

G2

多肢選択問題によるIRTに基づく 評価手法の構築

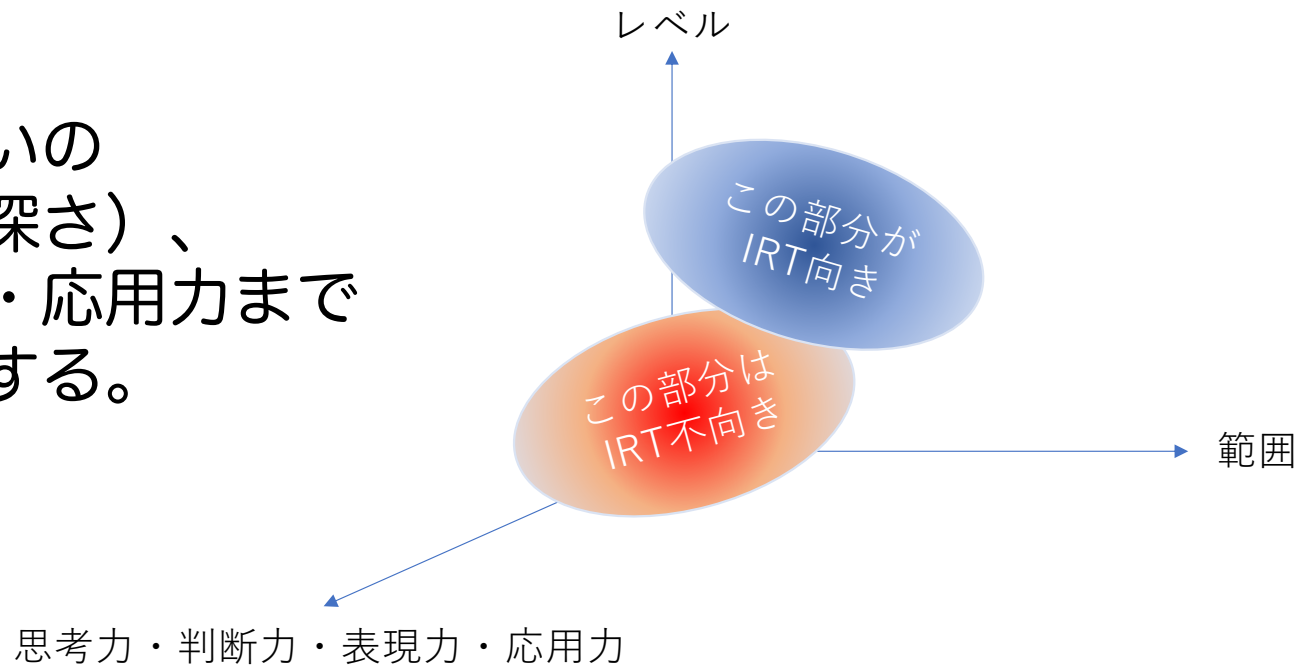
慶應義塾大学 環境情報学部
植原啓介 (kei@sfc.keio.ac.jp)

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方法についての検討状況

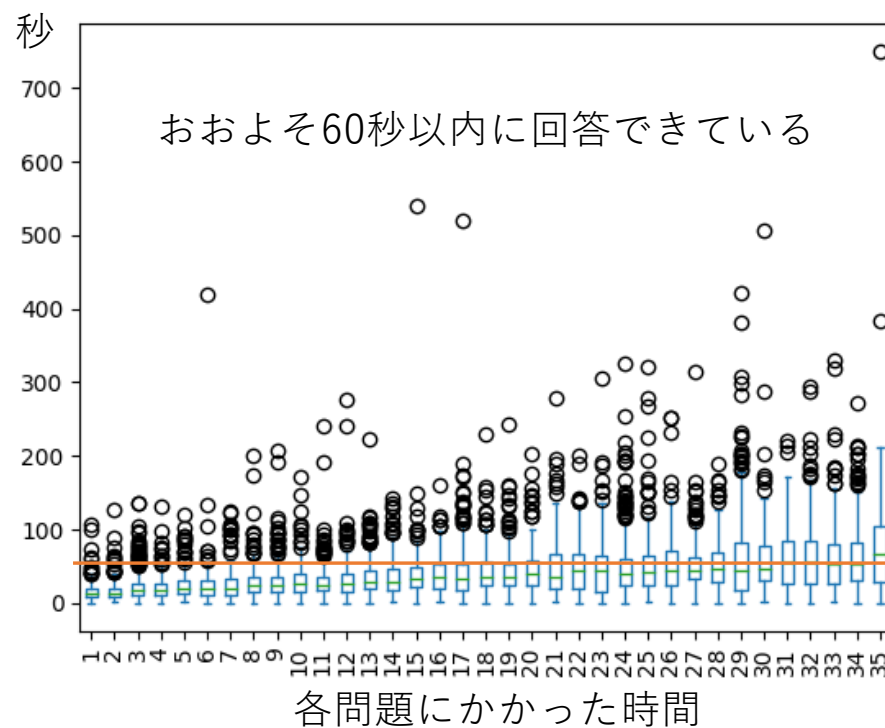
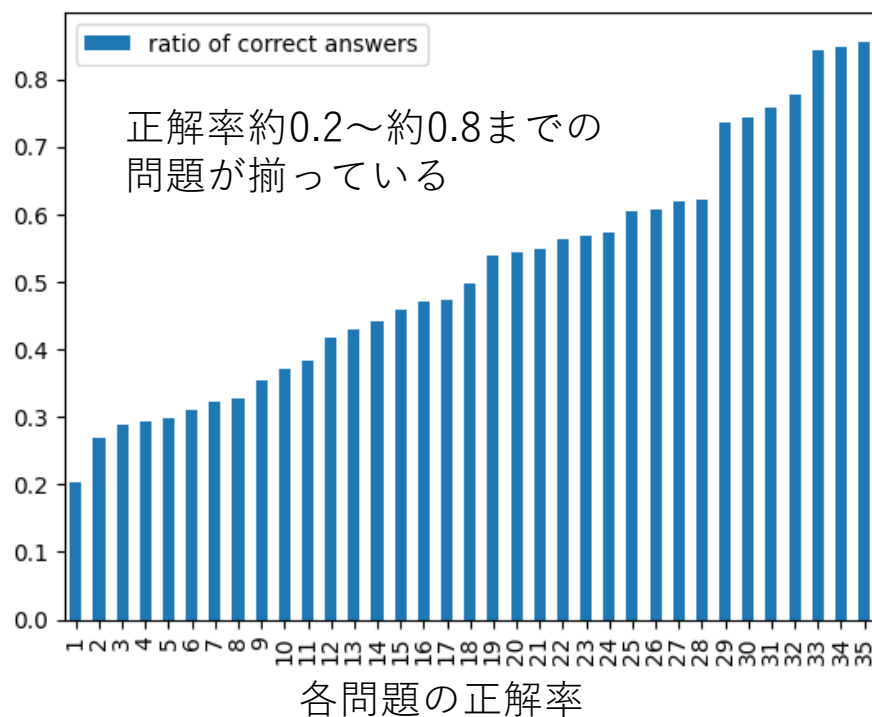
情報に関連する語句のリストがあるとした場合、下記のような作問の方法がある。

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）

現在、上記の方法に基づいて作問を進めている状況。

IRTを想定した多肢選択問題の結果

- 簡単な問題から難しい問題まで適切に出題できていたと考えられる
- 1問1分想定であったが、十分な時間が与えられたと考えられる



IRT問題の例

問題を解くための作業手順を過不足なく説明したもののことをなんと呼ぶか。最も適切なものを選びなさい。

1. アルゴリズム
2. アプリケーション
3. ソフトウェア
4. パッケージ

- 最も正答率が高かった問題（正答率84.4%）
- 前ページの1番目の出題方法。
- このような教科書に出てくる基本的な用語を問う問題は、正答率が高くなる傾向にある。

IRT問題の例

次のプログラムを実行した際に得られる出力はどれか。ただし、XおよびYは配列、iは変数である。また、配列の添字は0から始まるものとする。

X = [1, 5, 8, 9, 3]

Y = [3, 5, 2, 2, 9]

iを0から4まで1ずつ増やしながら繰り返す:

├ もし X[i] > Y[i] :

├ └ 表示する(X[i])

選択肢

1. 8 9

2. 5 8 9

3. 1 3

4. 1 5 3

- 正答率が平均に近かった問題（正答率57.4%）
- プログラミングの要素としてはループと条件分岐、そして配列が含まれている。
- 単純なループによる処理だけを扱う（配列や条件分岐がない）ものは、本問題より正答率が高くなる傾向にある。

IRT問題の例

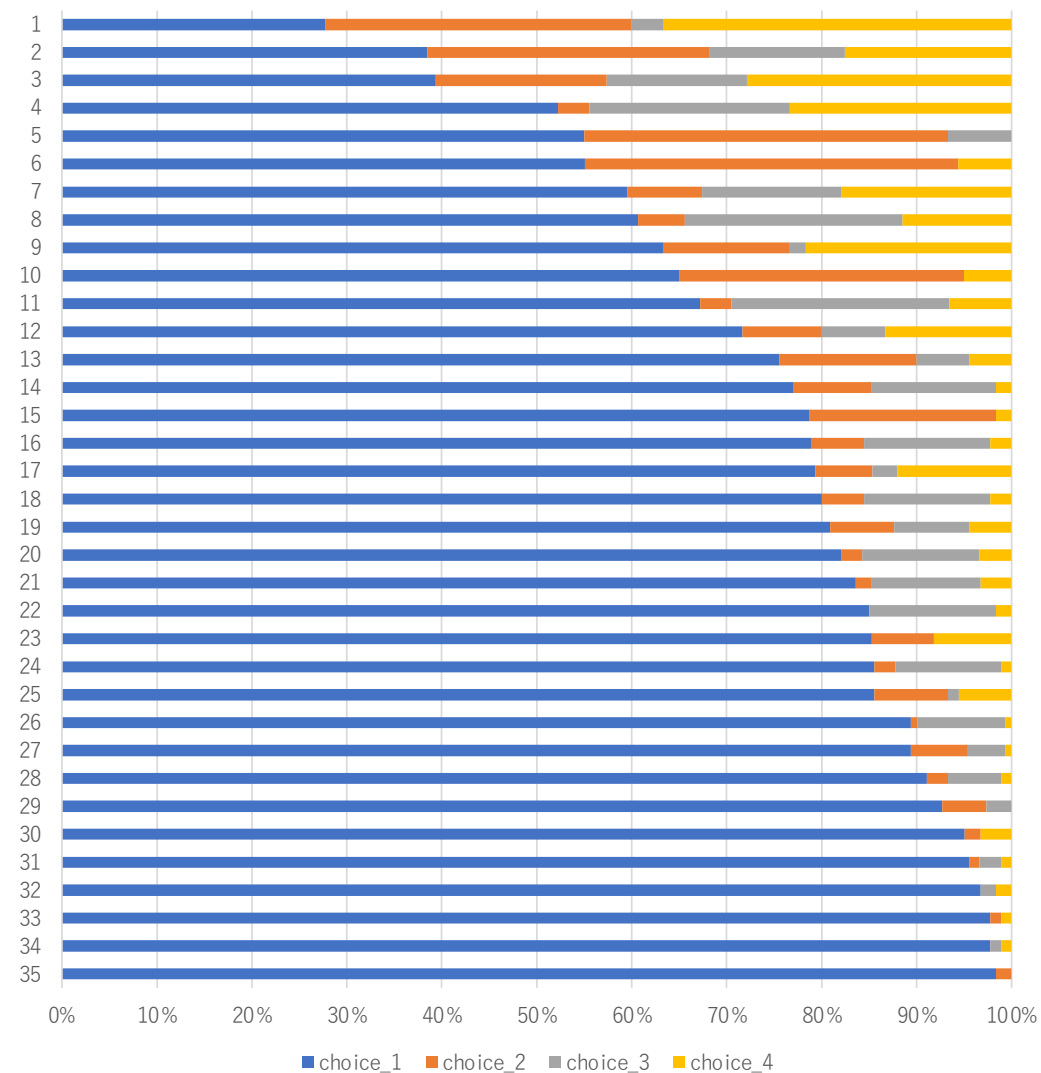
変数 a, b, c の値のなかで、最も小さな値を d にセットする処理（記述）はどれか。

- | | | | |
|--------------|-------------------|-----------------|-------------------|
| 1. $d = a$ | 2. $d = a$ | 3. もし $a < b$: | 4. もし $a < b$: |
| もし $d > b$: | もし $d > b$: | └ $d = a$ | └ $d = a$ |
| └ $d = b$ | └ $d = b$ | もし $b < c$: | そうでなくもし $b < c$: |
| もし $d > c$: | そうでなくもし $d > c$: | └ $d = b$ | └ $d = b$ |
| └ $d = c$ | └ $d = c$ | $d = c$ | そうでなくもし $c < a$: |
| | | | └ $d = c$ |

- 最も正答率が低かった問題（正答率20.3%）
- それぞれの選択率は1. 20.7%、2. 26.8%、3. 13.8%、4. 38.7%であった。選択肢4のプログラムが最も強く誤答に誘導したが、その理由は明らかでない。
- 変数の上書きが行われる問題の正答率は低くなる傾向にある。（手元に紙が無い?）
- 選択肢1において比較演算子が「大なり」を使っているが、「最も小さな値」という問題文との不整合が影響を及ぼした可能性もある。

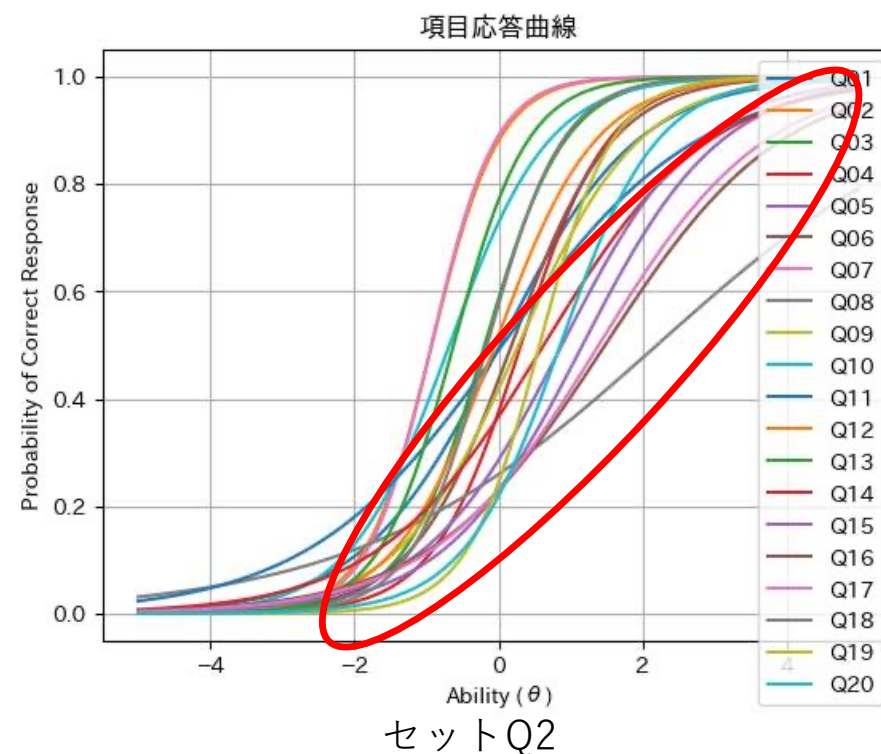
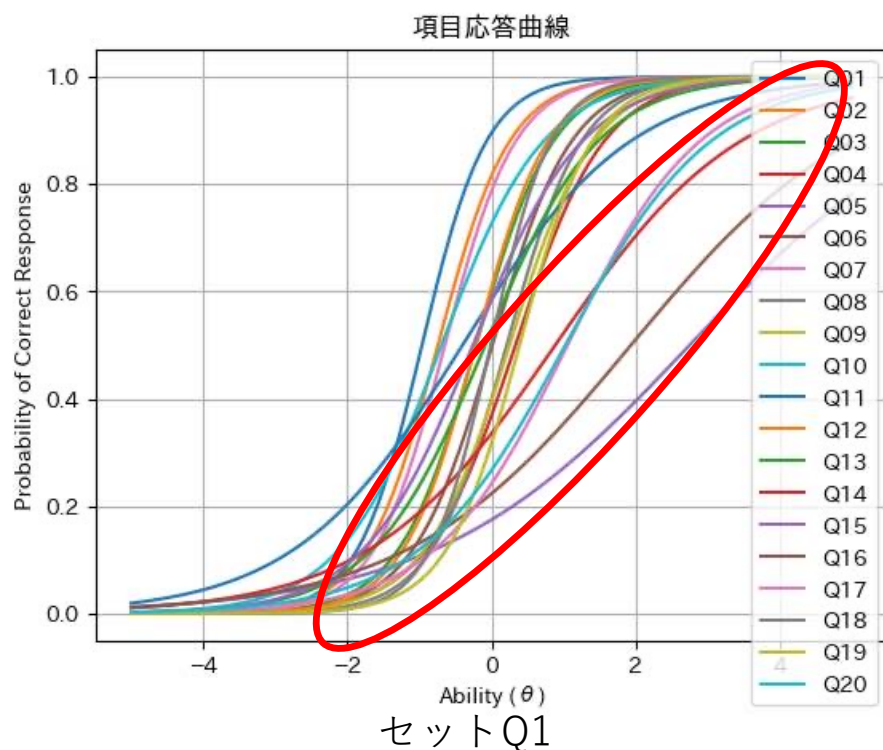
IRTを想定した 多肢選択問題の結果

- 難しい問題は上位25%の者でも正答率が28%程度



IRTを想定した多肢選択問題の結果

- 一部難しすぎる問題○も見られるが、概ね弁別性のある適切な問題が出題できている



今後について

- 今回得られた知見をもとに、例えば正答率が低かった問題はなぜ正答率が低かったのか、といった研究を進める。
- 他の分野の作問を継続する。