# PENGKODEAN DATA

Komunikasi Data

# Pendahuluan

- Karakter data yang akan dikirim dari suatu titik ke titik lain tidak dapat dikirimkan secara langsung. Perlu proses pengkodean pada setiap titik. Dengan kata lain, karakter-karakter data tersebut harus dikodekan terlebih dahulu dengan kode yang dikenal oleh setiap terminal yang ada.
- Tujuan dari pengodean adalah menjadikan setiap karakter dalam sebuah informasi digital kedalam bentuk biner agar dapat ditransmisikan. Suatu terminal yang berbeda menggunakan kode niner yang berbeda untuk mewakili suatu karakter.

- Kode-kode yang digunakan untuk keperluan komunikasi data pada sistem komputer dari sejak komputer ditemukan sampai pada komunikasi data modern memiliki perbedaan dari generasi ke generasi. Hal ini disebabkan oleh semakin besar dan kompleksnya data yang akan dikirim atau dipergunakan.
- Secara umum ada beberapa kode yang digunakan dalam komunikasi data diantaranya adalah:
  1. BCD (*Binary Coded Decimal*)
  2. SBCDIC (*Standard Binary Coded Decimal Interchange Code*)
  3. EBCDIC (*Extended Binary Coded Decimal Interchange Code*)
  4. BOUDOT
  5. ASCII (*American Standard Code for Information Interchange*)

# BCD

- Merupakan kode biner yang digunakan hanya untuk mewakili nilai digit desimal dari 0-9. BCD menggunakan kombinasi 4 bit sehingga ada 16 kombinasi yang bisa diperoleh dan hanya 10 kombinasi yang digunakan.
- Kode BCD sudah jarang digunakan untuk komputer dan transmisi data sekarang ini karena tidak dapat mewakili huruh atau simbol karakter khusus. BCD hanya digunakan oleh komputer generasi pertama.

## Tabel Binary Coded Decimal

| BCD 4 bit | Digit Desimal |
|---|---|
| 0000 | 0 |
| 0001 | 2 |
| 0010 | 3 |
| 0100 | 4 |
| 0101 | 5 |
| 0110 | 6 |
| 0111 | 7 |
| 1000 | 8 |
| 1001 | 9 |

# SBCDIC

- Merupakan kode biner yang dikembangkan dari BCD. SBCDIC menggunakan kombinasi 6 bit sehingga lebih banyak kombinasi yang bisa dihasilkan. Yaitu 64 kombinasi kode. Ada 10 kode untuk digit angka dan 26 kode untuk alphabet dan sisanya untuk karakter khusus tertentu. SBCDID digunakan pada komputer generasi kedua.

# Tabel SBCDIC

| SBCDIC BA8421 | Karakter | SBCDIC BA8421 | Karakter |
|---|---|---|---|
| 001010 | 0 | 100001 | J |
| 000001 | 1 | 100010 | K |
| 000010 | 2 | 100011 | L |
| 000100 | 4 | 100101 | N |
| 000101 | 5 | 100110 | O |
| 000110 | 6 | 100111 | P |
| 000111 | 7 | 101000 | Q |
| 001000 | 8 | 101001 | R |
| 001001 | 9 | 010010 | S |
| 110001 | A | 010011 | T |
| 110010 | B | 010100 | U |
| 110011 | C | 010101 | V |
| 110100 | D | 010110 | W |
| 110101 | E | 010111 | X |
| 110110 | F | 011000 | Y |
| 110111 | G | 011001 | Z |
| 111000 | H | | |

# EBCDIC

- EBCDID adalah kode 8 bit yang memungkinkan untuk mewakili karakter 256 kombinasi karakter. Pada EBCDID, *high order bits* atau 4 bit pertama disebut *Zone bits* dan *low order bits* atau 4 bit kedua disebut dengan *numeric bit.*

# Tabel Extended Binary Code Decimal Interchange Code

| | -0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 | -8 | -9 | -A | -B | -C | -D | -E | -F |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0- | NUL<br>00 | SOH<br>01 | STX<br>02 | ETX<br>03 | SEL | HT<br>09 | RNL | DEL<br>7F | GE | SPS | RPT | VT<br>0B | FF<br>0C | CR<br>0D | SO<br>0E | SI<br>0F |
| 1- | DLE<br>10 | DC1<br>11 | DC2<br>12 | DC3<br>13 | RES<br>ENP | NL | BS<br>08 | POC | CAN<br>18 | EM<br>19 | UBS | CU1 | IFS<br>1C | IGS<br>1D | IRS<br>1E | IUS<br>ITB<br>1F |
| 2- | DS | SOS | FS | WUS | BYP<br>INP | LF<br>0A | ETB<br>17 | ESC<br>1B | SA | SFE | SM<br>SW | CSP | MFA | ENQ<br>05 | ACK<br>06 | BEL<br>07 |
| 3- | | | SYN<br>16 | IR | PP | TRN | NBS | EOT<br>04 | SBS | IT | RFF | CU3 | DC4<br>14 | NAK<br>15 | | SUB<br>1A |
| 4- | SP<br>20 | RSP<br>A0 | â<br>E2 | ä<br>E4 | à<br>E0 | á<br>E1 | ã<br>E3 | å<br>E5 | ç<br>E7 | ñ<br>F1 | [<br>5B | .<br>2E | <<br>3C | (<br>28 | +<br>2B | !<br>21 |
| 5- | &<br>26 | é<br>E9 | ê<br>EA | ë<br>EB | è<br>E8 | í<br>E0 | î<br>EE | ï<br>EF | ì<br>ED | ß<br>DF | ]<br>5D | $<br>24 | *<br>2A | )<br>29 | ;<br>3B | ^<br>5E |
| 6- | -<br>2D | /<br>2F | Â<br>C2 | Ä<br>C4 | À<br>C0 | Á<br>C1 | Ã<br>C3 | Å<br>C5 | Ç<br>C7 | Ñ<br>D1 | ¦<br>A6 | ,<br>2C | %<br>25 | _<br>5F | ><br>3E | ?<br>3F |
| 7- | ø<br>F8 | É<br>C9 | Ê<br>CA | Ë<br>CB | È<br>C8 | Í<br>CD | Î<br>CE | Ï<br>CF | Ì<br>CC | `<br>60 | :<br>3A | #<br>23 | @<br>40 | '<br>27 | =<br>3D | "<br>22 |
| 8- | Ø<br>D8 | a<br>61 | b<br>62 | c<br>63 | d<br>64 | e<br>65 | f<br>66 | g<br>67 | h<br>68 | i<br>69 | «<br>AB | »<br>BB | ð<br>F0 | ý<br>FD | þ<br>FE | ±<br>B1 |
| 9- | °<br>B0 | j<br>6A | k<br>6B | l<br>6C | m<br>6D | n<br>6E | o<br>6F | p<br>70 | q<br>71 | r<br>72 | ª<br>AA | º<br>BA | æ<br>E6 | ¸<br>B8 | Æ<br>C6 | ¤<br>A4 |
| A- | µ<br>B5 | ~<br>7E | s<br>73 | t<br>74 | u<br>75 | v<br>76 | w<br>77 | x<br>78 | y<br>79 | z<br>7A | ¡<br>A1 | ¿<br>BF | Ð<br>D0 | Ý<br>DD | Þ<br>DE | ®<br>AE |
| B- | ¢<br>A2 | £<br>A3 | ¥<br>A5 | ·<br>B7 | ©<br>A9 | §<br>A7 | ¶<br>B6 | ¼<br>BC | ½<br>BD | ¾<br>BE | ¬<br>AC | \|<br>7C | ¯<br>AF | ¨<br>A8 | ´<br>B4 | ×<br>D7 |
| C- | {<br>7B | A<br>41 | B<br>42 | C<br>43 | D<br>44 | E<br>45 | F<br>46 | G<br>47 | H<br>48 | I<br>49 | SHY<br>AD | ô<br>F4 | ö<br>F6 | ò<br>F2 | ó<br>F3 | ő<br>F5 |
| D- | }<br>7D | J<br>4A | K<br>4B | L<br>4C | M<br>4D | N<br>4E | O<br>4F | P<br>50 | Q<br>51 | R<br>52 | ¹<br>B9 | û<br>FB | ü<br>FC | ù<br>F9 | ú<br>FA | ÿ<br>FF |
| E- | \<br>5C | ÷<br>F7 | S<br>53 | T<br>54 | U<br>55 | V<br>56 | W<br>57 | X<br>58 | Y<br>59 | Z<br>5A | ²<br>82 | Ô<br>D4 | Ö<br>D6 | Ò<br>D2 | Ó<br>D3 | Ő<br>D5 |
| F- | 0<br>30 | 1<br>31 | 2<br>32 | 3<br>33 | 4<br>34 | 5<br>35 | 6<br>36 | 7<br>37 | 8<br>38 | 9<br>39 | ³<br>83 | Û<br>D8 | Ü<br>DC | Ù<br>D9 | Ú<br>DA | EO |

# Kode Boudot

- Kode Boudot terdiri atas 5 bit yang dipergunakan pada terminal teletype dan teleprinter. Karena kombinasi ini terdiri dari 5 bit maka hanya terdiri dari 25 sampai 32 kombinasi dengan kode huruf dan gambar yang berbeda.
- Jika kode ini dikirim menggunakan transmisi serial tak sinkron, maka pulsa stop bit-nya pada umumnya memiliki lebar 1,5 bit. Hal ini berbeda dengan kode ASCII yang menggunakan 1 atau 2 bit untuk pulsa stop-bitnya.

# Tabel Kode Boudot

| Kode | Karakter Letter | Karakter Figure |
|---|---|---|
| 11000 | A | - |
| 10011 | B | ? |
| 01110 | C | : |
| 10010 | D | $ |
| 10000 | E | 3 |
| 10110 | F | ! |
| 01011 | G | & |
| 00101 | H | # |
| 01100 | I | 8 |
| 11010 | J | ‘ |
| 11110 | K | ( |
| 01001 | L | ) |
| 00111 | M | . |
| 00110 | N | , |
| 01101 | P | 0 |
| 11101 | Q | 1 |
| 01010 | R | 4 |
| 10100 | S | BELL |
| 00001 | T | 5 |
| 11100 | U | 7 |
| 01111 | V | ; |
| 11001 | W | 2 |
| 10111 | X | / |
| 10101 | Y | 6 |
| 10001 | Z | “ |
| 11111 | LTRS | LTRS |
| 11011 | FIGS | FIGS |
| 00100 | SPC | SPC |
| 00010 | CR | CR |
| 01000 | LF | LF |

# ASCII Code

- Kode ASCII memiliki 128 bit kombinasi yang selalu digunakan. Dari 128 kombinasi tersebut 32 kode diantaranya digunakan untuk fungsi-fungsi kendali seperti SYN, STX. Sisa karakter lain digunakan untuk karakter-karakter alphanumerik dan sejumlah karakter khusus seperti =, / . ?
- Pada dasarnya kode ASCII merupakan kode alfanumerik yang paling popular dalam teknik komunikasi data. Kode ini menggunakan tujuh bit untuk posisi pengecekan bit secara *even* atau *odd parity.*

| Dec | Hx | Oct | Char | Dec | Hx | Oct | Html | Chr | Dec | Hx | Oct | Html | Chr | Dec | Hx | Oct | Html | Chr |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 000 | NUL (null) | 32 | 20 | 040 |   | Space | 64 | 40 | 100 | @ | @ | 96 | 60 | 140 | ` | ` |
| 1 | 1 | 001 | SOH (start of heading) | 33 | 21 | 041 | ! | ! | 65 | 41 | 101 | A | A | 97 | 61 | 141 | a | a |
| 2 | 2 | 002 | STX (start of text) | 34 | 22 | 042 | " | " | 66 | 42 | 102 | B | B | 98 | 62 | 142 | b | b |
| 3 | 3 | 003 | ETX (end of text) | 35 | 23 | 043 | # | # | 67 | 43 | 103 | C | C | 99 | 63 | 143 | c | c |
| 4 | 4 | 004 | EOT (end of transmission) | 36 | 24 | 044 | $ | $ | 68 | 44 | 104 | D | D | 100 | 64 | 144 | d | d |
| 5 | 5 | 005 | ENQ (enquiry) | 37 | 25 | 045 | % | % | 69 | 45 | 105 | E | E | 101 | 65 | 145 | e | e |
| 6 | 6 | 006 | ACK (acknowledge) | 38 | 26 | 046 | & | & | 70 | 46 | 106 | F | F | 102 | 66 | 146 | f | f |
| 7 | 7 | 007 | BEL (bell) | 39 | 27 | 047 | ' | ' | 71 | 47 | 107 | G | G | 103 | 67 | 147 | g | g |
| 8 | 8 | 010 | BS (backspace) | 40 | 28 | 050 | ( | ( | 72 | 48 | 110 | H | H | 104 | 68 | 150 | h | h |
| 9 | 9 | 011 | TAB (horizontal tab) | 41 | 29 | 051 | ) | ) | 73 | 49 | 111 | I | I | 105 | 69 | 151 | i | i |
| 10 | A | 012 | LF (NL line feed, new line) | 42 | 2A | 052 | * | * | 74 | 4A | 112 | J | J | 106 | 6A | 152 | j | j |
| 11 | B | 013 | VT (vertical tab) | 43 | 2B | 053 | + | + | 75 | 4B | 113 | K | K | 107 | 6B | 153 | k | k |
| 12 | C | 014 | FF (NP form feed, new page) | 44 | 2C | 054 | , | , | 76 | 4C | 114 | L | L | 108 | 6C | 154 | l | l |
| 13 | D | 015 | CR (carriage return) | 45 | 2D | 055 | - | - | 77 | 4D | 115 | M | M | 109 | 6D | 155 | m | m |
| 14 | E | 016 | SO (shift out) | 46 | 2E | 056 | . | . | 78 | 4E | 116 | N | N | 110 | 6E | 156 | n | n |
| 15 | F | 017 | SI (shift in) | 47 | 2F | 057 | / | / | 79 | 4F | 117 | O | O | 111 | 6F | 157 | o | o |
| 16 | 10 | 020 | DLE (data link escape) | 48 | 30 | 060 | 0 | 0 | 80 | 50 | 120 | P | P | 112 | 70 | 160 | p | p |
| 17 | 11 | 021 | DC1 (device control 1) | 49 | 31 | 061 | 1 | 1 | 81 | 51 | 121 | Q | Q | 113 | 71 | 161 | q | q |
| 18 | 12 | 022 | DC2 (device control 2) | 50 | 32 | 062 | 2 | 2 | 82 | 52 | 122 | R | R | 114 | 72 | 162 | r | r |
| 19 | 13 | 023 | DC3 (device control 3) | 51 | 33 | 063 | 3 | 3 | 83 | 53 | 123 | S | S | 115 | 73 | 163 | s | s |
| 20 | 14 | 024 | DC4 (device control 4) | 52 | 34 | 064 | 4 | 4 | 84 | 54 | 124 | T | T | 116 | 74 | 164 | t | t |
| 21 | 15 | 025 | NAK (negative acknowledge) | 53 | 35 | 065 | 5 | 5 | 85 | 55 | 125 | U | U | 117 | 75 | 165 | u | u |
| 22 | 16 | 026 | SYN (synchronous idle) | 54 | 36 | 066 | 6 | 6 | 86 | 56 | 126 | V | V | 118 | 76 | 166 | v | v |
| 23 | 17 | 027 | ETB (end of trans. block) | 55 | 37 | 067 | 7 | 7 | 87 | 57 | 127 | W | W | 119 | 77 | 167 | w | w |
| 24 | 18 | 030 | CAN (cancel) | 56 | 38 | 070 | 8 | 8 | 88 | 58 | 130 | X | X | 120 | 78 | 170 | x | x |
| 25 | 19 | 031 | EM (end of medium) | 57 | 39 | 071 | 9 | 9 | 89 | 59 | 131 | Y | Y | 121 | 79 | 171 | y | y |
| 26 | 1A | 032 | SUB (substitute) | 58 | 3A | 072 | : | : | 90 | 5A | 132 | Z | Z | 122 | 7A | 172 | z | z |
| 27 | 1B | 033 | ESC (escape) | 59 | 3B | 073 | ; | ; | 91 | 5B | 133 | [ | [ | 123 | 7B | 173 | { | { |
| 28 | 1C | 034 | FS (file separator) | 60 | 3C | 074 | < | < | 92 | 5C | 134 | \ | \ | 124 | 7C | 174 | | | \| |
| 29 | 1D | 035 | GS (group separator) | 61 | 3D | 075 | = | = | 93 | 5D | 135 | ] | ] | 125 | 7D | 175 | } | } |
| 30 | 1E | 036 | RS (record separator) | 62 | 3E | 076 | > | > | 94 | 5E | 136 | ^ | ^ | 126 | 7E | 176 | ~ | ~ |
| 31 | 1F | 037 | US (unit separator) | 63 | 3F | 077 | ? | ? | 95 | 5F | 137 | _ | _ | 127 | 7F | 177 | &#127; | DEL |

Source: www.LookupTables.com

# Unicode

- Orang-orang di negara-negara yang berbeda menggunakan karakter berbeda untuk menuliskan kata-kata dalam bahasa ibu mereka. Sekarang ini kebanyakan aplikasi, mencakup sistem email dan web browser, menggunakan sistem 8 bit yang mana mereka dapat beroperasi yang tepat sesuai ketentuan, seperti ISO-8859-1.
- Unicode memiliki lebar per karakter sebesar 20 bit. Akan menjadi boros jika kita mengirim data Unicode yang berisi teks huruf Latin menggunakan 20 bit per karakter. Oleh karena itu maka Unicode ditransformasikan terlebih dahulu menjadi UTF-8 atau UTF-16 (*Unicode Transformation Format*) dengan UTF-8 maka karakter-karakter pada U+0000 (Notasi U+abcd) digunakan untuk mengacu pada karakter bernomor abcd pada tabel Unicode.

- Pada dasarnya ada 4 cara untuk mengkodekan karakter Unicode, yaitu:
  1. UTF-8: 128 karakter digunakan untuk mengkode 1 byte (karakter ASCII). 1.920 karakter digunakan mengkode 2 byte (untuk karakter Roma, Yunani, Cyrilic, Coptic, Armenian, Ibrani dan Arab). 63.488 karakter digunakan untuk mengkde 3 byte (Cina dan Jepang). 247.418.112 karakter yang lain, yang belum digunakan, dapat digunakan untuk mengkpde 4, 5, 6 karakter.
  2. UCS-2: Tiap-tiap karakter direpresentasikan oleh 2 byte. Pengkodean ini digunakan untuk merepresentasikan 65.536 karakter Unicode yang pertama.
  3. UTF-16: Ini adalah perluasan dari UCS-2 dimana dapat direpresentasikan 1.112.064 karakter Unicode. 65.536 karakter Unicode yang pertama diwakili 2 byte, yang lainnya 4 byte.
  4. UCS-4: Tiap-tiap karakter direpresentasikan oleh 4 byte.

- Unicode Bahasa Armenian

