2018年度　Sセメスター　情報(文系用)　シケプリ

1、このシケプリは、文系用に作られたものであり、理系とは試験範囲が異なる。
2、試験は例年大問3つで、前の2つが理系との共通問題、後ろの1つが文系専用問題(実際の試験ではA、Bと選択形式になっておりAを選ぶことになる)である。
3、このシケプリでは、授業内容であっても試験範囲ではない部分はすべて除外し授業内容でなくても試験範囲である部分はすべて含めてある。なお教科書(山口和紀『情報 第2版』東京大学出版会、2017)から適宜引用している。
4、このシケプリを読む前に以下の「必修学習項目」に必ず目を通すこと。なお試験範囲は赤字と緑字の部分であり、青字と黒字の部分はこれに該当しない。

必修学習項目

「第１章」（学習内容の概観）

１、情報とは

事物のあり方の変化や、構成状態の変化に関連する概念

２、情報の多面性

A 人間に関わる側面→情報の表現、伝達、理解が求められる側面

B 問題解決に関わる側面→様々なデータの収集・分析・比較・評価が求められる側面

C 社会に関わる側面→コンピュータ、ネットワーク、情報システムの性質と振舞いについて、
自分自身で正しい理解をしておくことが求められる側面

３、情報の表現

表現の対象、表現の目的、表現の方法

（＋αとして）人間の認知に関する事項、機械処理に関する技術的諸側面

４、情報の伝達

情報理論(クロード・シャノン)→今日の情報量の基礎

プロトコル→送り手と受け手の間に成立している共通の理解、情報伝達の前提条件

５、モデル化

モデル→さまざまな状況で使われる代替物

モデル化→モデルを作ること

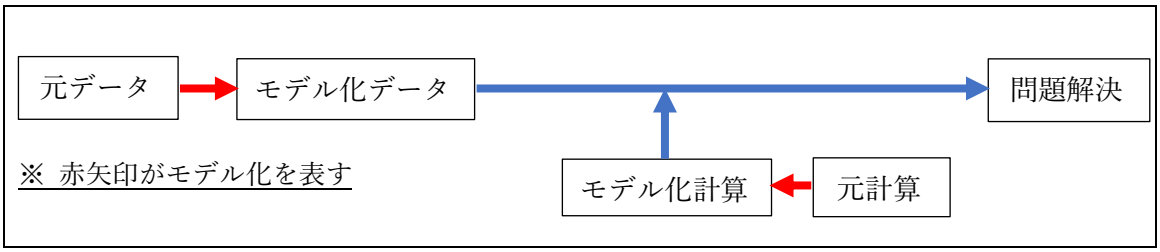
データのモデル→さまざまな“ものおよびその状態”の代替

計算のモデル→モデル化されたデータを操作する　cf.プログラム言語

６、問題解決

問題解決→様々な状況をモデル化し表現することで、対象とする“問題”を処理できる状況を作る

計算の意味(→元の問題と解決状況との対応関係)が明確であるよう注意することが求められる。



アルゴリズム→ある問題を解決するために計算モデルの上で構築される“やり方”

計算量→ある問題をあるやり方で処理する場合に必要な手間、アルゴリズムに依存

７、計算の機構

A コンピュータ→今まで数多くの自動機械が行なっていたことを一つにまとめた機械
「モデル化→処理→再解釈」の流れを通じて問題解決を行う

B ２進数モデル→理論としてブール代数、実現素子として論理回路を使用

コンピュータは、データモデルと共に計算モデル(計算要素としてのプログラム)も２進数モデルで表現することにより、内部構造を変えることなく問題解決できる→プログラム内蔵方式

８、情報システム

ソフトウェア→プログラムを集めて全体として一つのサービス・業務を行うようにしたもの

情報システム→ソフトウェアを含む情報処理機器と情報伝達のためのネットワークを組み合わせることで、様々なサービスや機能を提供するような複合的システム

９、ユーザインタフェース

ユーザインタフェース→利用者と情報システムが関わる部分

ユニバーサルデザイン→広範囲の知識水準・能力を持つユーザを対象としたシステムデザイン

１０、社会との関わり

情報システムは既成の枠組みと相容れないことも多い

むしろ情報システムが既成の枠組みを脅かすことすらある

→情報社会人の基本的素養としての情報についての知識・経験が重要

「第２章」（情報システム）

１、情報システムの実例

例　ポケモン GO

複数のシステムの連携(GPS、ジャイロセンサー、カメラシステム、データプロバイダなど)

拡張現実感(AR)の応用→先端的な研究成果の活用

モバイル端末とネット上のサーバの組み合わせ→ユビキタスコンピューティング

ビッグデータの処理

cf. 先端的な研究成果の活用例→人工知能（AI、artificial intelligence）

２、集中と分散

A 集中型→中央集権的に全体の制御処理を司るホストコンピュータを中心にして、利用者が端末からホストコンピュータに接続し情報システムを使う形式

B 分散型→ネットワーク上に分散した多くのプロセッサがそれぞれ処理を行い、互いが必要な情報を交換して全体としての整合をとっていく形式

集中型の利点/分散型の欠点	全体の整合性の確保が、効率よくできる/効率よくできない。
集中型の欠点/分散型の利点	柔軟性や拡張性が、低い/高い。

３、クライアントサーバモデル

サーバ→ネットワーク上でサービスを提供するプログラム、又はそれを有したコンピュータ

クライアント→サーバに対してサービスを要求するプログラム

クライアントサーバモデルは、ネットワーク上にサーバとクライアントという役割が分散している点では分散型だが、サーバにサービス提供の機能が集中している点では集中型とも言える

４、クライアントサーバ間のやり取り

URL→あるページの場所を直接指定する記号表記

HTTP→クライアントサーバ間で従われるプロトコル(通信規約)、GET・POST など

「第３章」（情報の表現）

１、情報の表現

A1 自然言語→特別な訓練なしに日常的に使用している言語

A2 人工言語→プログラミング言語など人工的に作られた言語

B1 手続き的表現→時間を追った手順に基づく

B2 宣言的表現→対象間の関係・対象の属性に基づく

C1 記号表現→与えられた記号の集合とそれらを解釈する規則体系に基づく

C2 パターン表現→構成要素間の時空間的パターンに基づく

２、情報の表現とモデル

モデル→単純化・抽象化された事物・事象・概念

モデル化→実際の事物・事象に対応したモデルを構築する過程

モデルの表現形式は目的に応じて多様である

A 表

B 図→ベン図、オイラー図　cf. 認知科学における図的推論

C グラフ→ノード(頂点)とエッジ

エッジの種類→ラベル付きエッジ(ラベル付きグラフ)、有向エッジ(弧)

グラフを用いた例→組織図、PERT 図(作業工程)、意味ネットワーク(人の記憶構造)

３、情報の表現における注意

受信側→多くの表現の適切な解釈が求められる

発信側→多くの表現手段の適切な選択が求められる

A 表現の対象

B 表現の目的

C 表現の方法

４、記号と表現

A 図記号(ピクトグラム)→パターン表現と記号表現が混在

提喩(ある事物の表現にそれと意味的包含関係にある事物を代用)

隠喩(ある事物の特徴を直接別の事物で代用表現)

記号の恣意性

→情報の受信者の解釈の枠組みへの配慮が重要

→情報の受信者自身の情報表現のコンテクスト的解釈も重要

cf. コンピュータでの文字表現→文字コードでの対応づけ、標準化・統一化が重要

例　ASCⅡ(アルファベット)、UTF-8・JIS・シフト JIS・EUC-JP(日本語)、Unicode

B 数表現→情報表現間のトレードオフ(ある側面が優先される一方別の側面では問題が生じる関係のこと)に対する配慮が重要　　例 ローマ数字とアラビア数字の表記

ビット列による表現(0 と 1 しか用いられない)

５、アナログとディジタル

アナログ表現→ある情報を連続量として表すこと

ディジタル表現→ある情報に一定間隔の尺度を導入し元の値をそれに近似し離散的に表すこと

アナログ量→アナログ表現で表されたもの

ディジタル量→ディジタル表現で表されたもの

アナログ量(表現)	→→→	ディジタル量(表現)
無限の精度が必要 データの劣化が不可避 正確な復元が困難	もともとアナログ量が持つ連続的情報が切り捨てられる	精度が有限 データが劣化しにくい 誤り検出・訂正が容易

量子化→アナログ量をある一定間隔の離散的な表現に変換すること、目盛りを導入すること

標本化→アナログ量をある抽出間隔をもってディジタル量に変換すること

cf. 標本空間→標本化対象となる情報が定義される時間や領域のこと

６、ディジタル符号化

２進数による２進符号化

グレイ符号による２進符号化

cf. グレイ符号→値が隣接する符号間で対応桁の異なる 0,1 の個数（ハミング距離）が常に 1

圧縮可能性→符号化などの工夫により、より少ないデータ量で同等の情報を表現できること

可逆圧縮・非可逆圧縮　　cf. JPEG 圧縮　　cf.2 遺伝的アルゴリズム

７、情報量

情報理論→すべての情報に共通するような普遍的性質を捉え情報伝達を定量的に評価

情報伝達は受信側の状態変化に依存する

受信側の状態変化とはメッセージによる選択肢の減少である

情報量は「事前の場合数」と「事後の場合数」を使ってどう表されるべきか？

差→不適。なぜなら、例えば 100→97 と 4→1 が同列の扱いになってしまうから。

商→不適。なぜなら、情報量の加法性を満たさないから。

cf. 情報量の加法性→ある情報の受信では段階の数によらず情報量は一定であるという性質

対数→適切。ゆえに、情報量の定義式は、ビット単位で、

$$-\log_2\{\frac{\text{事後の場合数}}{\text{事前の場合数}}\}=\log_2\{\frac{\text{事前の場合数}}{\text{事後の場合数}}\}$$

である。次にメッセージ全体の情報量について考えてゆく。これは期待値のように考えればよく、

$$(\text{平均情報量})=-\sum_{k=1}^{n} p_k(\log_2 p_k)=\sum_{k=1}^{n} p_k(\log_2 \frac{1}{p_k})$$

である。ただし上の式は n 個のメッセージを考えており、$p_k=\frac{\text{メッセージ}k\text{の事後の場合数}}{\text{メッセージ}k\text{の事前の場合数}}$ とする。

平均情報量の最大値は、$p_1,\cdots,p_n=1/n$ のときである。（つまりすべて等確率のときである）

平均情報量はエントロピーと呼ばれることもある。

８、符号化と平均情報量

復元可能なデータの圧縮を考える

メッセージ m_k の長さ l_k→あるメッセージの符号化によるビット数のこと

メッセージ m_k の確率 p_k→あるメッセージの出現確率のこと

平均符号長→(平均符号長)＝$\sum_{k=1}^{n} p_k l_k$

復元可能なデータの圧縮の限界は平均符号長の最小値である。ここで、平均符号長の最小値は、情報理論によると、メッセージ全体の平均情報量である。つまり、各メッセージをその情報量に等しい長さで符号化すれば、復元可能な範囲では、最もデータを圧縮できるということである。このことを情報源符号化定理という。

「第４章」（情報伝達と通信）

１、プロトコル

プロトコル→自分の意図を相手が理解し相手の意図を自分が理解するための決め事

プロトコルの階層性→階層ごとに独立に考えることができる、インターネット通信でも使用

URL→ウェブ文書を指定するための記号表記、文書を保持するウェブサーバのホスト名を含む

ブラウザ→ウェブ文書を表示するもの

HTTP→ブラウザ(クライアント)の要求、ウェブサーバ(サーバ)の応答におけるプロトコル

HTTPS→HTTP を拡張したプロトコル

２、安全な通信

HTTPS の接続の有効性の確認

盗聴の防止→HTTPS におけるデータの暗号化

「なりすまし」防止→認証

cf. フィッシング(偽のサイトを作り情報を盗もうとする攻撃)

→HTTPS において認証局が署名したディジタル証明書を通じてサーバを確認(→PKI)

他にも、生体認証、２段階認証、ワンタイムパスワードなど

３、インターネット

インターネットの大まかな仕組み

→小規模なネットワークがありその接続点にあたるルータを通じてネットワーク間で通信

ネットワーク内の通信→通信媒体による

ネットワーク間の通信→TCP/IP(共通プロトコル群)

TCP/IP の階層モデル(上位のプロトコルは下位のプロトコルを通じて実現)

アプリケーション層	HTTP、DNS	URL
トランスポート層	TCP	IP アドレス
インターネット層	IP	
ネットワークインタフェース層	（イーサネット）	(MAC アドレス)

A　DNS→ホスト名(文字による表記)と IP アドレス(通信の宛先としての表記)を結びつける
　　ホスト名と IP アドレスとの対応はドメインごとに分散して管理する
　　IP アドレスを調べる際には反復問合せを行う
　　反復問合せによる結果を再利用できるように一定時間記憶しておく(キャッシュ)
B　IP アドレス→32 ビット、読む際には 8 ビットずつ区切って 10 進数に変換して読む
　　ネットワークアドレス(IP アドレスの上位ビット)
　　ホスト番号(IP アドレスの下位ビット)
cf.1 DHCP→機器を接続したときに自動で適切な IP アドレスを割り振るプロトコル
cf.2 ポート番号→同じ機器内でアプリケーションを区別するための数値(16 ビット)
C　TCP→仮想的に双方向に通信可能な通信路を設定
　　送信側→データをパケットに分割、TCP ヘッダ付与(シーケンス番号、誤り検出符号)
　　→カプセル化
　　受信側→パケットを並びかえ結合しデータを復元
　　確認応答→パケット消失に備え送信元に再送や速度調整を促す
cf. UDP→再送を行わない
D　IP→IP パケットをどのルータを通じて適切に配送すればよいかという経路制御が目的
　　静的経路制御→経路表をあらかじめ決めておく
　　動的経路制御→経路表を自動的に更新してゆく
cf. 防火壁(ファイアウォール)→パケットの選択・通信制御

<u>４、機密性と認証</u>

機密性保持のための暗号技術、改ざん・否認防止や認証のためのディジタル署名
A 暗号化
平文(元データ)を暗号文にすること(逆の作業は復号)、計算手順と鍵を用いる
共通鍵暗号→暗号化と復号が同じ鍵
公開鍵暗号→暗号化と復号が別の鍵、公開する方の鍵が公開鍵、別の方の鍵が秘密鍵
B ディジタル署名
一方向ハッシュ関数→ある文書を特徴づける数値の計算、この計算結果を電子指紋と言う
ディジタル署名→電子的な署名の技術、電子指紋を秘密鍵で暗号化し公開鍵で復号する
公開鍵への署名→公開鍵の検証、第三者の公開鍵へのディジタル署名が公開鍵の信頼に
PGP→共通の知り合いの署名が輪となって仲介することで公開鍵を信頼するという形の認証
C PKI
ディジタル証明書→ウェブサイトが本物であることを示すために提示するもの
認証局→ディジタル証明書を発行しその発行に際しディジタル署名を行なっている
ルート証明書→ウェブブラウザにあらかじめ登録されている信頼性の高いディジタル証明書
PKI(公開鍵暗号基盤)→あるディジタル証明書のデジタル署名からたどりルート証明書にディジタル署名のある認証局の署名が得られれば信頼性を認めるモデル

「第５章」(計算)

１、計算の一例－計数－

集合の要素数の計算→取り出し型(一つずつ数える)、分割型(集合を分割して考える)

２、計算の記述

変数→変化していく値

変数名→変数そのものの名称

代入→変数に値を設定する操作、「変数名←式」の形で表す

逐次処理の重要性

A 条件付き処理

B 反復処理→添字つき変数を用いて複数の値を一列に並べる「配列」を利用

３、アルゴリズム

アルゴリズム→計算手順

プログラム→アルゴリズムに基づき記述されるもの

→アルゴリズムとプログラムは別物、アルゴリズム次第でプログラムの実行時間は変化する

４、アルゴリズムの例－平方根－

※精度→小数位のこと、精度 0,0001 ならば小数第四位まで求める

A 解の候補から絞り込む

B 解の存在範囲から絞り込む→二分法

C 推測値から絞り込む→ニュートン・ラフソン法

精度 0.0001 で$\sqrt{2}$を求める際に、A は 14142 回、B は 15 回、C は 3 回の反復を要する

→アルゴリズムによってプログラムの実行時間は大きく変化することがわかる

５、計算量

計算量→アルゴリズムを基にしたプログラムの実行時間の見積もり

計算量のオーダーと呼ばれる大まかな尺度を用いて考える

４の例を再び考える

A は、もとの数$\sqrt{x}$と精度δに対して反復回数は$\sqrt{x}/\delta$より、計算量のオーダーは$\sqrt{x}/\delta$

B は、反復回数 n が $x/2^n < \delta$ を満たす形になるので、計算量のオーダーは$\log_2 x/\delta$

C は、計算量のオーダーを求めることは難しい(が B よりもよい)

計算量のオーダー(「何に比例するか」の尺度)からも、アルゴリズムの重要性は見えてくる

「第６章」(計算理論) は試験範囲外のため省略する。

「第７章」(データの扱い)

１、データモデル

データ→コンピュータの処理対象となる符号化された情報

データモデル→データを体系的に扱うためのモデル

２、ネットワークモデル

グラフ(ノードとエッジで構成される)のような形でつながり方を表すモデル

経路→順にたどってゆけるエッジの列

オイラー路→すべてのエッジを重複なくたどる経路

ウェブをネットワークモデルとして見ることができる(有向エッジを用いる)

３、階層モデル

木構造→分類を表すときに用いる根を中心に枝分かれする構造、部分では部分木と言う

階層モデル→木構造のように階層的に表されるモデル

住所のある部分に至るまでの名称はパス名として住所を指定し区別する

例　「港区」の前に書かれる「東京都」「名古屋市」「大阪市」のことをパス名と言い、これらパス名によって「東京都港区」「名古屋市港区」「大阪市港区」は各々指定され区別される

「第８章」(コンピュータ)

１、プログラム内蔵方式

ソフトウェア→プログラムとデータ両方の情報のこと

ハードウェア→ソフトウェアを処理する物理的な機構

プログラム内蔵方式→メモリ上にプログラムとデータを保持しプログラムに従い計算する方式

これを有するコンピュータをフォン・ノイマン型コンピュータと言う

２、コンピュータの基本構成

A 演算装置(データの計算処理)

B 主記憶装置(データ・プログラムの記憶)→情報の選択にはアドレスを使用

C 制御装置(演算装置の駆動、主記憶装置へのデータの読み書き)

A、C を中央処理装置(CPU)と言い、その半導体集積回路はマイクロプロセッサ(MPU)と言う

cf. 演算レジスタ→データ保持の装置、かつてはアキュムレータとも呼ばれた

「第９章」(インタフェース)

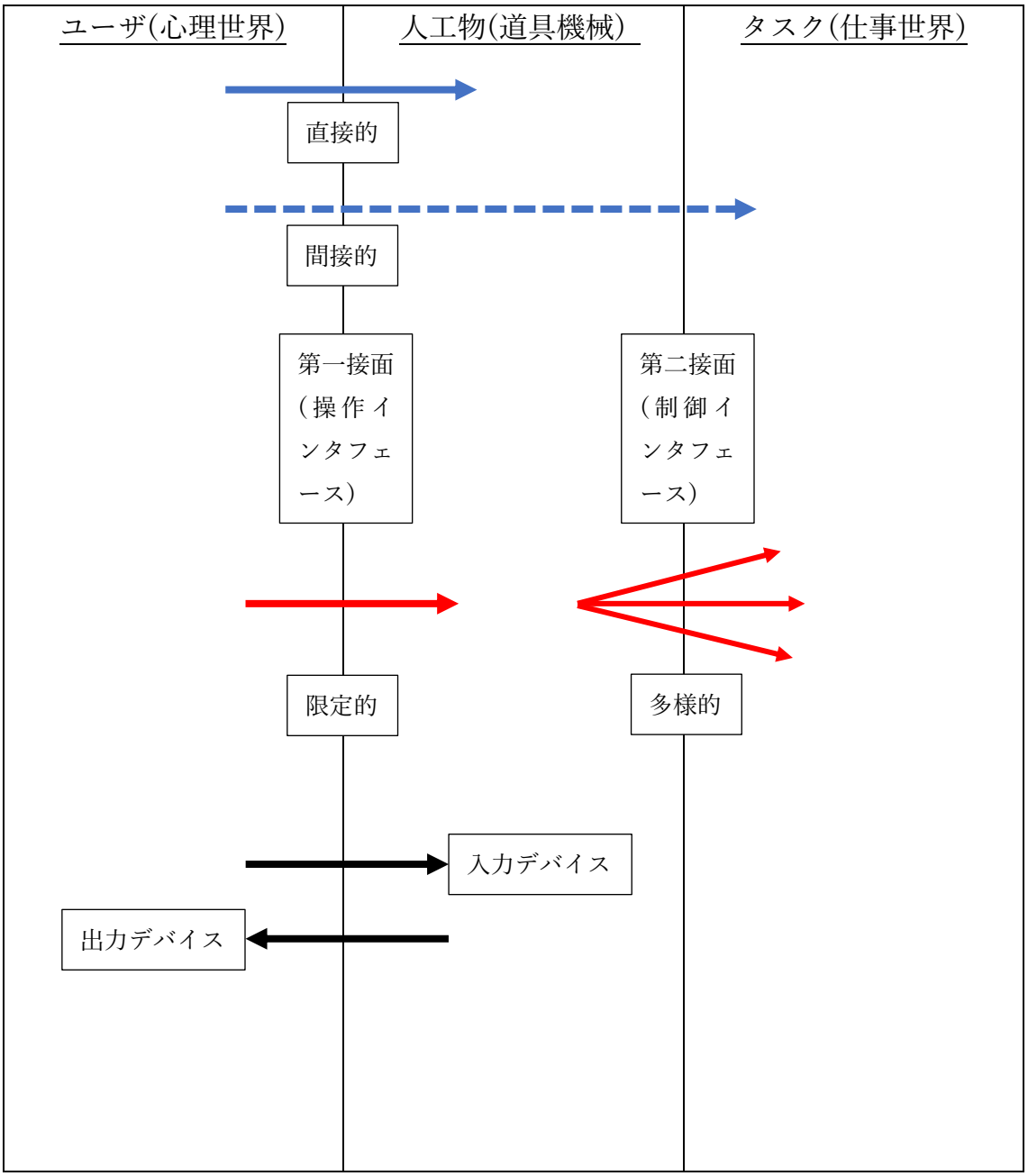
１、インタフェースとは

境界面、コンピュータ(人工物)とユーザ(人間)の間、ユーザ(ヒューマン)インタフェース

ユーザはインタフェースを通じてコンピュータを操作する以上人工物が機能を最大限発揮するには使いやすいインタフェースの存在が不可欠

通例、道具と対象への働きかけは一対一だが、今回はそうではない(等価に扱えないことも多い)

2、インタフェースの二重接面性



（上の図を通して理解することを求める）

補足

A 高度な人工物には大抵の場合二重接面性が存在している

B 第二接面において多様的な広がりを見せるのは、ソフトウェアの汎用性による

C 二重接面性が存在せず処理対象に直接働きかけるように操作可能なことを直接操作と言う

３、入出力デバイス

A 入力デバイス→キーボード(QWERTY 配列)

ポインティング・デバイス(直接入力型→タッチ○○、間接入力型→マウス)

cf. ダグラス・エンゲルバート→マウスの考案者

B 出力デバイス→ディスプレイ(ブラウン管(CRT)→液晶ディスプレイ(LCD))

プリンタ(近年の 3D プリンタ)

４、GUI と CUI

GUI→情報表示においてグラフィカルなオブジェクトを多用

ポインティングデバイス、直感的、直接操作

CUI→情報表示において文字を使用

キーボード、迅速性、文字列命令

５、GUI の歴史

構想から実用化(1940～60 年代)

アイバン・サザランド、アラン・ケイ

Mac、Windows などで採用

タッチパッド、タッチスクリーン、タブレット・コンピュータ(タブレット端末)

６、GUI の特徴

A　デスクトップメタファ→机上に書類を広げるような感覚

B　WIMP→ウィンドウ、アイコン、メニュー、ポインターを構成要素とする

cf. アフォーダンス→外界の環境や事物が生体の活動に供するべく持っている情報のこと

７、インタフェースの評価

ユーサビリティテスト

A パフォーマンステスト→システムの動作情報とユーザの操作履歴から評価

B ガイドライン法→既定の使いやすさのガイドラインを基準に評価

C モデル法→ユーザの行動をユーザがシステムを使う際の行動モデルに即して評価

D インスペクション法→複数の専門家による例から使い方を想定して問題点を発見

キーストローク・レベル・モデル

モデル法の一例→GOMS(Goal, Operator, Method, Selection rules)

GOMS-KLM(キーストローク・レベル・モデル)

→ユーザがコンピュータを使って特定の作業を行う際の作業時間を、その作業を構成する各操作をコンピュータが実行する時間の合計とみなす手法

(例 ファイルの削除→ポインタ移動、クリック、ドラッグ、クリック解除、ポインタ移動の合計)

フィッツの法則

デバイスの性能評価の手法、ある目標物までポインタを移動するのに要する時間を予測

ポインタ始点から目標物中心までの距離 D、目標物の大きさ W、ポインタの移動開始時間と停止時間を a、ポインタの移動速度を b とすれば、時間 $T = a + b\log_2(\frac{D}{W} + 1)$ と表せる

「第１０章」(情報技術と社会)

情報技術の基礎を知り、人間や社会への影響を理解して思考するための基礎知識を持つことは、現代を生きる私たちの必須の条件→「情報」を学ぶ意義

情報リテラシー→自分の操作の裏で何が動いているのかについてある程度論理的に考える能力

１、情報技術の影響

メディアの発展が社会体制・社会構造の転覆・変革を起こした事実がある(メディア論的思考)

A　インターネット技術がもたらした変化

「中央集権的→分散的」→場所制約、時間制約、経路制約、輸送コストからの解放

→コミュニケーション形態の変容(地理的要素への依存からの解放)

→権威構造の転覆(制度、法、経済、検閲、宗教、文化など、権力構造を越えた形)

B　SNS、GPS、ビッグデータがもたらした変化

炎上、サイバーカスケード現象(集団極性化、異質排除的傾向)

責任問題、プライバシー問題

深層学習、IoT の普及(インターネットに通信機能を備えた機器を接続)

２、社会への影響

A 無形性、複製可能性

手にすることができない→所有概念への影響

簡単に複製できる→権利概念への影響

B 広い通信射程、匿名性

３、社会への影響－境界の変容－

A 権利と所有の境界

利用者意識の希薄さ→知的財産権の問題

著作権法→プログラムやコンテンツ

プログラム→著作権法改正により保護

コンテンツ→DRM の問題(Winny の著作権侵害幇助問題)

B プライバシーとセキュリティの境界

プライバシー→保護されることがやはり重要、個人情報保護法(利用、管理、提供、開示)

セキュリティ→情報セキュリティ(情報システムにおける安全性の確保)

情報セキュリティの側面

a 機密性(認可された者にのみ情報がアクセスできる)

b 完全性(情報が正確かつ安全である)

c 使用可能性(必要なとき必要な情報資源にアクセスできる)

情報セキュリティの対策→未然対策、想定対策、事後対策、再発対策

不正アクセス禁止法は機密性に対する対策の一例

情報セキュリティ確保のための技術
個人認証技術→アクセスした人間がシステム利用を許可された利用権者かどうかを識別
　　アカウント(ID)、パスワードが有名
暗号化技術→元データを解読できないデータにする、暗号化⇆復号の流れ
公開鍵暗号方式→公開鍵と秘密鍵を用いたシステム
PKI→公開鍵暗号方式を利用したセキュリティインフラストラクチャのこと
電子署名→特定の文書の発信人を保証、なりすましの防止や本人確認が可能
国家のセキュリティ
セキュリティとプライバシーが対立することも多い
４、インターネット上の議論
A サイバースペース(バーチャルスペース)での議論における民主性
「多対多の通信」の観点、「情報という力」の観点、「機構からの脱却」の観点
以上のような観点から考えればインターネットは民主性を助長すると言える
だが現実は？
荒らし、匿名性を利用した無責任な書き込み、なりすましなど
→情報技術には民主性を高める可能性もある一方これを現実化するための規則の整備は必要
B サイバースペースは公共空間か共同体か？
公共性→アクセス可能性、異質価値、関心事の差異、多元的、社会運動
海外ではサイバースペースが公共空間である認識は強い
一方日本ではサイバースペースが共同体と化していることも多い
C ネットの功罪
大津市のいじめ自殺事件の例
インターネットの果たした役割→ネット炎上、全国的な議論、地域の権力構造の転覆
多対多の広い通信射程→遠隔地との短時間・容易なコミュニケーションが成立
匿名性→不可視性の感覚により通常行わないような行動を起こさせる可能性
複製可能性→問題の悪化・深刻化を助長
炎上、サイバーカスケード現象から生じる倫理的問題は以上の特徴から発生することが多い

以上