

第2章

情報の表現 記号・符号化

記号表現

▼ 記号表現

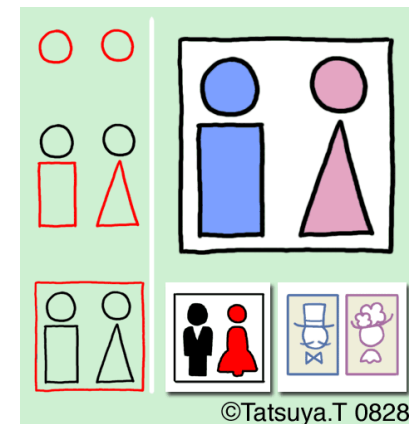
- ◆ 事物/事象,心的概念を抽象化したもの

▼ 記号表現の実際の形式

- ◆ 図記号(ピクトグラム), 数の表現など

▼ 記号が表す2側面

- ◆ 意味されるもの(シニフィエ)…(例)サービスエリア
- ◆ 意味するもの(シニフィアン)…(例)図記号



記号表現

↓ 記号表現

- ◆ 事物/事象, 心的概念を抽象化したもの

↓ 記号表現の実際の形式

- ◆ 図記号(ピクトグラム), 数の表現など

↓ 記号が表す2側面

- ◆ 意味されるもの(シニフィエ)…(例)サービスエリア
- ◆ 意味するもの(シニフィアン)…(例)図記号

サービスエリアの図記号



✦ 抽象化された図形によるデザイン

- ◆ 瞬時に表示内容を認識できる

✦ 記号表現とパターン表現の混在

- ◆ パターン表現は常に具体的/直接的であればいいわけではない

サービスエリアの図記号



↓ 抽象化された図形によるデザイン

- ◆ 瞬時に表示内容を認識できる

↓ 記号表現とパターン表現の混在

- ◆ パターン表現は常に具体的/直接的であればいいわけではない

図記号の修辞法

▼ 提喩に相当する表現方法

- ◆ ある事物を表現するのに、それと意味的包含関係にある事物を代わりに用いる比喻
 - ・ ナイフ, フォークの図でサービスエリアを表現する

▼ 隠喩に相当する表現方法

- ◆ GUIにおけるゴミ箱アイコンは「ゴミを捨てる」という行為の隠喩

日本語文字コード

- ✦ 文字と計算機上の符号(数値)を対応づけるための枠組み
- ✦ 現在, **JIS**,シフト**JIS**,**EUC**などの異なった日本語文字コードが混在している
- ✦ 解釈の枠組みが異なれば記号の意味が異なってしまう例
 - ◆ プログラムが想定するコード体系と異なると,「文字化け」が起こる
- ✦ コード体系の標準化・統一化には困難も多い
 - ◆ 歴史的経緯、利害関係、処理の都合、拡張性、文字セットの違い

コンピュータでの数の表現

- ▼ 「0」と「1」の2種類の記号を用いたビット列で表現される
- ▼ 表現できる数値はコンピュータに依る
 - ◆ 表現できる正の整数
 - ・ 16ビットのシステム: 0~65535までを表現できる
 - ・ 32ビットのシステム: 0~4294967295を表現できる

数値の表現方法

- ↘ 数値といっても色々ある（整数・有理数・小数点数・複素数）
- ↘ 「表現」する際には精度が問題となる
 - ◆ 無限に大きな数、無限に細かい小数を「書き表わす」には、無限の文字が必要
- ☒ まずは、有限の正の整数を考える
（例：0から10000までの数）

アナログ表現とデジタル表現

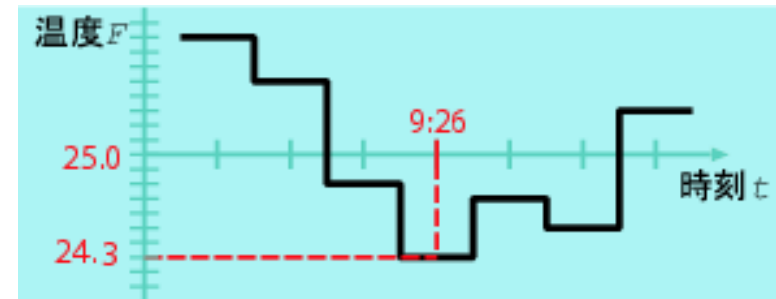
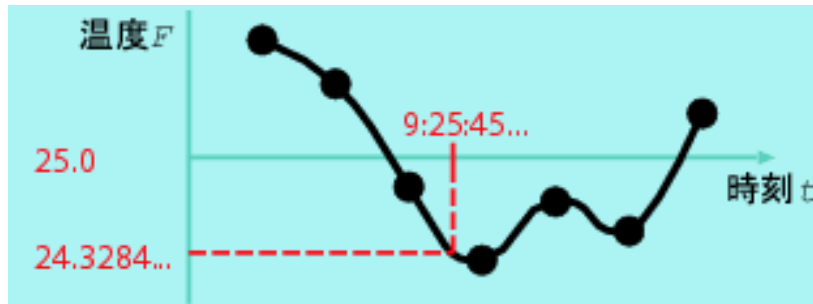
▼ アナログ表現

- ◆ ある情報を連続量(アナログ量)として表すこと
- ◆ 無限の精度を必要とするため, データの複製は元のデータの近似にしかない

▼ デジタル表現

- ◆ ある情報を離散的に表すこと(デジタル量)
 - ある情報に対して一定の間隔の尺度を導入し, その尺度の値に近似して表現する
- ◆ 複製時にデータが劣化しにくい
- ◆ 情報コンテンツの著作権保護への問題をもたらす

アナログ表現とデジタル表現の実際



- ✚ (左の図) 気温のアナログ表現
- ✚ (右の図) 気温のデジタル表現
- ✚ アナログ量をデジタル量に変換する際には、情報を離散化する間隔を選択し、表現する必要がある
 - ◆ 離散化する2つの軸 → 量子化, 標本化
 - ・ 量子化 測定値をある間隔ごとに表現する。
 - ・ 標本化 一定時間感覚ごとの計測。

標本化定理(シャノン)

- ↘ 求められる情報の精度から、標本化に必要な間隔を決める定理
- ↘ 標本化の対象となるアナログ関数 F が、周波数の異なる複数の周期関数の重ね合わせで表現できる事を基礎とする
- ↘ F を表すために必要な周期関数(周期 T , 周波数 $\omega = 1/T$)の周波数が W 以下であるとする、 $1/2W$ の間隔で標本化すれば、標本化された値から元のアナログ関数 F を完全に復元できる

▼ エイリアシング 画像の標本化において標本化間隔 D の $1/2$ より高い周波数成分が原画像に含まれる場合、標本化された画像中に原画には存在しなかった空間周波数成分が含まれたり、空間周波数スペクトルがゆがんだりする現象。

(「岩波情報科学辞典」より)

エイリアシングの例

- ✦ 縞模様の間隔と、
標本化の周期が近い所に注目



標本化の実際

▼ 電話の標本化の例

- ◆ 会話をするのが目的
 - 標本化周波数: 8kHz
 - ナイキスト周波数: 4kHz

周期関数への分解



▼ フーリエ解析

- ◆ 与えられた信号を個々の異なる周波数成分の波に分解

▼ 音声や画像などの情報表現と圧縮に用いる

- ◆ (図)画像の低周波成分から高周波成分へと足し合わせていったもの. 復元に必要な情報量が分かり, データ量を圧縮できる

デジタル符号化(1)

↘ 2進符号

- ◆ 10進数を2進数に変換したもの

↘ ハミング距離

- ◆ 2つの符号間で対応する桁の記号が異なる個数
 - ・ (0000)と(0001)では1
 - ・ (0011)と(0100)では3
- ◆ 2進符号では数値の差とハミング距離が一致しない

JPEG圧縮

- ↘ 画像データの圧縮方法(非可逆圧縮)
- ↘ 要求される精度の周波数成分までを符号化する
 - ◆ 人間が必要としない高周波成分に対する情報を切り落とすことでデータ量を圧縮