

(演習2_1)

- ・ 空欄に入る用語を選んでください
- ・ スロフィッキーは、集合知がうまくいくには以下の4つが必要であると述べている

1. 分散性

- 問題を抽象化せず、各参加者が直接得られる情報に基づいて判断する必要がある。参加者ごとに得られる情報の種類は異なると予想されるが、2.を維持するためにも、各参加者に共通する属性のみで判断すべきでない。

2. 多様性

- 各参加者がそれぞれに独自の視点を持っていれば、総体として多くの候補解を列挙することができる。探索空間が狭い場合には、その探索空間内に適切な解が存在しない可能性がある

3. 独立性

- 各参加者の持つ意見や提案が他の参加者の影響を受けないよう、各参加者の3.が確保されている必要がある。とくに小集団で議論を行う場合には、2.が低いために偏った結論に集約される危険性がある。

4. 集約性

- 1.~3.の特性を生かして得られた知識を参加者全体で共有し、比較検討して最終的な結論を導く仕組みが必要である。

集合性、自律性、分散性、頑健性、多様性、柔軟性、集約性、協調性、独立性

(演習2-2)

「ビンの中のジェリービーンズ」実験が成功したのか（or 失敗したのか）理由を説明してください



- 上手くいかなかった
- 数が少なかった（群衆とまではいかなかった）
- 相談してしまったから（独立性がなかった）
- みんな同じ学生だったから（多様性がなかった）
- 外れ値があったから（別の集約方法が必要かもしれない）
- . . .

(演習2_3)

・以下の説明が正しければ () に○、間違っていれば×をつけてください

SEO

- (×) ・ ~~検索エンジン~~とは、検索結果でより上位に現れるようにウェブページを書き換えることである

クエリー処理

- (×) ・ ~~クローラー~~とは、ユーザーからの検索キーワードを処理し、スコアリングををもとに検索順位を決め、結果をリスト表示する処理である

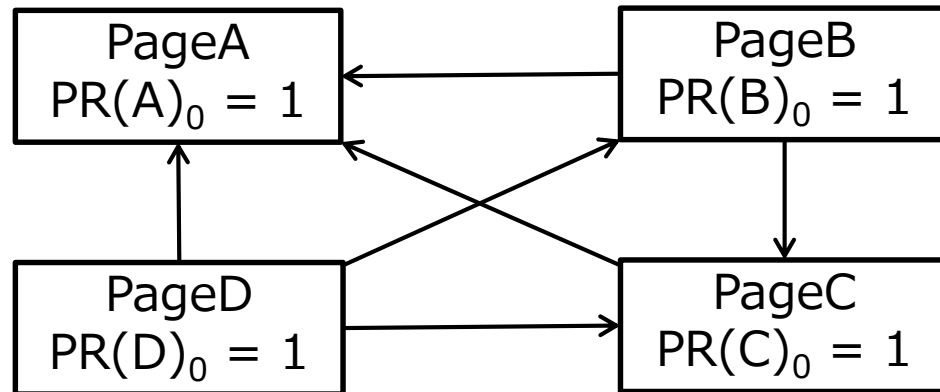
- (○) ・ インデックス付け (Indexer) とは、クローラーが収集したWebページのデータを、データベースに (処理しやすいように整理された状態で) 格納することをいう

- (×) ・ PageRankとは、Google社が開発したウェブページの重要度の判定を、ページの内容を分析して指標化する技術である

ページの内容を分析しないで

(演習2_4)

- PageRankを計算して以下の表の () を埋めてください



$$PR(pi)_t = (1 - d) + d \sum_{pj \in M(pi)} \frac{PR(pj)_{t-1}}{L(pj)}$$

反復回数	PR(A)	PR(B)	PR(C)	PR(D)
t=0 (初期値)	1	1	1	1
t=1	1.708	0.433	0.858	0.15
t=2	1.106	0.193	0.377	0.15
t=3	0.595	0.193	0.275	0.15
t=4	0.508	0.193	0.275	0.15
t=5	0.508	0.193	0.275	0.15