

All-Star Contest**The Soundex Indexing System**

As found at www.avotaynu.com/soundex.html and written by Gary Mokotoff

PROBLEM: The latest significant improvement to soundexing is the Daitch-Mokotoff soundex system. In 1985, Gary Mokotoff indexed the names of some 28,000 Eastern Europeans who had changed their names after WWII. There were numerous spelling variants of the same basic surname and the list needed to be soundexed. Using the conventional U.S. government system, Eastern European names that sounded the same did not soundex the same. The most prevalent were those names spelled interchangeably with the letter *w* or *v*, for example, the names Moskowitz and Moskovitz.

A modification to the U.S. soundex system was then created and published. Randy Daitch read the article and expanded on the rules of the new system. It included the following improvements over the conventional system:

Rules of the Daitch-Mokotoff Soundex System

The rules for converting names into D-M code numbers are listed below. The Coding Chart follows them. Turn to the chart briefly to familiarize yourself with the concept and then return to the specific instructions.

1. Names are coded to six digits, each digit representing a sound listed in the Coding Chart below.
2. The letters A, E, I, O, U, J and Y are always coded at the beginning of a name, as in Augsburg (054795). In any other situation, they are ignored except when two of them form a pair and the pair comes before a vowel, as in Breuer (791900), but not Freud. The letter "H" is coded at the beginning of a name, as in Halberstadt (587943) or preceding a vowel as in Mannheim (665600), otherwise it is not coded.
3. The adjacent letters combinations as shown in the chart must be coded that way as in Chernowitz, which is not coded Chernowi-*t- z* (496734) but Chernowi-*tz* (496740).
4. When adjacent letters have the same code number, they are coded as one sound, as in Cherkassy, which is not coded Cherka-*s-sy* (495440) but Cherkassy (495400). Exceptions to this rule are the letter combinations "MN" and "NM" whose letters are coded separately, as in Kleinman which is coded 586660 not 586600.
5. When a name consists of more than one word, it is coded as if one word, such as Nowy Targ, which is treated as Nowytarg (673950).
6. Several letters and letter combinations pose the problem that they may sound in one of two ways. The letter and letter combinations CHS, KS (see chart below), are assigned two code numbers.

7. When a name lacks enough coded sounds to fill the six digits, the remaining digits are coded "0" as in Berlin (798600) that has only four coded sounds (B-R-L-N).

SAMPLE INPUT

SAMPLE OUTPUT

- | | |
|---------------|-----------|
| 1. CENIOW | 1. 467000 |
| 2. TSENYUV | 2. 467000 |
| 3. HOLUBICA | 3. 587740 |
| 4. GOLUBITS | 4. 587400 |
| 5. PSHEMESHIL | 5. 746480 |
| 6. LAPSCHD | 6. 874300 |

Daitch-Mokotoff Soundex Coding Chart (modified and abbreviated)
NC=NOT CODED

Letter	Alternate Letters	Start of a name	Before a vowel	All other cases
AI	AJ, AY,EI, EY, IE, IO, IU, OI, OJ, OY	0	1	NC
AU		0	7	NC
A, E, I, O ,U		0	NC	NC
B, F, P, V, W	FB, PF, PH	7	7	7
D, T	DT, TH	3	3	3
G, K, Q	KH, CK	5	5	5
H		5	5	NC
L		8	8	8
M, N		6	6	6
R	RZ,RS	9	9	9
C, S, Z	TCH, CS, CSZ, CZS, DS,DZ, CZ,TS, SH, SCH, SZ, TZ	4	4	4
X	CHS, KS	5	54	54
Y		1	NC	NC
	EU	1	1	NC
J		2	2	2
	MN,NM	66	66	66
	SD, SHD, ST, SZD, SCHAT, SHT, SCHED	2	43	43

INPUT: There will be 10 lines of input: Each line will contain a string representing a name to be coded.

OUTPUT: For each line of input, print the D-M code.