G2
多肢選択問題によるIRTに基づく評価手法の構築

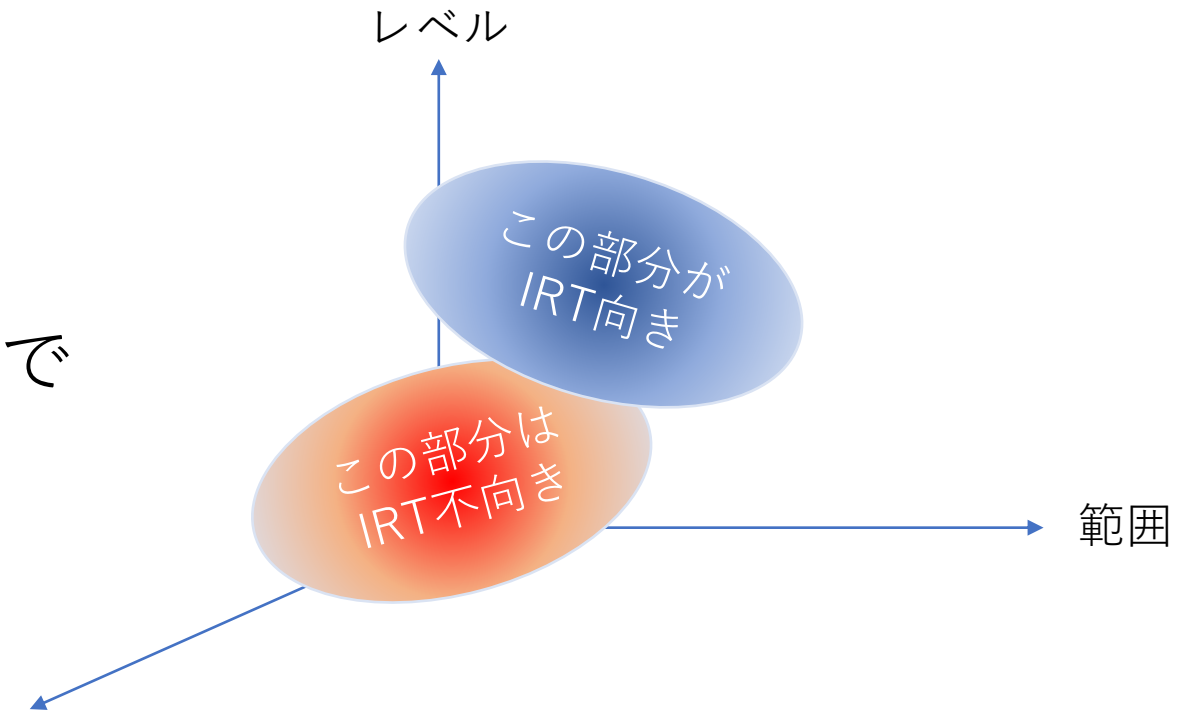
慶應義塾大学 環境情報学部

植原啓介 (kei@sfc.keio.ac.jp)

概要

- 目的
  - 多肢選択問題等の自動採点可能な問題による評価手法を開発する
  - 作問のBest Current Practiceとしての手順書を作成する

- 多肢選択問題でどれくらいの
  範囲（分野）、レベル（深さ）、
  思考力・判断力・表現力・応用力まで
  確認できるかを明らかにする。



多肢選択問題の作問指針

情報に関連する語句のリストがあるとした場合、下記のような作問の方法がある。

1. 説明文を読んで適切な句を選ばせる。（語句を知っている）
   同義語（あるいは反対語や非同義語）を選ばせる。
2. 語句の説明として正しいものを選ばせる。（語句を説明できる）
   語句があるクラスを表している場合はそのインスタンス（あるいは間違ったインスタンス）を答えさせる。
3. 語句の説明文の空欄を埋めさせる。（語句を説明できる）
4. 基数変換などの操作をさせる。（語句の内容を使える）
5. 知識を使って美しく解答を導く。（語句の内容を応用できる）

最近の活動

- EMIU情報模試2026冬に向けた作問
  - 前ページの作問指針仮説に基づいて、比較検討ができるような問題を作成
- 情報I 作題支援のための主要学習項目・キーワード集
- 自動作問エージェントの作成

情報I 作題支援のための主要学習項目・キーワード集

Yutaka Yasuda, Kyoto Sangyo University

IRT問題（小問×多数）作成時の疑問・不安

- 次はどんな問題を作ろうか？
  - 「ひらめき」に頼りがち
- どんな問題を作るか、より先に「何の理解を問いたいか」ではないか？
  - 「トリビアな知識」に寄せない
  - 「トリッキーな思考」に頼らない
- 出発点：「情報」で学ぶ「単元」集がほしい

単元集としての学習指導要領

・本質的にはそうなっているが、作題には使いにくい

(3) コンピュータとプログラミング

コンピュータで情報が処理される仕組みに着目し，プログラミングやシミュレーションによって問題を発見・解決する活動を通して，次の事項を身に付けることができるよう指導する。

ア　次のような知識及び技能を身に付けること。

(ｱ) コンピュータや外部装置の仕組みや特徴，コンピュータでの情報の内部表現と計算に関する限界について理解すること。

(ｲ) アルゴリズムを表現する手段，プログ
を活用する方法について理解し技能を身

(ｳ) 社会や自然などにおける事象をモデル
改善する方法について理解すること。

指導要領は
切り詰めすぎ

その解説は
細かすぎ

ア(ｱ) コンピュータや外部装置の仕組みや特徴，コンピュータでの情報の内部表現と計算に関する限界について理解することでは，コンピュータの特性を踏まえて活用するために，コンピュータの基本的な構成や演算の仕組み，オペレーティングシステムによる資源の管理と入力装置や出力装置などのハードウェアを抽象化して扱う考え方，コンピュータ内部でのプログラムやデータの扱い方，値の範囲や精度について理解するようにする。その際，ソフトウェアはオペレーティングシステムの機能を利用して動作していること，コンピュータでは定められたビット数のデータが扱われ，表現できる値の範囲や精度が有限であることで，計算結果は原理的に誤差を含む可能性があることなどを理解するようにする。

情報量としては充分だが具体性
が低く問題までの距離がある

単元集としての学習指導要領とその解説

- 本質的にはそうなっているが、作題には使いにくい
  - 具体的な問題を想像するためのキーワードが欲しい
  - →「情報科全教科書用語リスト」は「単元」の内容が浮かびにくい
- 適度な一覧性（縦覧性？）がほしい
  - 学習指導要領の各領域を適切にカバーする問題群を作りたい
  - 「次はこの領域からひとつ」「ここがまだ不十分」で方向を決めたい
- そうした組織的な検討をするためのツールが欲しい

作題用単元・用語リスト

- 学習指導要領解説を項目単位で短く
- キーワードを併記
  多くは教科書用語リストから抽出
- (0.89)などの数値は教科書用語リストの総意率
- 単元並びは網羅性が重要
- キーワードは網羅性を重視せず（有用性重視）

（3）コンピュータとプログラミング

ア（ア）・イ（ア）：コンピュータの仕組みと情報の内部表現

- **コンピュータの基本構成**：コンピュータは中央処理装置（CPU）、主記憶装置（メモリ）、補助記憶装置（ストレージ）、入力装置、出力装置などから構成され、それらの間でデータをやり取りしながら情報を処理していることを理解する。
  - データ(1.00) | 情報(0.92) | 補助記憶装置(0.89) | 主記憶装置(0.84) | CPU(0.83) | 入力装置(0.77) | 出力装置(0.77) | 制御装置(0.69) | 演算装置(0.69) | メモリ(0.66) | コンピュータ(0.63) | ハードディスク(0.56) | フラッシュメモリ(0.54) | 記憶装置(0.54) | 5大機能(0.47) | ストレージ(0.18) | 入力(0.18) | 出力(0.18)
- **CPUとメモリの役割**：メモリはコンピュータがデータを処理する作業場所として働き、CPUは機械語で表されたプログラムをデータとして読みながら実行することを理解する。
  - データ(1.00) | プログラム(1.00) | CPU(0.83) | メモリ(0.66) | コンピュータ(0.63) | 機械語(0.63) | レジスタ(0.61) | クロック周波数(0.56) | 記憶装置(0.54) | プログラムカウンタ(0.45) | マルチコアCPU(0.42)
- **ハードウェアとソフトウェアの関係**：コンピュータのハードウェアは、オペレーティングシステムやアプリケーションソフトウェアによる指示や制御の下で動作していることを理解する。
  - オペレーティングシステム(1.00) | ハードウェア(1.00) | ソフトウェア(0.92) | 基本ソフトウェア(0.80) | アプリケーションソフトウェア（アプリ）(1.00) | デバイスドライバ(0.54)

作成者判断で入れた要素

- 学習指導要領解説には言及がないが、重要と思われる項目は明示して補記
- 教科書用語リスト（＝索引語）には無いが教科書には登場するキーワードで作題に有用そうなものは（0.00）として補記

学習指導要領解説には登場しないが重要な項目

- **プログラムの基本要素**：データを扱う変数や配列と、処理を順番に実行する順次構造、条件によって処理を選ぶ分岐構造、処理を繰り返す反復構造を組み合わせてプログラムを構成する。
  - データ(1.00) | プログラム(1.00) | ソースプログラム（ソースコード）(0.72) | 変数(1.00) | 配列(0.83) | 反復構造(0.80) | 分岐構造(0.80) | 順次構造(0.80) | 演算子(0.77) | 比較演算子(0.56) | 論理演算子(0.42) | 代入(0.54) | 反復(0.37) | 順次(0.37) | 分岐(0.18)
- **論理回路**：論理演算と論理回路の関係を理解する。
  - 論理回路(0.72) | AND回路（ANDゲート）(0.72) | OR回路（ORゲート）(0.72) | NOT回路（NOTゲート）(0.72) | 半加算器(0.66) | 全加算器(0.54) | 真理値表(0.61) | 論理演算（ブール演算）(0.54)

- **デジタル画像の表現**：画像の解像度やファイル形式によって画質やファイルサイズが変化し、画質とファイルサイズにはトレードオフの関係があること、ラスタ形式（ビットマップ形式）では画像を画素（ピクセル）の集まりとして表現し、ラスタ形式とベクタ形式には異なる特性があり、用途に応じた使い分けが必要なことを理解する。
  - 符号化(1.00) | ビット(1.00) | バイト(0.92) | 画素（ピクセル）(1.00) | 解像度(1.00) | 階調(1.00) | トレードオフ(0.80) | 光の3原色(0.77) | 色の3原色(0.69) | ディジタル(0.72) | ファイル(0.66) | RGB(0.54) | CMY(0.37) | ベクタ形式(0.54) | ラスタ形式(0.54) | ファイル形式(0.18) | ビットマップ(0.00) | ファイルサイズ(0.00) | 画質(0.00)

（期待される）使い方と効果

- リストを眺めて問いたい単元を選び、用語をピックアップして作題
  → ヒラメキ駆動型からの脱却

- 不足している分野、過剰な分野などをリストを見て調整
  → バランス確認・抜け防止

- 何かの用語から問題を思いついたが、それは本来どのような文脈で（単元のなかで）問われるべきかを確認してから作題
  → トリビア問題化の抑止

- 「組織的な検討と、それに基づく作題」のツールになり得るか？

ご興味のある方、
お試し＋ご意見く
ださい

誤り指摘など大歓迎です

https://ylb.jp/joho1-topics08.html

（YLB.jp です）

AIによる作問システムSakuAの開発

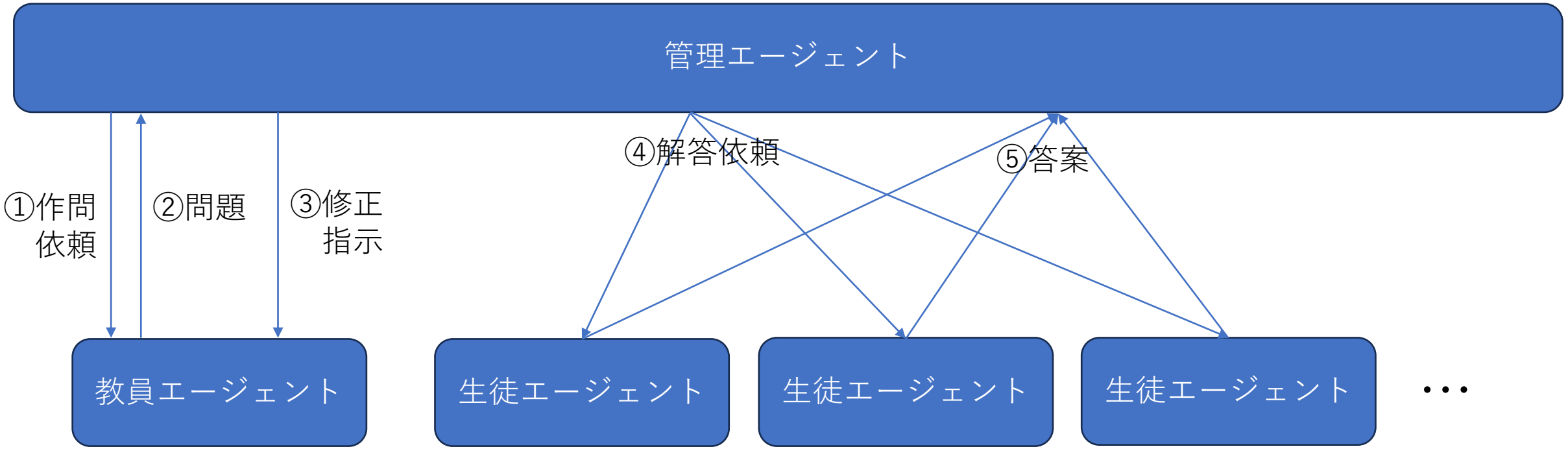
慶應義塾大学 環境情報学部
植原啓介 (kei@sfc.keio.ac.jp)

動機

- 「作問指針」と「情報I 作題支援のための主要学習項目・キーワード集」があれば、AIである程度作問できるのではないだろうか。

- アイディア
  - 教員エージェントが「作問指針」と「情報I 作題支援のための主要学習項目・キーワード集」を元に作問をする。
  - 生徒エージェントが問題を解いてみる。
  - 適切な問題ができるまで、上記を繰り返す。

システム概要


管理エージェント
①作問
依頼
②問題
③修正
指示
④解答依頼
⑤答案
教員エージェント
生徒エージェント
生徒エージェント
生徒エージェント
・・・

動作の様子

全分野からまんべんなく計 20 問（4択）を作成します。
目標正解率: 50.0% (生徒エージェント 10人の平均目標正解数: 10.0±1.0問 / 全 20問)

```
============================================================
--- サイクル 1 / 10 （問題作成 → レビュー → 解いてみる） ---
============================================================
[ステップ1: 問題作成] 教員エージェント（モデル: gemini-3.8-flash）が初期問題（Level 1〜2、4択）を作成中...
  -> 「（1）情報社会の問題解決」の問題（5問）を生成中...
  -> 「（2）コミュニケーションと情報デザイン」の問題（5問）を生成中...
  -> 「（3）コンピュータとプログラミング」の問題（5問）を生成中...
  -> 「（4）情報通信ネットワークとデータの活用」の問題（5問）を生成中...

[ステップ2: レビュー] 管理エージェントが作成された問題（20問）を審査中...
  [管理エージェント] 作成された問題（20問）をレビュー中...（レビュー 1/5回目）
    -> [差し戻し] 問007 [Lv.1] テーマ「メディアとコミュニケーション手段を活用したコンテンツ制作」: 問題ID 5と正解語句（アクセシビリティ）および問題文で問うている内容・定義が実質的に完全に同一であり重複しています。また本問のテーマ「メディアとコミュニケーション手段を活用したコンテンツ制作」のカリキュラム内容（目的や受け手に応じたメディア選択、コンテンツの設計・制作手順など）に沿っていません。コンテンツ制作における設計手順やメディア表現技法（ワイヤーフレーム、ストーリーボード、プロトタイピングなど）を問う問題に作り直してください。
  [管理エージェント] 望ましくない問題が 1 問見つかったため、差し戻して教員エージェントに作り直させます...（差し戻し 1/5回目）
  [管理エージェント] 作成された問題（1問）をレビュー中...（レビュー 2/5回目）
  [管理エージェント] レビュー完了: 全 1 問が適切な問題であると判定されました。

[ステップ3: 解いてみる] 10人の生徒エージェントが問題（20問）を解答中...
  [採点結果] 生徒エージェント(10人)の平均得点: 13.7 / 20 点 (平均正解率: 68.5% | 各生徒: [14点, 18点, 7点, 17点, 11点, 17点, 11点, 17点, 7点, 18点] | 目標: 10.0±1.0点)
  -> 目標得点帯に未収束のため、次のサイクル（2/10）へ進みます。

============================================================
--- サイクル 2 / 10 （問題作成 → レビュー → 解いてみる） ---
============================================================
[ステップ1: 問題作成（難化）] 前回の平均得点が目標より高いため、教員エージェントが正答率の高い問題を中心に 6 問を難化作成中...

[ステップ2: レビュー] 管理エージェントが作成された問題（6問）を審査中...
  [管理エージェント] 作成された問題（6問）をレビュー中...（レビュー 1/5回目）
  [管理エージェント] レビュー完了: 全 6 問が適切な問題であると判定されました。

[ステップ3: 解いてみる] 10人の生徒エージェントが問題（6問）を解答中...
  [採点結果] 生徒エージェント(10人)の平均得点: 12.6 / 20 点 (平均正解率: 63.0% | 各生徒: [13点, 18点, 8点, 16点, 10点, 16点, 10点, 17点, 4点, 14点] | 目標: 10.0±1.0点)
  -> 目標得点帯に未収束のため、次のサイクル（3/10）へ進みます。
```

動作の様子

```
============================================================
--- サイクル 5 / 10 （問題作成 → レビュー → 解いてみる） ---
============================================================
[ステップ1: 問題作成（難化）] 前回の平均得点が目標より高いため、教員エージェントが正答率の高い問題を中心に 3 問を難化作成中...

[ステップ2: レビュー] 管理エージェントが作成された問題（3問）を審査中...
  [管理エージェント] 作成された問題（3問）をレビュー中...（レビュー 1/5回目）
  [管理エージェント] レビュー完了: 全 3 問が適切な問題であると判定されました。

[ステップ3: 解いてみる] 10人の生徒エージェントが問題（3問）を解答中...
  [採点結果] 生徒エージェント(10人)の平均得点: 11.0 / 20 点 (平均正解率: 55.0% | 各生徒: [12点, 14点, 8点, 13点, 6点, 16点, 9点, 13点, 6点, 13点] | 目標: 10.0±1.0点)
  -> 目標の得点帯に収束しました！テスト作成を完了します。

--- 最終生成された問題の出題テーマと難易度一覧 ---
問001 [Lv.1] (想定正解率: 70% [7/10人]) [（1）情報社会の問題解決] テーマ: 情報に関する法規や制度
問002 [Lv.4] (想定正解率: 60% [6/10人]) [（1）情報社会の問題解決] テーマ: 文章や図への可視化
問003 [Lv.1] (想定正解率: 60% [6/10人]) [（1）情報社会の問題解決] テーマ: グループによる問題解決と合意形成
問004 [Lv.4] (想定正解率: 30% [3/10人]) [（1）情報社会の問題解決] テーマ: 情報技術が社会に及ぼす影響
問005 [Lv.1] (想定正解率: 40% [4/10人]) [（1）情報社会の問題解決] テーマ: 誰もが利用できる情報技術
問006 [Lv.5] (想定正解率: 40% [4/10人]) [（2）コミュニケーションと情報デザイン] テーマ: 情報の表現方法の選択
問007 [Lv.3] (想定正解率: 60% [6/10人]) [（2）コミュニケーションと情報デザイン] テーマ: メディアとコミュニケーション手段を活用したコンテンツ制作
問008 [Lv.4] (想定正解率: 60% [6/10人]) [（2）コミュニケーションと情報デザイン] テーマ: 色の表現と配色
問009 [Lv.2] (想定正解率: 40% [4/10人]) [（2）コミュニケーションと情報デザイン] テーマ: コミュニケーション手段の特徴
問010 [Lv.2] (想定正解率: 70% [7/10人]) [（2）コミュニケーションと情報デザイン] テーマ: 情報デザインに基づく表現・制作
問011 [Lv.2] (想定正解率: 80% [8/10人]) [（3）コンピュータとプログラミング] テーマ: モデル化の長所と限界
問012 [Lv.4] (想定正解率: 50% [5/10人]) [（3）コンピュータとプログラミング] テーマ: 意図しない計算結果の検討
問013 [Lv.4] (想定正解率: 60% [6/10人]) [（3）コンピュータとプログラミング] テーマ: プログラムの基本要素
問014 [Lv.2] (想定正解率: 50% [5/10人]) [（3）コンピュータとプログラミング] テーマ: 身近なソフトウェアの仕組みの理解
問015 [Lv.4] (想定正解率: 50% [5/10人]) [（3）コンピュータとプログラミング] テーマ: プログラムのデザインと改善
問016 [Lv.2] (想定正解率: 60% [6/10人]) [（4）情報通信ネットワークとデータの活用] テーマ: データに基づく問題の発見と予測
問017 [Lv.2] (想定正解率: 50% [5/10人]) [（4）情報通信ネットワークとデータの活用] テーマ: マルウェア
問018 [Lv.2] (想定正解率: 70% [7/10人]) [（4）情報通信ネットワークとデータの活用] テーマ: ネットワークの構成要素とプロトコルの選択
問019 [Lv.4] (想定正解率: 40% [4/10人]) [（4）情報通信ネットワークとデータの活用] テーマ: 相関係数と散布図
問020 [Lv.4] (想定正解率: 60% [6/10人]) [（4）情報通信ネットワークとデータの活用] テーマ: データベース管理システム

難易度別内訳: Level 1: 3問, Level 2: 7問, Level 3: 1問, Level 4: 8問, Level 5: 1問 (計 20問)

CSVファイル「questions.csv」に全 20 問を出力しました。

--- 実行統計 ---
所要時間　　: 9分30.5秒 (570.5秒)
API呼び出し回数 : 80 回
使用トークン数 : 合計 271,367 トークン (入力: 180,807 / 出力: 58,143 / 思考: 32,417)
```

作題例（Level 1　70%）

知的財産権のうち、自然法則を利用した技術的思想の創作のうち高度なものを保護対象とし、特許庁への出願・審査を経て登録されることで発生する権利として、最も適切なものを1つ選べ。

a. **特許権**
b. 実用新案権
c. 著作権
d. 意匠権

作題例（Level 4　60%）

あるプロジェクトにおいて、各作業の依存関係と所要日数が次のように定義されている。・作業A（要件定義）：所要日数4日（先行作業なし）・作業B（画面設計）：所要日数3日（作業Aが完了した後に開始可能）・作業C（データベース設計）：所要日数5日（作業Aが完了した後に開始可能）・作業D（画面開発）：所要日数4日（作業Bが完了した後に開始可能）・作業E（データベース構築）：所要日数6日（作業Cが完了した後に開始可能）・作業F（総合テスト）：所要日数3日（作業Dと作業Eの両方が完了した後に開始可能）。この作業日程をアローダイアグラムとして可視化して分析したとき、プロジェクト全体の最短所要日数と、作業Bが予定より2日遅れて完了した場合の全体日程への影響に関する記述として、最も適切なものはどれか。

a.　最短所要日数は25日であり、作業Bが2日遅延するとプロジェクト全体の完了日は27日になる。

b.　最短所要日数は18日であり、作業Bが2日遅延するとプロジェクト全体の完了日も2日遅れる。

c.　最短所要日数は14日であり、作業Bが2日遅延するとプロジェクト全体の完了日は16日になる。

d.　最短所要日数は18日であり、作業Bが2日遅延してもプロジェクト全体の完了日は変化しない。

作題例（Level 3　60%）

Webコンテンツの設計・制作において考慮すべき手法や概念に関する次の文章の空欄［ ア ］～［ ウ ］に入る語句の組合せとして、最も適切なものはどれか。

「Webサイトの制作において、利用者が求めている情報へ迷わず効率的に到達できるよう、コンテンツの論理的構造や分類、移動経路を整理して設計することを［ ア ］という。また、現在のページがWebサイト全体の階層構造の中でどの位置にあるかを視覚的に示し、上位の階層へ簡単に戻れるように画面上部などに配置されるリンクの並びを［ イ ］という。さらに、利用者の心身の特性や障害の有無、利用環境に関わらず、誰もが等しくコンテンツの情報にアクセスし利用できるように配慮されている度合いを［ ウ ］という。」

a.　ア：インフォグラフィックス　　イ：サイトマップ　　ウ：アクセシビリティ

b.　ア：インフォグラフィックス　　イ：パンくずリスト　ウ：ユーザビリティ

c.　ア：情報アーキテクチャ（IA）　イ：サイトマップ　　ウ：ユーザビリティ

d.　ア：情報アーキテクチャ（IA）　イ：パンくずリスト　ウ：アクセシビリティ

作題例（Level 4　60%）

次のプログラムを実行したとき、変数 count に格納されて出力される値として、最も適切なものはどれか。なお、配列 arr の添字は 0 から始まるものとする。

```
arr ← [5, 2, 9, 1, 6]
count ← 0
i を 0 から 3 まで 1 ずつ増やしながら繰り返す:
    j を (i + 1) から 4 まで 1 ずつ増やしながら繰り返す:
        もし arr[i] > arr[j] ならば:
            count ← count + 1
count を表示する
```

a.　6

b.　10

c.　4

d.　<u>5</u>

今後の予定

- 「情報I 作題支援のための主要学習項目・キーワード集」のブラッシュアップ
- 作問システムSakuAの有用性検証