

レポート提出票

科目名: 情報工学実験3

実験課題名: 課題3 教育システム

実施日: 2024年 6月 13日

学籍番号: 4622045

氏名: 小澤 翼

共同実験者:

1 要旨

今回は、アプリケーションを作成することを主として古典的な方法を実装し、主観評価やインターフェイスの重要性について理解することが出来た。

2 目的

アプリケーションを作成することを主として古典的な方法を実装し、主観評価やインターフェイスの重要性について理解すること。

3 理論 (テスト理論)

古典テスト理論は一般的に行われているテストの採点方法である。受験者の能力値 θ は以下の式で定式化される。

$$x = \theta + se$$

観測反応 x には真の能力値 θ と偶然誤差 se が含まれているという意味である。この観測を多数行い、平均を取れば x は θ に一致する。

$$\theta = E(x)$$

ただし、 se は単なる偶然誤差であり、多要因とは無関係である。実装は単に正答率を能力値とすればよい。

```
1 const theta = x.reduce((a, b) => a + b) / x.length;
```

この計算は x の取りうる範囲に依存し、これまで通り $x \in 0, 1$ の時は加点方式となる。これを、正答時に $x = 1$ 、無回答時に $x = 0$ 、誤答時に $x = -1$ とすると誤びょう率となり、回答が多くても間違いが多いと能力値が低下する。SPI で「分からない問題は回答しないで下さい」という指示があった場合は誤びょう率を計測している可能性が高い。

古典理論は問題の性質を考慮していないという点で限界があり、さまざまなテスト理論が提案されている。一般的にテストは、その目的別に表 1 のように分類できる。能力測定型テスト

表 1: テストの分類

種類	規模	例	理論
測定	広	センター試験	IRT
選抜	広	入試	IRT
診断	中	期末テスト	ネットワーク型 IRT
形式	狭	授業中の課題	SP 表

は受験者の能力値を計算することが目的であり、大学入試センター試験や TOEFL、TOEIC などが代表例として挙げられる。選抜型テストは、定数内の受験者の選抜やある能力基準以上の受験者の選抜が目的であり、入社・入学試験や資格試験などが代表例として挙げられる。診断型テストは学習者の学習行き詰まりの原因を調査するためのテストで、学期末などに行われる定期試験などで実施されることが多く、直接的には能力測定を目的としていない。ネットワーク型 IRT とは局所独立性を仮定しない IRT である。形式的テストは、教育現場で学習者・教師

の日々の改善のために実施されるテストで、授業で配布される演習問題プリントなどを指す。テストの結果より、教師は個々の学習者の習熟度を確認したり、授業ペースや授業難易度の適切性を確認したりする。SP 表とは Student-Problem 表の略であり、授業改善を目的とした順序統計的な分析法の一種である。

4 課題

4.1 演習 2 - 1

`getItem(0).then(item⇒console.log(item));` で問題が表示されるか確認せよ。

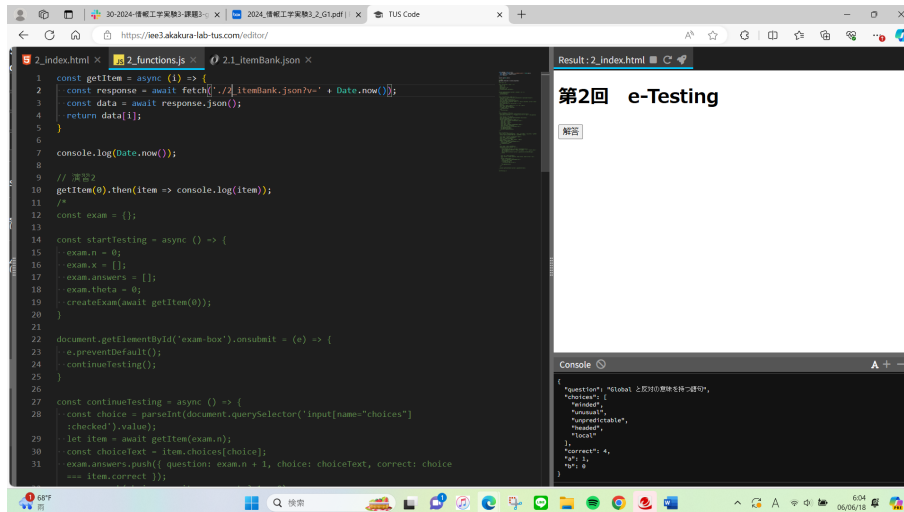


図 1: 演習 1 - 1 の結果

4.2 演習 2 - 2

ブラウザに問題文とラジオボタンと選択肢が表示されること、および能力値（正答率）が正しく計算されることを確認せよ。



図 2: 演習 2 - 2 の結果

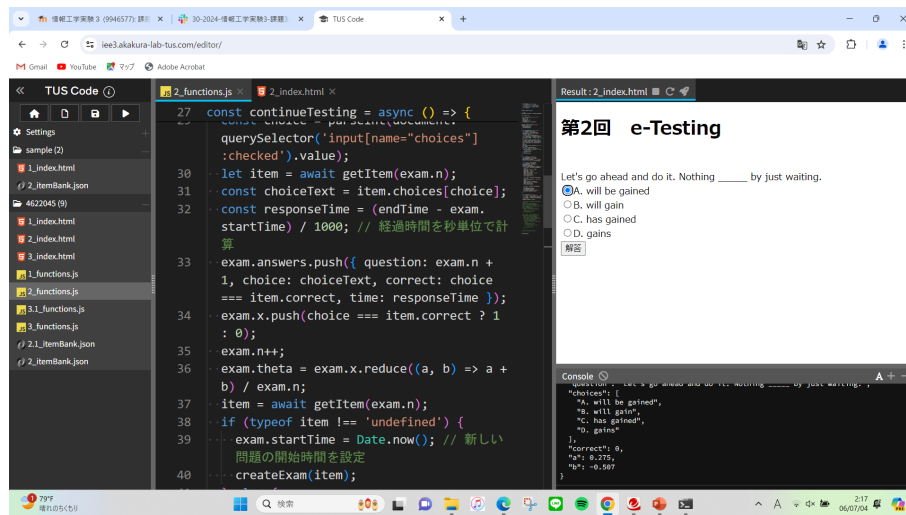


図 3: 問題文とラジオボタンと選択肢

4.3 演習 2 - 3

与えられた英語の問題セット 5 問を解き、自身の解答を記録しなさい。その後、古典テスト理論による能力値を算出するテストアプリケーションを完成させ、自身の能力値を求めなさい。

私の結果は、第 1 問を A と選択し正解、第 2 問を A と選択し正解、第 3 問を C と選択し不正解、第 4 問を A と選択し正解、第 5 問を D と選択し不正解であった。以下の図の 4 は、この結果から出された古典テスト理論に基づく能力値である。

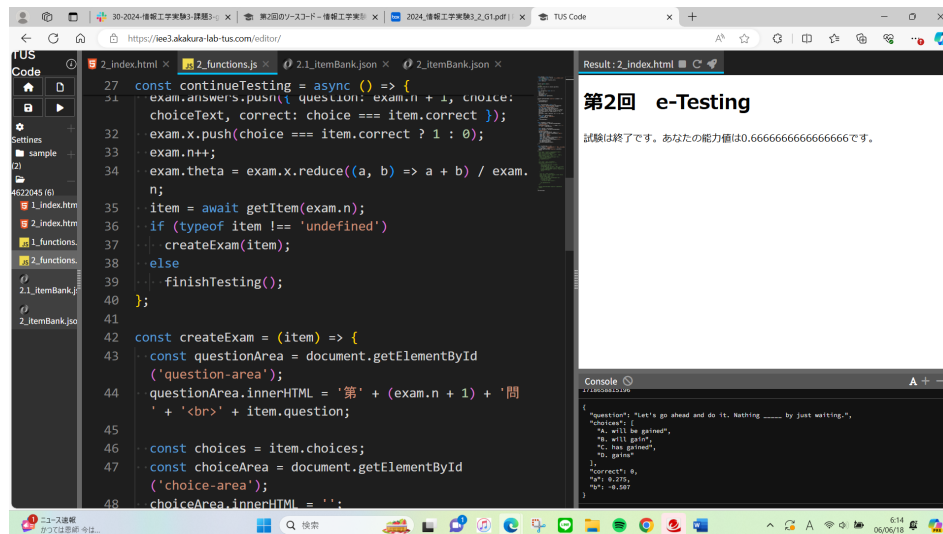


図 4: 演習 2 - 3 の結果

5 課題2 - 1

演習2 - 3に以下の条件を付けくわえよ。

- 問題番号の表示
- 結果のページに解いた問題番号と選択した解答、およびその正誤を以下の図5のように表示

問	あなたの解答	正誤
1	A. Rikadai	正
2	B. Kagurazaka	誤
3	C. Noda	正
4	D. Katsushika	誤

図 5: 課題2 - 1 の例

問題番号は以下の図6のように表示されている。

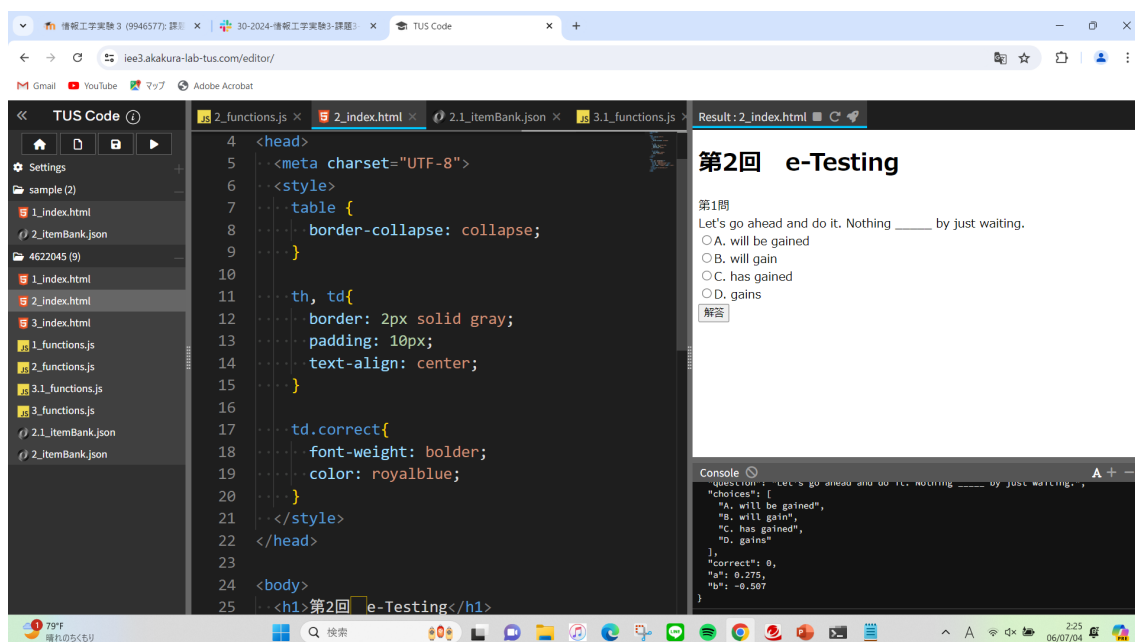



図 6: 課題2 - 1 の結果

結果は以下の図7の通りである。



問	あなたの解答	正誤
1	A. will be gained	×
2	A. patience	×
3	A. suggest	×
4	A. had been	×
5	A. I think it's a very good idea.	×

図 7: 課題 2 - 1 の結果

5.1 課題 2 - 2

全ての問題を解答後にテスト全体に対する自信度を 5 件法で取得し、取得した自信度を結果ページで表示せよ

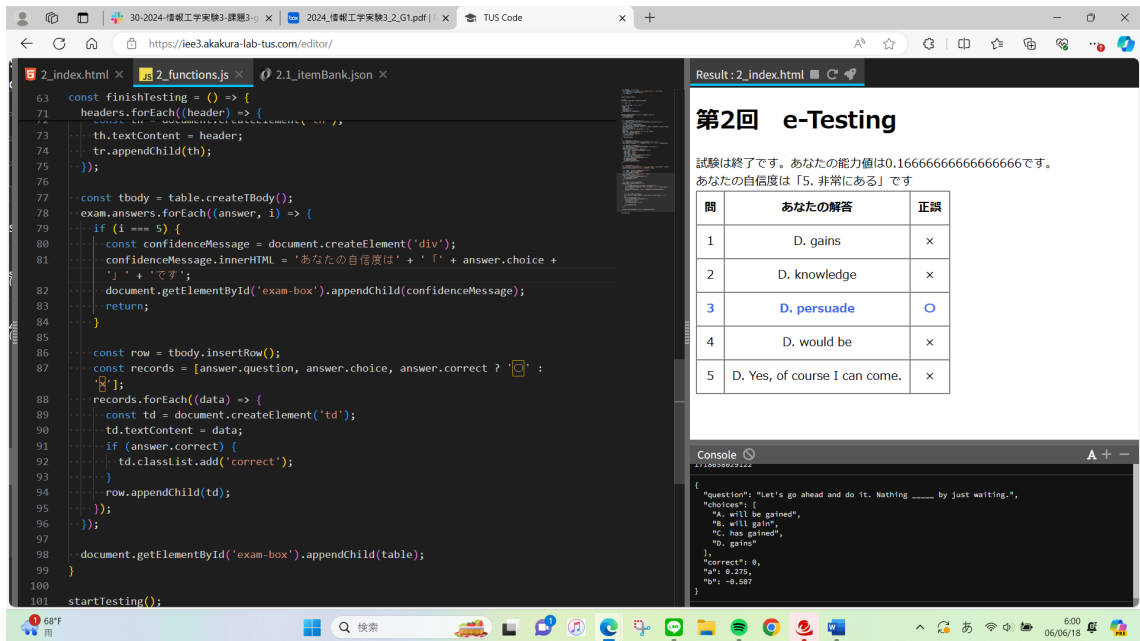
自信度を答える画面は以下の図 8 のようになっている。



問	あなたの解答	正誤
1	A. will be gained	×
2	A. patience	×
3	A. suggest	×
4	A. had been	×
5	A. I think it's a very good idea.	×

図 8: 課題 2 - 2 の結果

結果は以下の図9の通りである。



第2回 e-Testing

試験は終了です。あなたの能力値は0.16666666666666666です。
あなたの自信度は「5. 非常にある」です

問	あなたの解答	正誤
1	D. gains	×
2	D. knowledge	×
3	D. persuade	○
4	D. would be	×
5	D. Yes, of course I can come.	×

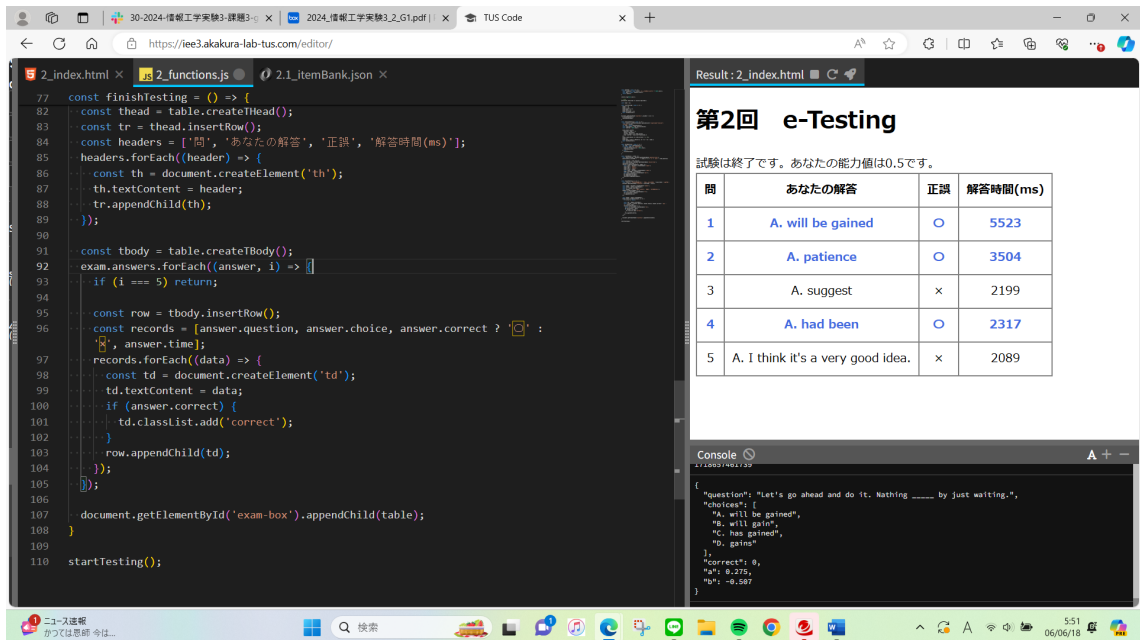
Console

```
{
  "question": "Let's go ahead and do it. Nothing ____ by just waiting.",
  "choices": [
    "A. will be gained",
    "B. will gain",
    "C. has gained",
    "D. gains"
  ],
  "correct": 0,
  "m": 0.275,
  "n": -0.587
}
```

図 9: 課題 2 - 2 の結果

5.2 課題 2 - 3

各問題の解答時間の記録を課題 2 - 1 の表に出力せよ。



第2回 e-Testing

試験は終了です。あなたの能力値は0.5です。

問	あなたの解答	正誤	解答時間(ms)
1	A. will be gained	○	5523
2	A. patience	○	3504
3	A. suggest	×	2199
4	A. had been	○	2317
5	A. I think it's a very good idea.	×	2089

Console

```
{
  "question": "Let's go ahead and do it. Nothing ____ by just waiting.",
  "choices": [
    "A. will be gained",
    "B. will gain",
    "C. has gained",
    "D. gains"
  ],
  "correct": 0,
  "m": 0.275,
  "n": -0.587
}
```

図 10: 課題 2 - 2 の結果

この解答時間が記録できることから問題の解答時間の制限をお K することが出来る。また、大勢の受験者から解答時間を算出することで、どの問題が難しく多くの時間を強いているか分かるので、出題者が問題の難易度の把握や調整が出来ると期待できる。

6 まとめ

私は、今まで多くのオンライン上での今回実装したような問題を解いて来ました。問題を解いている間は、実装に必要なソースコードのことは考えていなかったが、今回実験を行ってかなり手間がかかるものだと認識した。今後はもっと場数を踏み、一人で一からこの程度の問題が作れるようになりたいと感じた。