coding required designing the card layout so it was clear what was punched in each column or set of columns. Preprinting cards for routine operations aided this process because it made it easier for the keypunch operator to identify what was to be punched.

First used in 1887, machine-sorted punched cards are still in use today, particularly in voting booths, tollbooths, and a few business operations, particularly as aperture cards. At least one company, Cardamation (http://www.cardamation.com/), carries a wide range of machine-sorted cards and equipment. However, the 2000 US presidential election in Florida and the notoriety received from the use of punched-card systems (most of which used preperforated cards) may well bring the story of machine-sorted punched cards to a complete end.6

Hand-sorted punched-card systems

The distinction between machine-sorted and hand-sorted is not completely pure because some 1950s systems used mechanized approaches to the selection process (generally with a motor that shook the container holding the cards in order to separate those punched in the desired position from those that had not been punched in that position). In general, however, the distinction is useful because all these systems rely on manually handling the cards during the sorting or retrieving processes. This manual aspect meant that systems employing hand-sorted cards were generally much smaller than those using machine-sorted cards. However, some surprisingly large systems, as will be discussed in this article, did use hand-sorted cards.

The cards varied widely in size and shape but did not usually have to be manufactured to exact specifications, as with machine-sorted cards. Hand-sorted card systems could be installed relatively inexpensively because they required less outlay for equipment. The individual cards, however, were generally more expensive than machine-sorted cards. There were two basic kinds of hand-sorted cards: those punched in one or more rows on one or more edges of the card and those in the card body. The edge-notched cards were generally manipulated with a "needle," and the center-punched cards were used in optical coincidence retrieval systems, electronic sorters, or other computer input devices. Hand-sorted cards generally gave the user considerably more freedom in designing the system to fit specific needs, such as displaying text on the card or devising a unique classification or indexing system.

A variety of names, generic and trademarked, have been used for marginal punched-card systems: edge-notched cards, slotted cards, E-Z Sort, Zatocards, McBee, McBee Keysort, Flexisort, Velom, Rocket, and many others (see Figures B and C.) Marginal punched-card systems may or may not have had a central organizing principle, such as by author, chemical compound name, or some other dominant interest of the system designer. A useful feature of edge-notched cards was that filing cards in a particular order was not essential. However, careful design of the coding system to be used was essential and required considerable thought about the retrieval aspects. The coding system could be simple (called direct), such as using numbers to specify month and year, or complex (indirect), such as using an assigned series of numbers to represent a particular "subject." Users of Calvin Mooers' Zatocoding system7 often used four random numbers (for example, 12-17-25-36) to represent a subject. A number of subjects could be entered in the same field, with only a small risk of creating

"false" selections. The center portion of these cards generally contained the relevant information and could be preprinted, handwritten, recorded on microfilm (this would have been unusual), or typewritten. The card might contain all the information required by the user or refer to some other file of information (such as a journal article or book). Marginal punched cards relevant to a search were generally selected by manual manipulation (preferably, using a card holder or tray while doing the sorting) of one or more knitting-needle-type rods in the precoded holes or slots, which had been notched (or cut open with some type of simple punching device or scissors), so that the desired cards fell out from the other nonrelevant cards in the collection, which remained in place on the needle. These manipulations, using either simple or complex (such as Boolean searching) sorting, could take place as many times as needed until the desired level of specificity had been achieved.

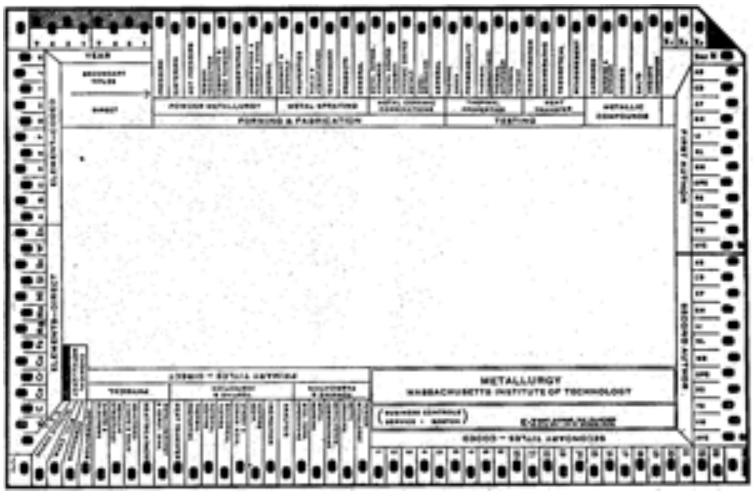

Figure B.
"E-Z sort" edge-notched card (single hole), preprinted for an information file on metallurgical literature. (Source: R.S. Casey and J.W. Perry, eds., Punched Cards: Their Applications to Science and Industry, 1951. Courtesy of John Wiley & Sons.)

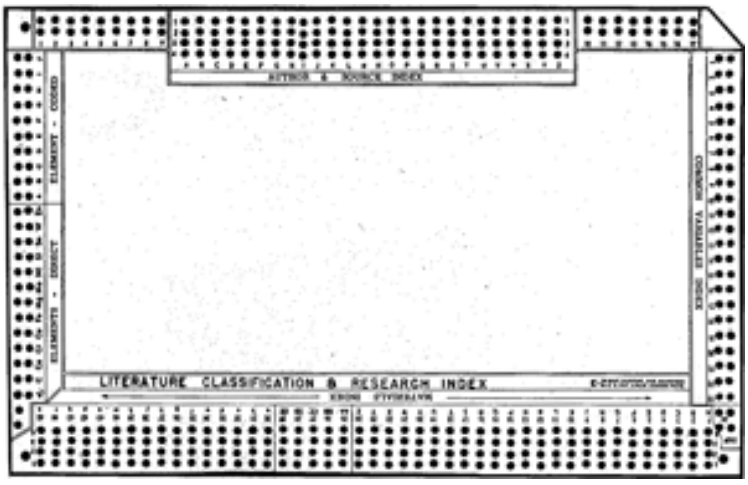

Figure C.
"E-Z sort" edge-notched card (multiple lines of holes), preprinted for specific coding by author, classification, words, and so on for literature application; the blank space is used for entering bibliographic or other information.

Edge-notched cards were first developed in 1896 by Henry P. Stamford, who patented a simple device for searching one hole at a time (the holes were not notched) to locate, for instance, insurance premium due dates.8 Similar devices for bottom-notched cards were developed shortly thereafter, but the most successful variation was the edge-notched card system developed by Alfred Perkins in Birmingham, England, for the Dunlop Rubber Company sometime prior to 1925 (when he received a US patent). The Copeland-Chatterson Company patented Perkins' device in the UK, and the McBee Corporation bought the US rights in 1932.9 Both companies successfully marketed the invention and continued development through a series of patents over the next 50 years. The McBee company changed ownership several times in the last 30 years and, now owned by New England Business Services, no longer produces the edge-notched card.

Optical coincidence card systems were generally designed around aspects, or subjects, but any desired organizational system (such as author or media type) could be used. These systems went by many names: feature cards, aspect cards, peek-a-boo, fiches superposables, superimposables, Batten cards, Cordonnier cards, and many others. These systems also used a wide variety of card sizes, depending on the size of the collection of information to be indexed.

Each card was given a subject (or similar) heading from an approved or free language list of terms. The cards were usually marked off in tiny squares and each square assigned a specific number. The numbers usually corresponded to a file of information (articles, books, patents, and so on) stored separately. Each item in this file was then analyzed, or indexed, and a hole punched (or drilled) in the square corresponding to that item number. When a particular subject term (or chemical compound or animal species, for instance) was heavily represented in the collection, there could be many holes in the card with that term. A term with little coverage in the collection would have few holes punched in the card. Simple, single-term searching was accomplished by removing that term card from the file, superimposing it over a light source, and recording the numbers of the items that showed light. Multiple-term searching, using the AND aspect of Boolean retrieval, was accomplished by superimposing several term cards simultaneously over the light source. Boolean NOT and OR searches could also be made using a similar technique with the addition of something like different-colored transparent sheets of plastic and the appropriate term cards.

The basic idea for an optical coincidence card system was first developed and patented by Horace Taylor in 1915.10 Refinements were made by others in the next few years, but significant development, at least for literature searching, did not occur until the work of W.E. Batten11 in the UK, about 1940, and of G. Cordonnier, about 1945, in France. (A significant development in mechanized information retrieval, although not with punched cards, occurred in 1927 when Germany's Emanuel Goldberg filed for a patent on a "statistical machine" that involved Boolean searching using a light beam and a photoelectric tube for finding information stored on microfilm.12)

The Batten and Cordonnier systems are discussed in the main text of the article (Robert V. Williams' "The Use of Punched Cards in US Libraries and Documentation Centers, 1936-1972.")because they directly relate to the use of punched cards in libraries and documentation centers. Developments in the US during the 1950s, by Mortimer Taube and the Jonkers Termatrex Company, will also be covered. Although no one is apparently manufacturing and selling punched-card-based optical coincidence systems currently, some of these systems are still in active use.13

References and notes

This "classification system" was also used in the two editions of what became the classic handbooks on punched cards for many years: R.S. Casey and J.W. Perry, eds., Punched Cards: Their

Applications to Science and Industry, Reinhold, New York, 1951, 2nd ed., R.S. Casey et al., eds., Reinhold, New York, 1958.

R.S. Casey et al., Punched Cards: Their Applications, 1958.

L.E. Truesdell, The Development of Punch Card Tabulation in the Bureau of the Census, 1890-1940, US Government Printing Office, Washington, D.C., 1965, pp. 49-50.

R.S. Casey and J.W. Perry, Punched Cards: Their Applications, 1951, p. 56.

For example, almost all of the articles in the first extensive treatment of the use of the "Punched Card Method" in educational institutions (G.W. Baehne, ed., Practical Applications of the Punched Card Method in Colleges and Universities, Columbia Univ. Press, New York, 1935) contain extensive discussions of coding systems used in various administrative and research functions described.

For an interesting and brief cultural history of the machine-sorted punched card, see Steven Lubar, of the Smithsonian Institution, "'Do not fold, spindle or mutilate': A cultural history of the punch card," http://ccat.sas.upenn.edu/slubar/fsm.html. Other interesting details about it may be located at the Web site of Doug Jones, Univ. of Iowa, http://www.cs.uiowa.edu/~jones/cards/.

C.W. Brenner and C.N. Mooers, "A Case History of a Zatocoding Information Retrieval System," Punched Cards: Their Applications to Science and Industry, 2nd ed., R.S. Casey et al., eds., Reinhold, New York, 1958, pp. 340-356.

H.P. Stamford, Information Card, US patent 564,117, to Henry Stamford, Patent and Trademark Office, Washington, D.C., 14 July 1896.

F. Reichman, "Notched Cards," The State of the Library Art, vol. 4, part 1, R.R. Shaw, ed., Rutgers Univ., Graduate School of Library Service, New Brunswick, N.J., 1961, pp. 12-13.

L.S. Thompson, "Feature Cards," The State of the Library Art, vol. 4, part 2, R.R. Shaw, ed., Rutgers Univ., Graduate School of Library Service, New Brunswick, N.J., 1961, pp. 57-58.

W.E. Batten, "Specialized Files for Patent Searching," Punched Cards: Their Applications to Science and Industry, R.S. Casey and J.W. Perry, eds., 1951, pp. 169-181.

M.K. Buckland, "Emanuel Goldberg, Electronic Document Retrieval, and Vannevar Bush's Memex," J. Am. Soc. for Information Science, vol. 43, no. 4, 1992, pp. 284-294.

I recently received an email from a company in Pennsylvania, asking if I knew of any way to convert an actively used large file (in excess of 100,000 cards) of Jonkers Termatrex cards to a computer-based system!