

世界を変えた Googleの

PageRank
アルゴリズム

Google

perfume



[すべて](#)[動画](#)[ニュース](#)[画像](#)[ショッピング](#)[もっと見る](#)[設定](#)[ツール](#)

約 430,000,000 件 (0.49 秒)

Perfume Official Site

www.perfume-web.jp/ ▼

所属事務所アミューズによる**Perfume**オフィシャルサイト。
このページに複数回アクセスしています。前回のアクセス: 18/10/04

FUTURE POP - Perfume Official Site

<https://www.perfume-web.jp/index-jpn.php>

所属事務所アミューズによる**Perfume**オフィシャルサイト。

Perfume - YouTube

<https://www.youtube.com/channel/UCxOjoraUPd0Dq9PAyIhC6tQ> ▼

Perfume YouTube Official Channel. ... Tickets are now on-sale for the Asia & North American **Perfume** WORLD TOUR 4th "FUTURE POP" shows. Get tix ↪ <http://smarturl.it/PrfmWorl...> ▽詳細はこちら
More info: **【Perfume Official Site】** ...

Perfume - ウィキペディア

<https://ja.wikipedia.org/wiki/Perfume> ▼

Perfume（パフューム）は、中田ヤスタカがプロデュースする広島県出身の3人組テクノポップユニット。ダンスと舞台演出に特徴がある。アミューズ所属。レコードレーベルは**Perfume Records** / ユニバーサルJ。公式ファンクラブは「P.T.A.」。

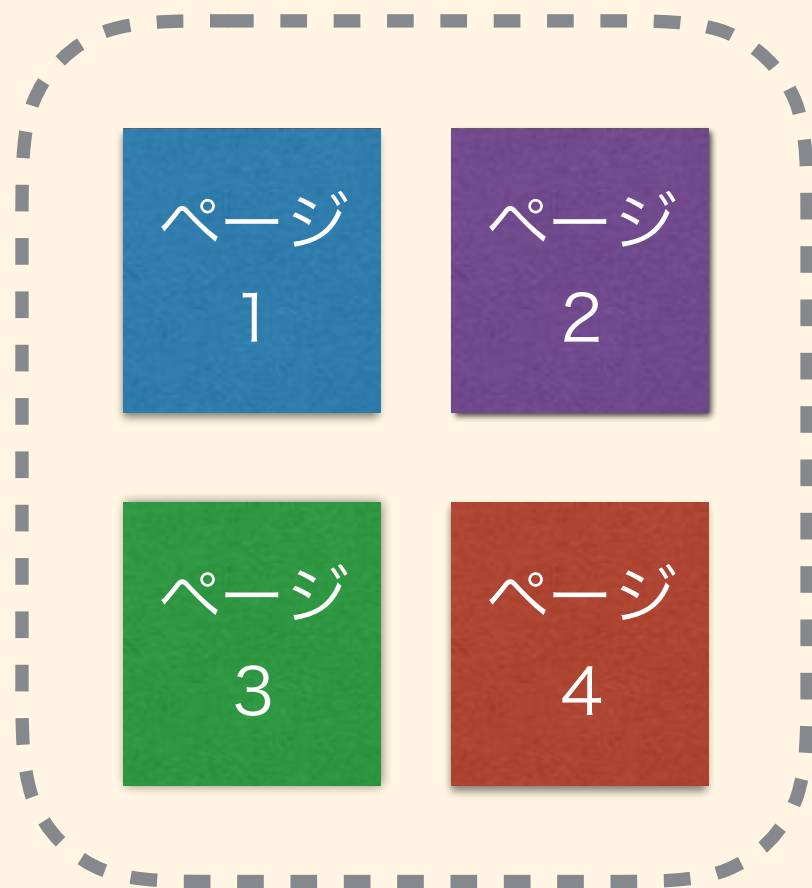
[Perfume \(曖昧さ回避\)](#) · [Perfume Livefolio](#) · [Perfume LOCKS!](#) · [Perfume Clips](#)

4億のページから
重要なページを
選び出す

ランキングアルゴリズム

- Webページを重要な順番に並び替える

検索結果



1 位

2 位

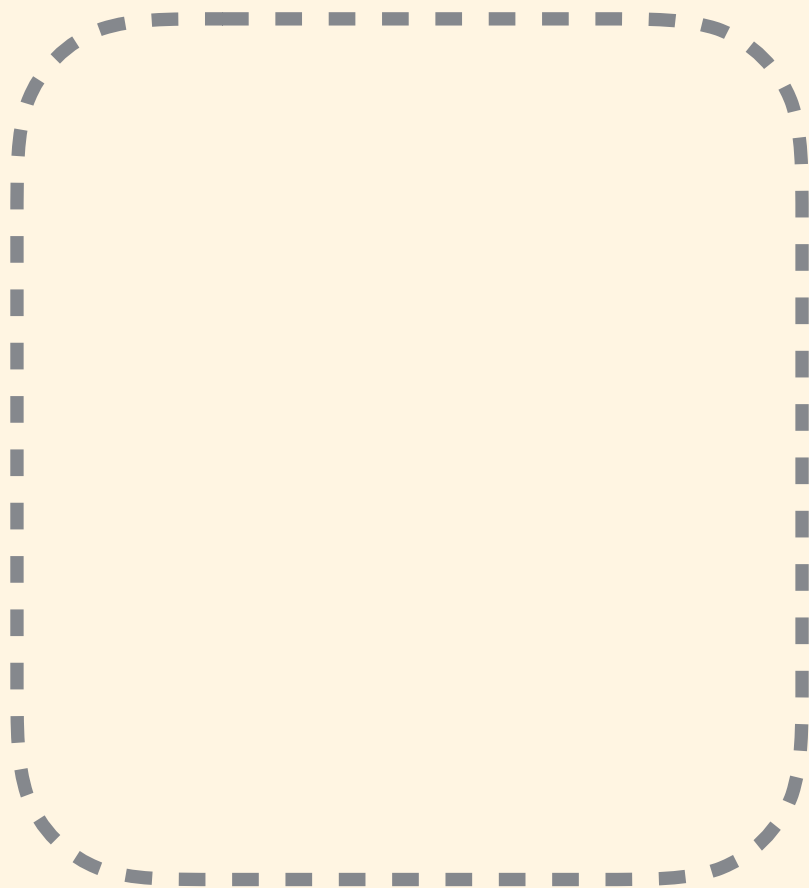
3 位

4 位

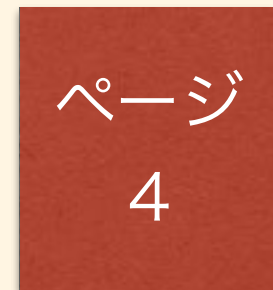
ランキングアルゴリズム

- Webページを重要な順番に並び替える

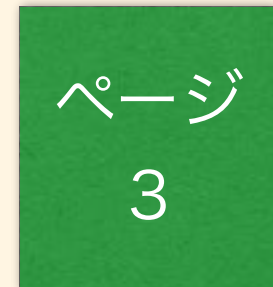
検索結果



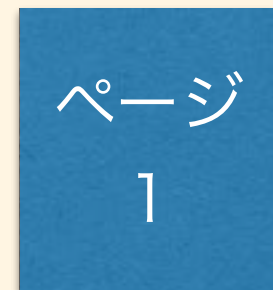
1 位



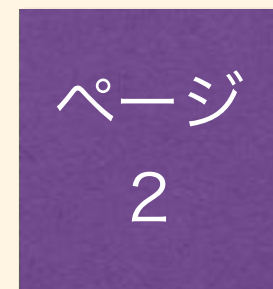
2 位



3 位



4 位



ランキングアルゴリズム (例)

- 検索ワードの出現数でランキング

ページ 1

... .. perfume ...
... ..
... perfume
perfume

ページ 2

... ..
... ..
... ..
... perfume

ページ 3

... perfume
... .. perfume
... .. perfume ...
perfume

出現数

3 個

1 個

4 個

順位

2 位

3 位

1 位

Web検索

最大の難問



Google's Larry Page and Sergey Brin ask an unusual interview question | CNBC make it
<https://www.cnbc.com/2017/05/31/google-larry-page-and-sergey-brin-ask-an-unusual-interview-question.html>

PageRank

アルゴリズム

NOTICE

- 中学生でもわかる直感的なイメージ
- プログラミングというよりアルゴリズムのお話
- Processingによるシミュレーション

突然ですが

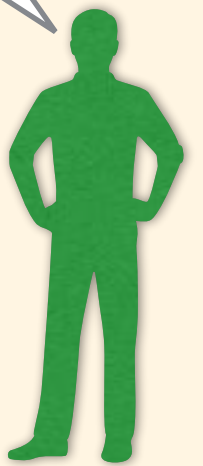
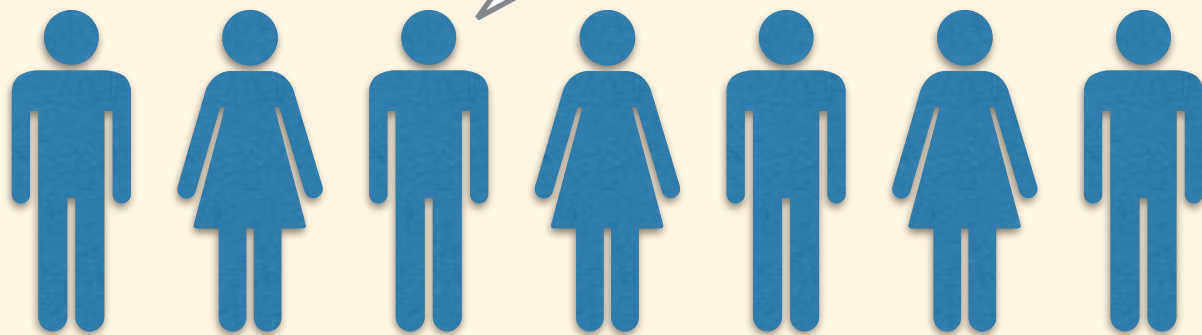
学級委員長の
良い決め方は？

学級委員長の決め方

- ・ 立候補の問題点

はい！

委員長に
なりたい人！

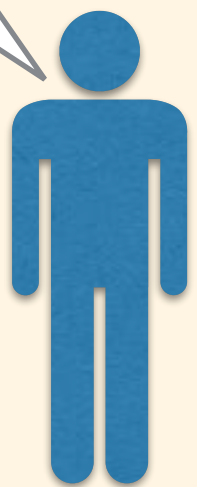
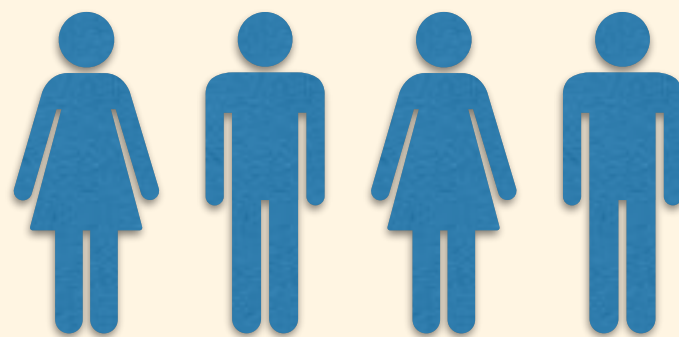
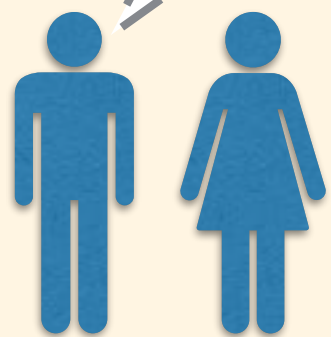


学級委員長の決め方

- 立候補の問題点

あいつ全然
仕事しねー...

俺が
委員長！



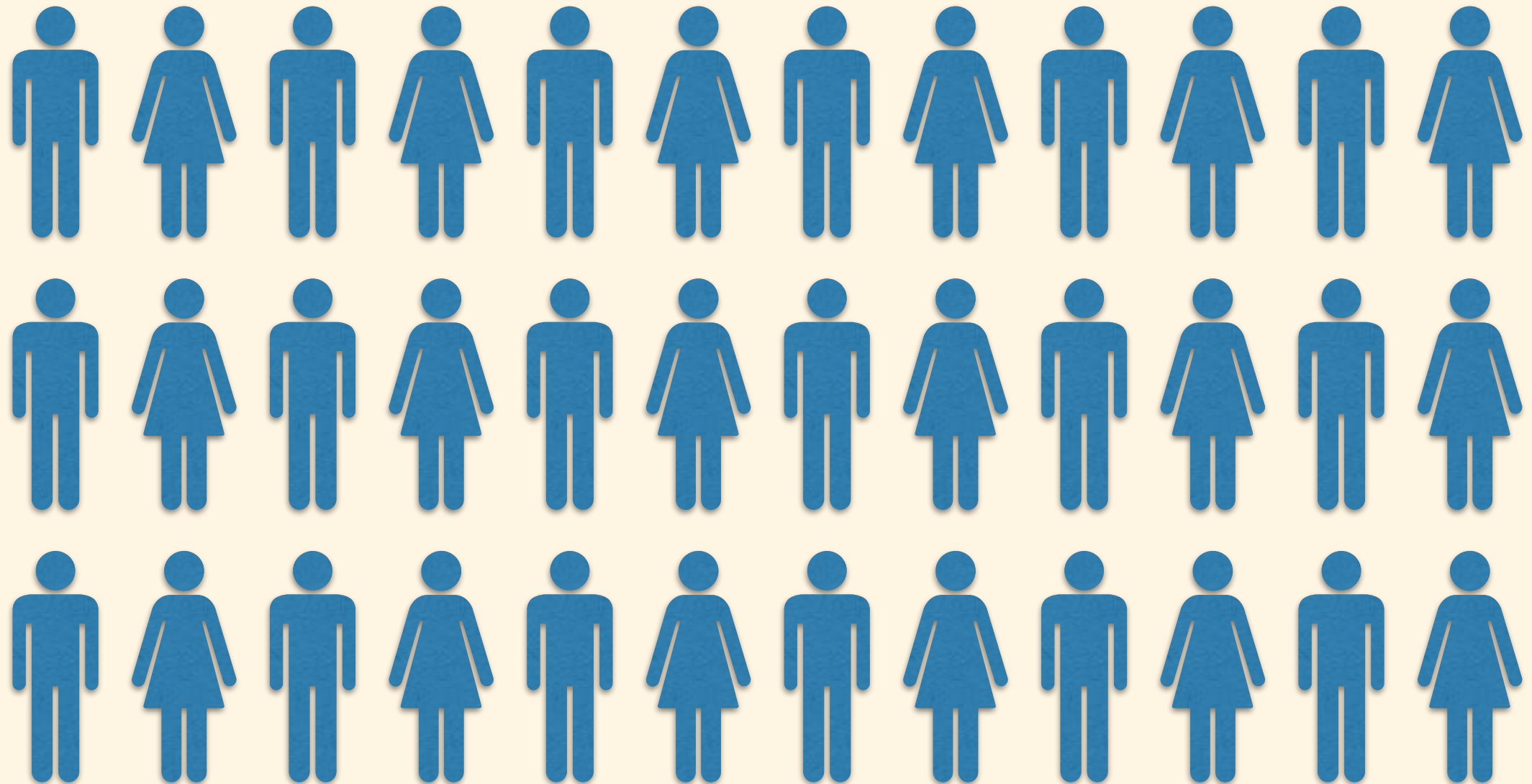
学級委員長の決め方

- 立候補の問題点
 - その人が信頼できるかどうかに関わらない

信頼できる人をどう選ぶか？

最も信頼できる人の選び方

- クラス全体から**最も**信頼できる人を選ぶ



最も信頼できる人の選び方（例）

- 多数決（1人1票 or 1人複数票）

得票数	
Aさん	4票
Bさん	12票
Cさん	7票
...	...

1票の
重みは同じ

本気出して
考えてみた

新ルール

大原則

1 票の重みが異なる

新ルール

原則 1

- ・ 多くの人から支持される人は信頼できる

原則 2

- ・ 信頼できる人からの支持に、より価値がある

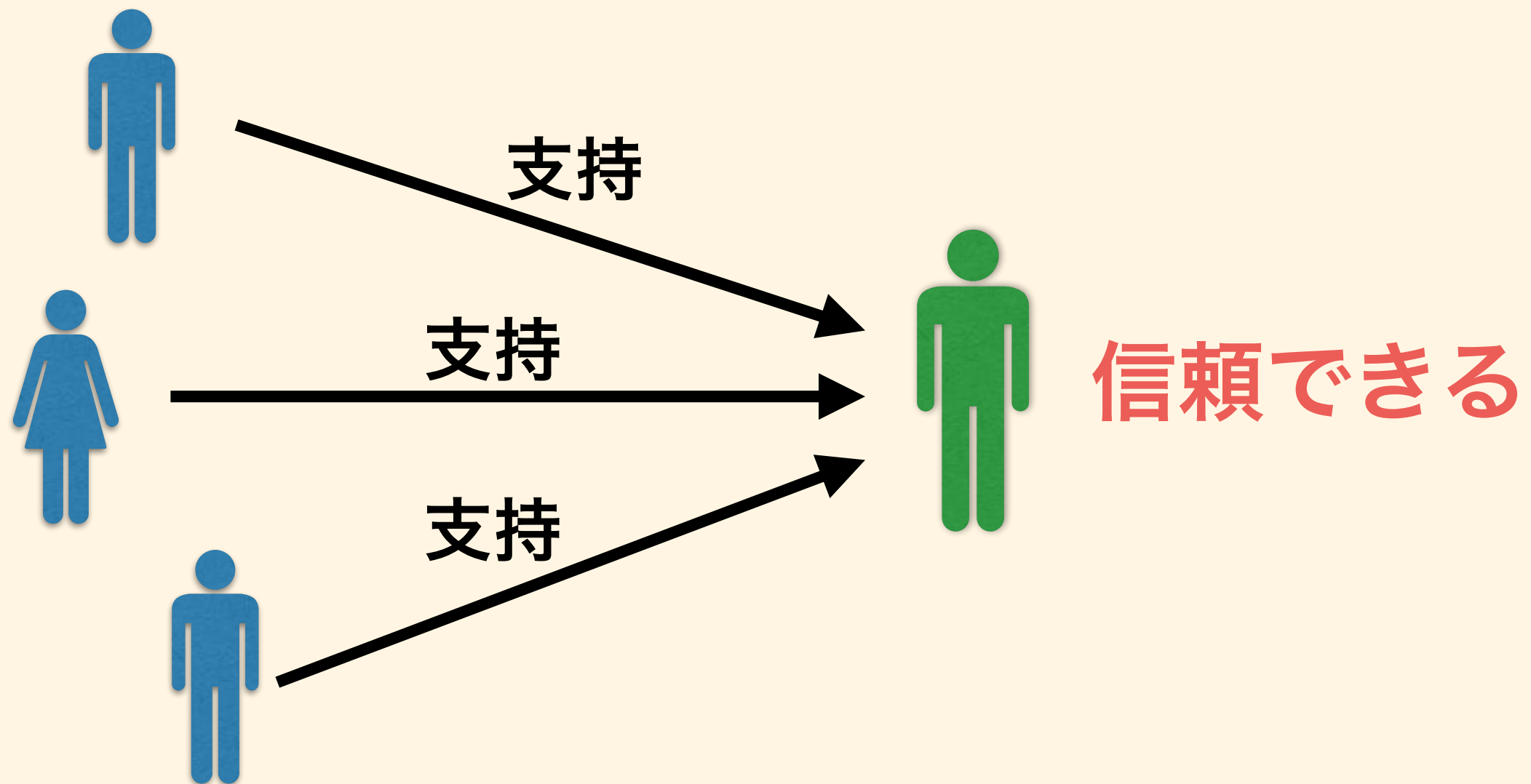
原則 3

- ・ 多数への支持より、少数への支持に価値がある

新ルール

原則 1

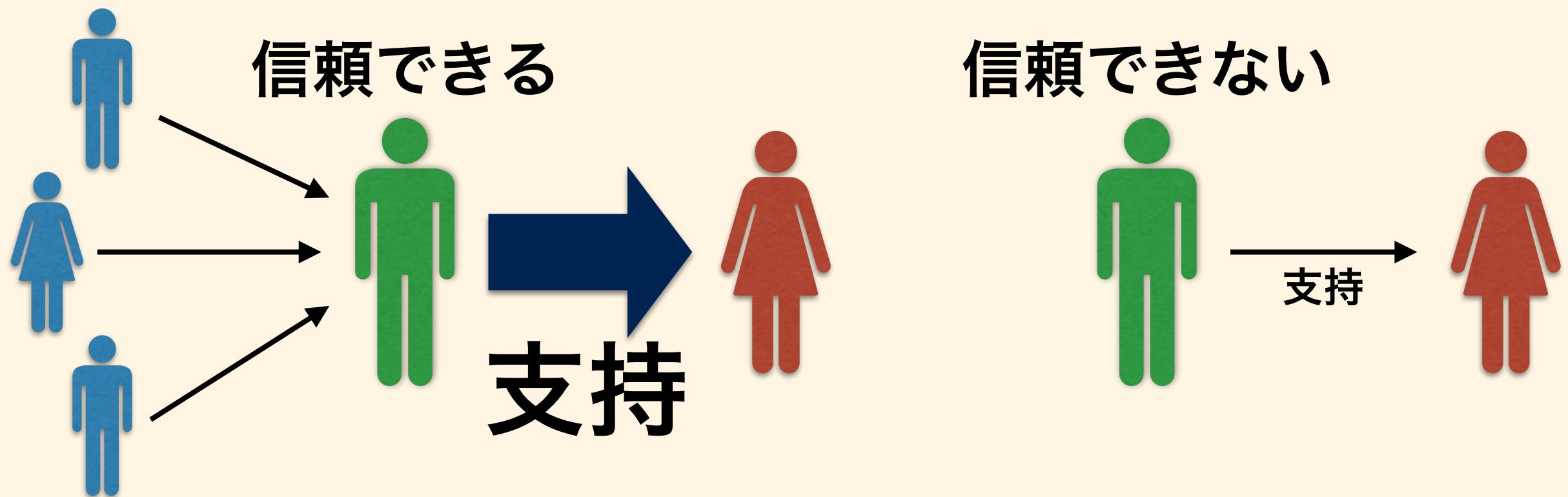
- ・ 多くの人から支持される人は信頼できる



新ルール

原則 2

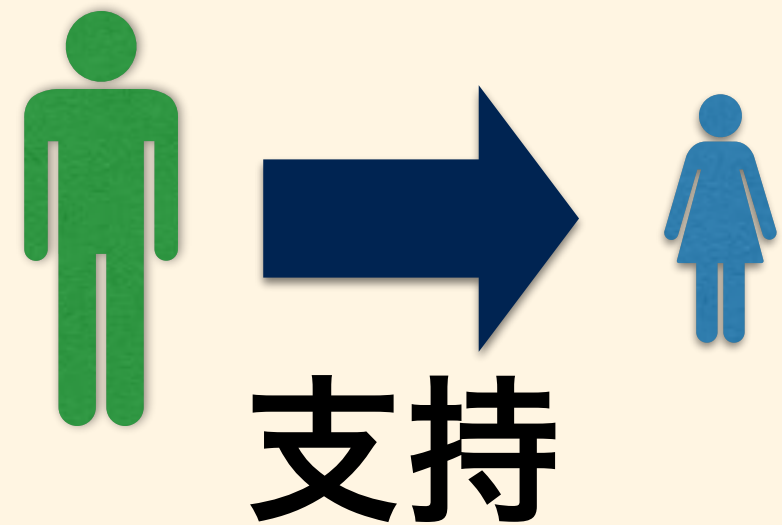
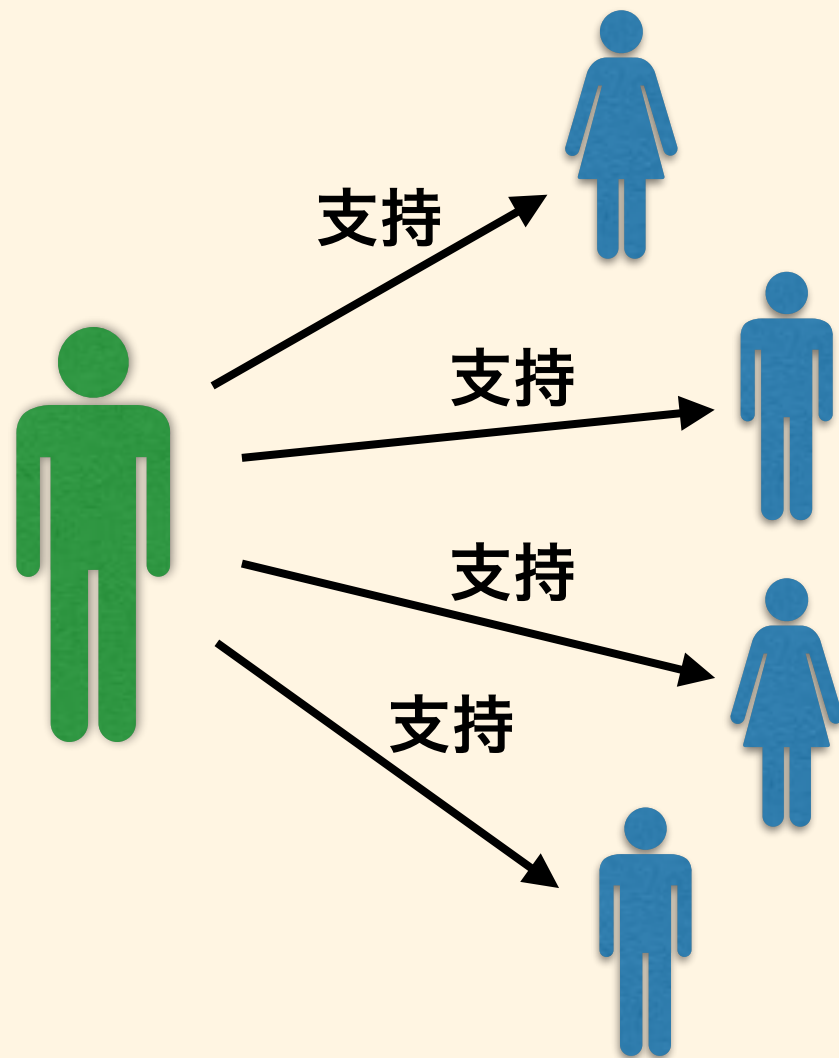
- ・ 信頼できる人からの支持に、より価値がある



新ルール

原則 3

- ・ 多数への支持より、少数への支持に価値がある



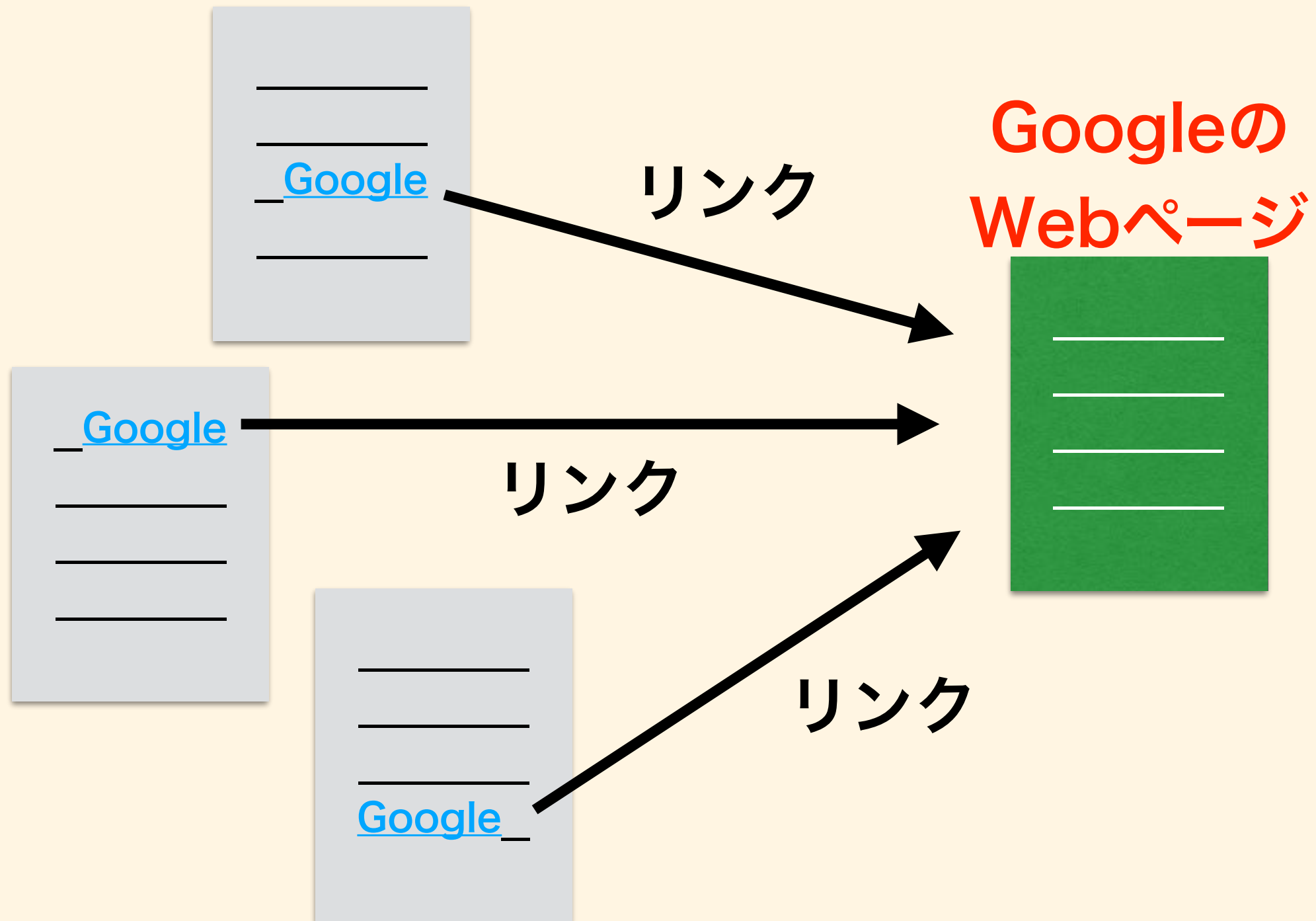
閑話休題

PageRankアルゴリズム

最大の特徴

PageRank 最大の特徴

- 「リンク」をそのページへの「支持」と考える



PageRank 3つの原則

原則 1

- ・多くのページからリンクされるページは重要である

原則 2

- ・重要なページからのリンクに、より価値がある

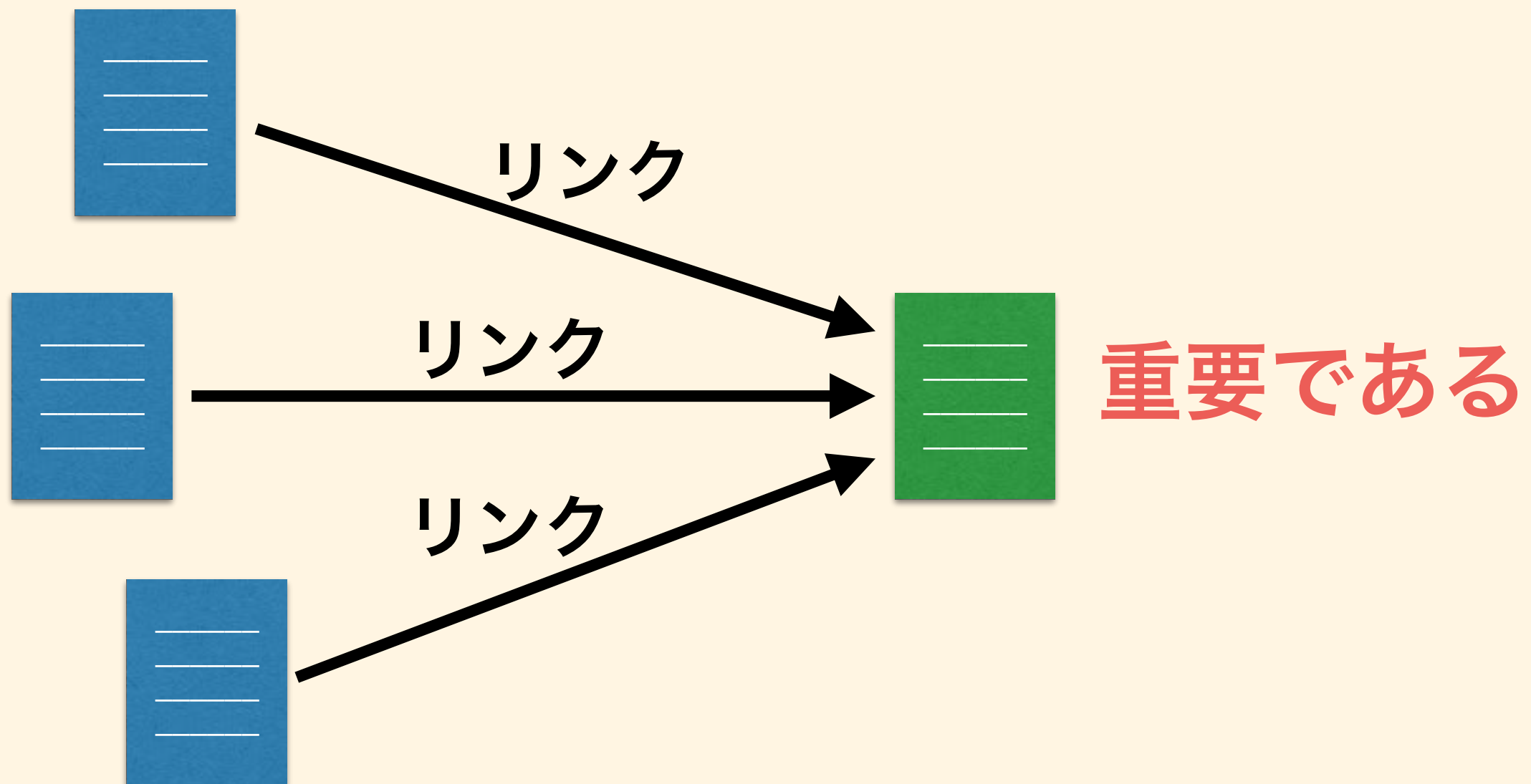
原則 3

- ・多数へのリンクより、少数へのリンクに価値がある

PageRank 3つの原則

原則 1

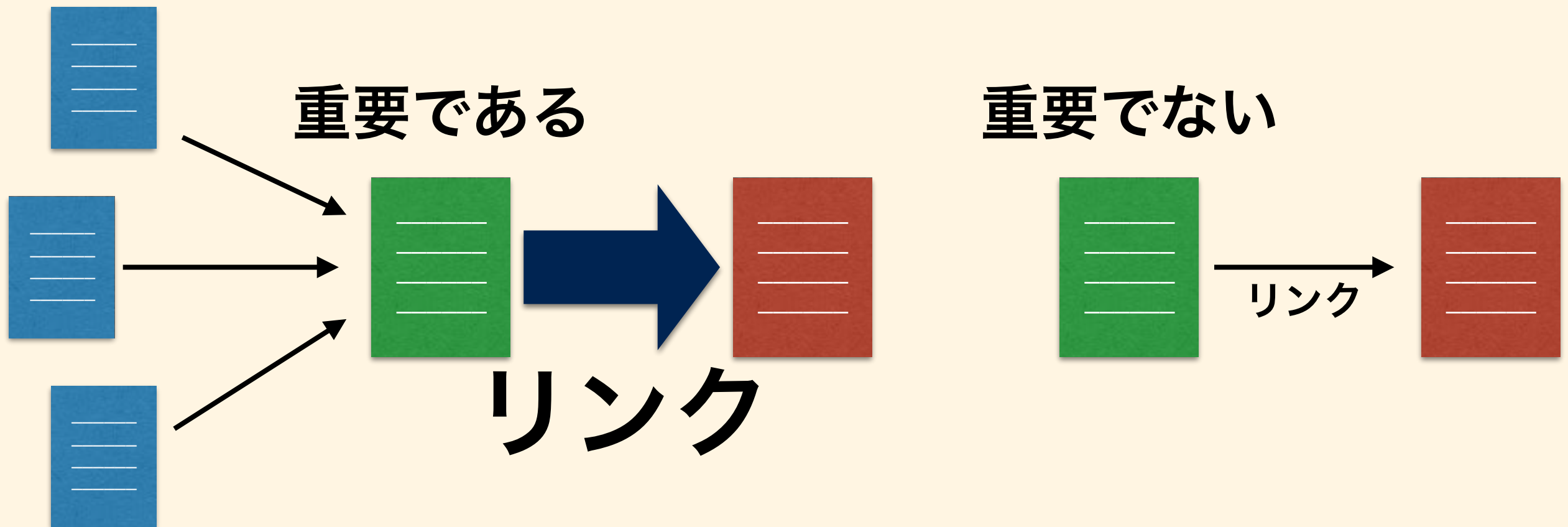
- ・多くのページからリンクされるページは重要である



PageRank 3つの原則

原則 2

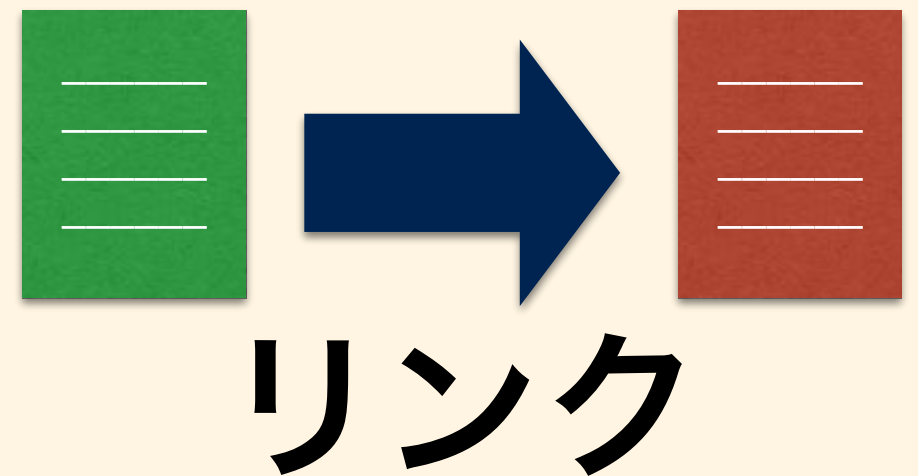
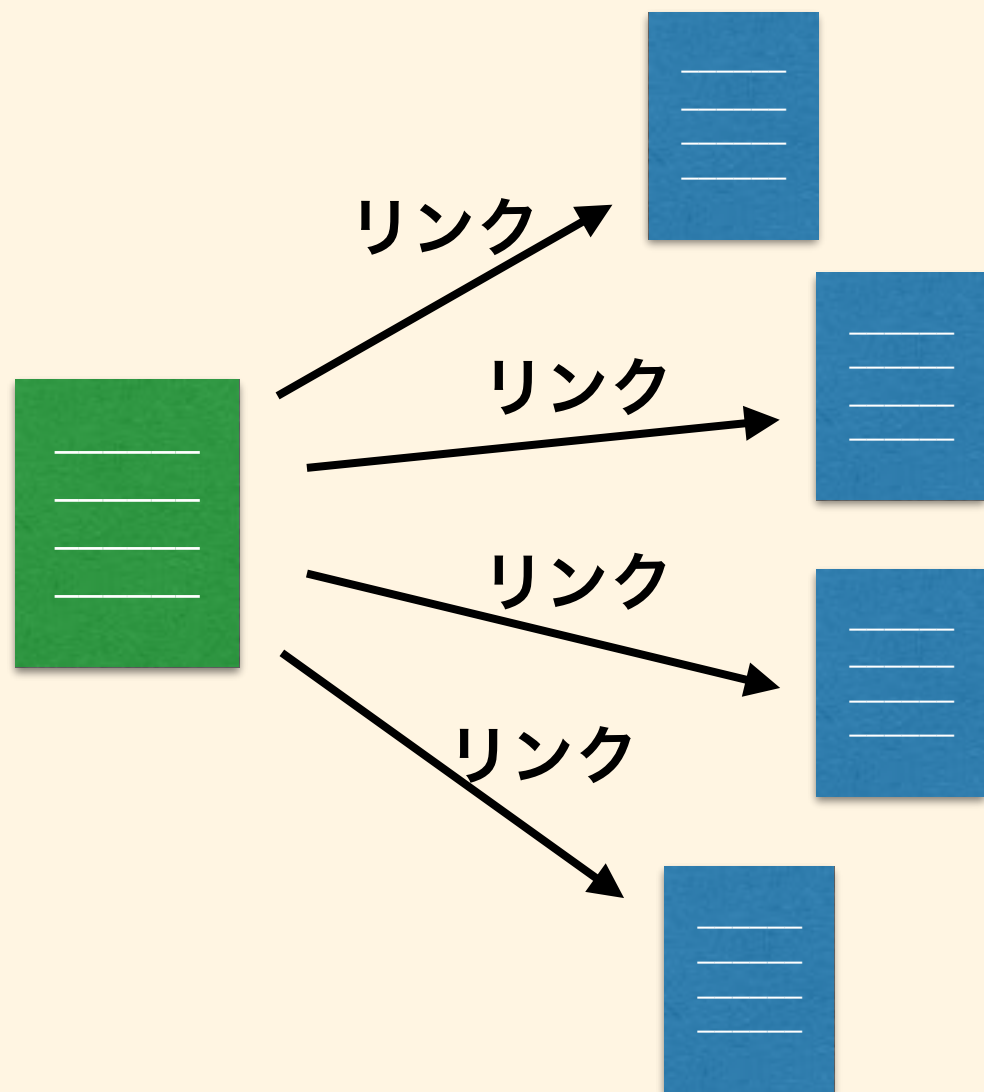
- ・ 重要なページからのリンクに、より価値がある



PageRank 3つの原則

原則 3

- ・ 多数へのリンクより、少数へのリンクに価値がある

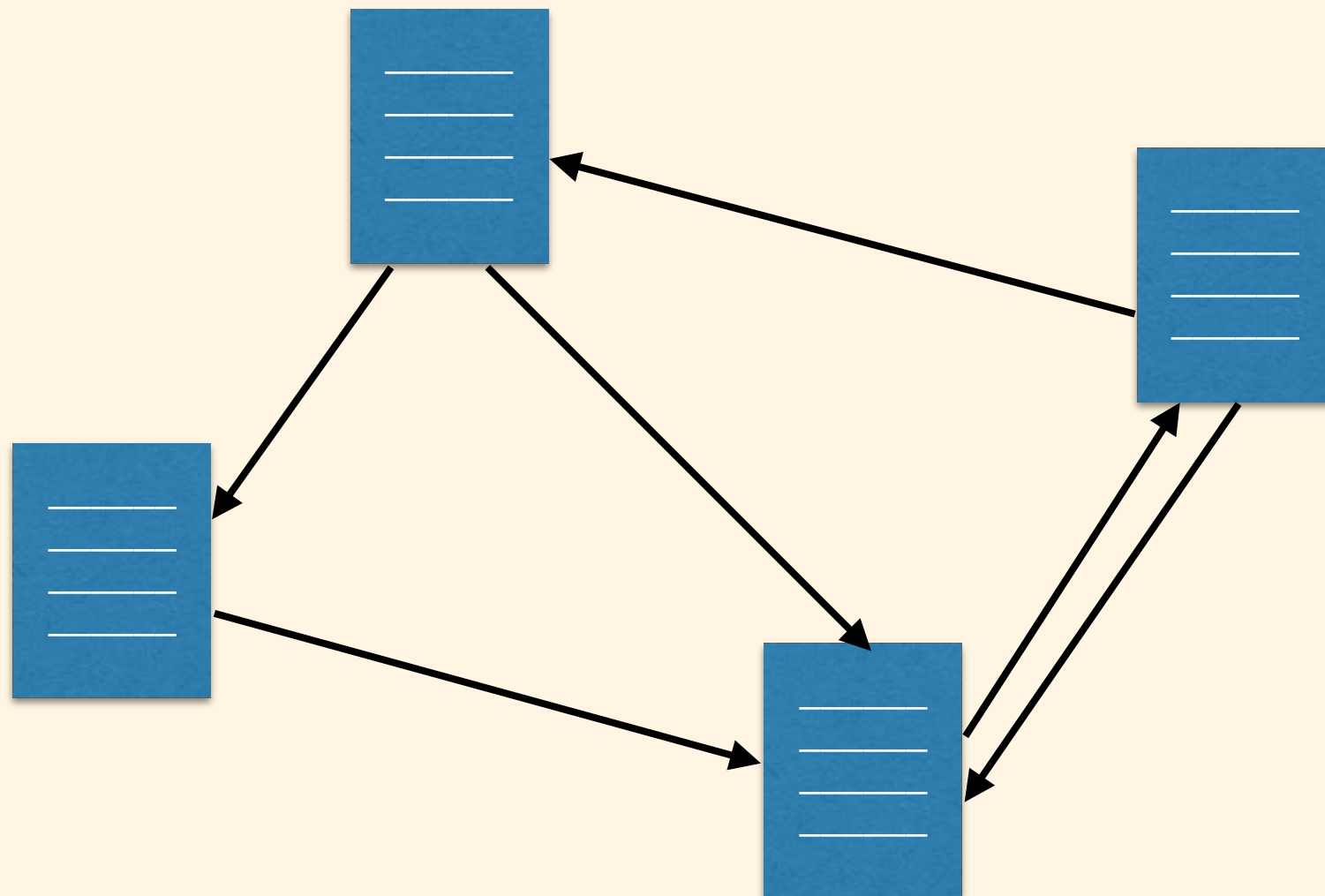


この3つの原則を元に
重要度を計算するのが
PageRankアルゴリズム

PageRank アルゴリズム

初期条件

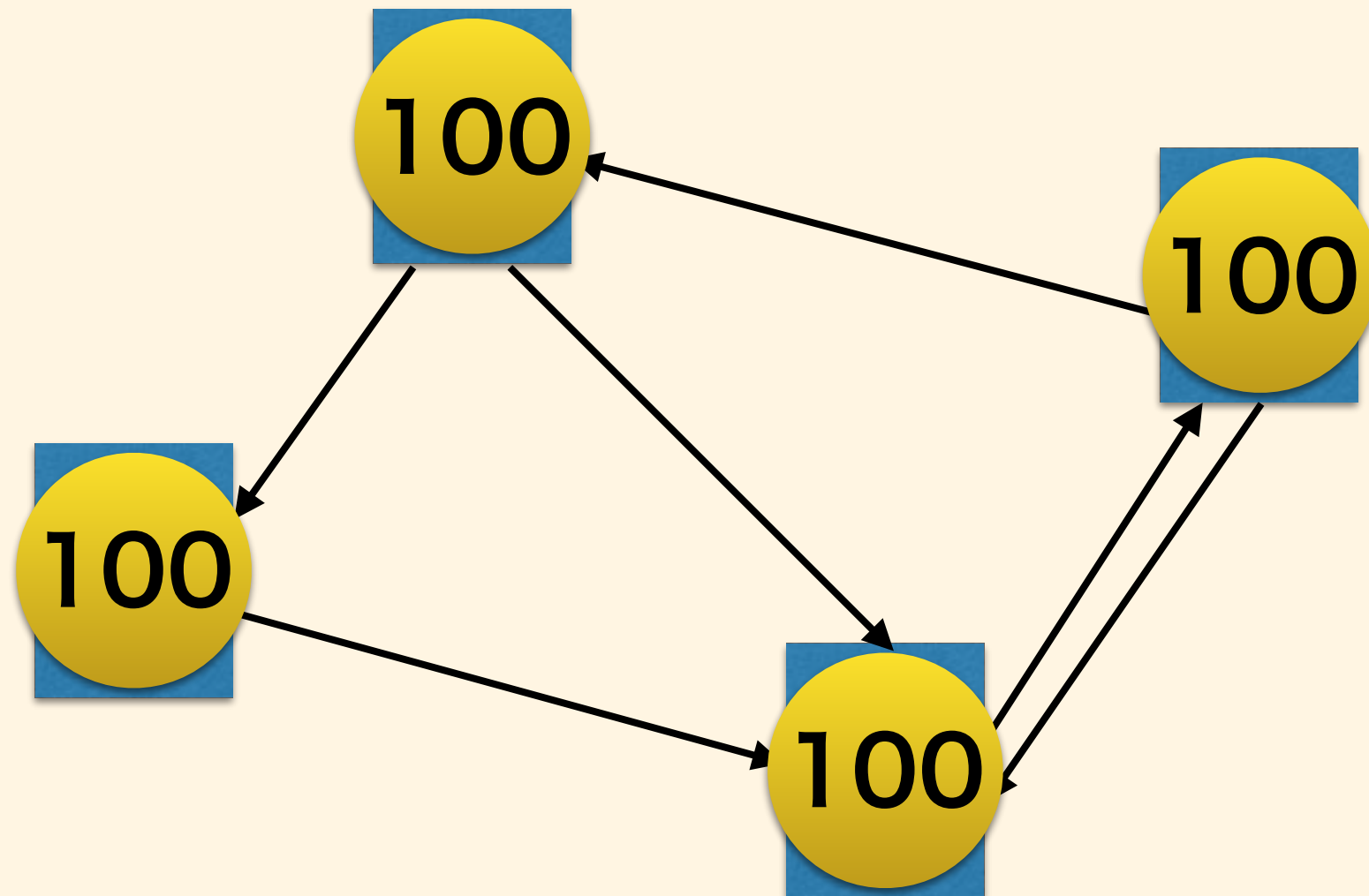
- ・ 全てのページに同じスコアを与える



PageRank アルゴリズム

初期条件

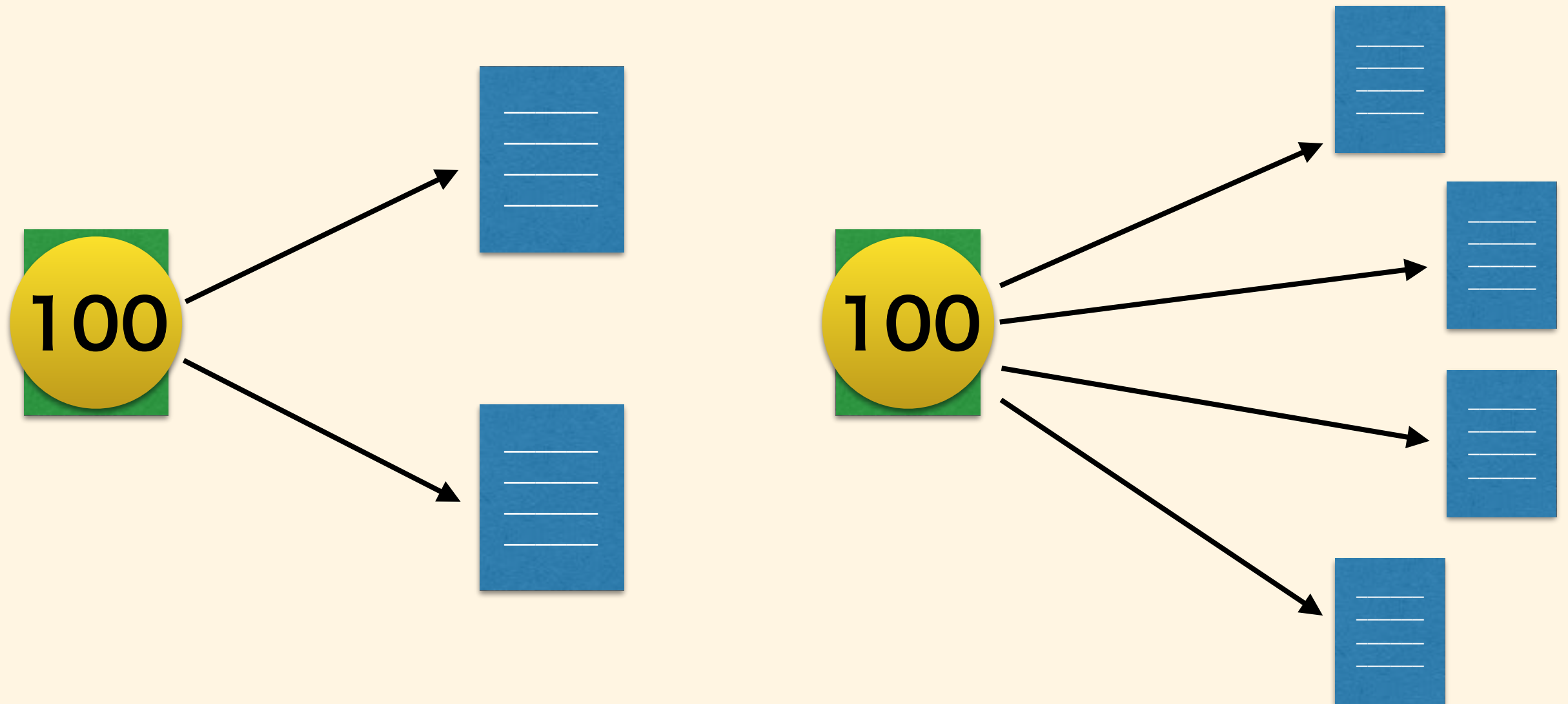
- ・ 全てのページに同じスコアを与える



PageRank アルゴリズム

スコアの計算

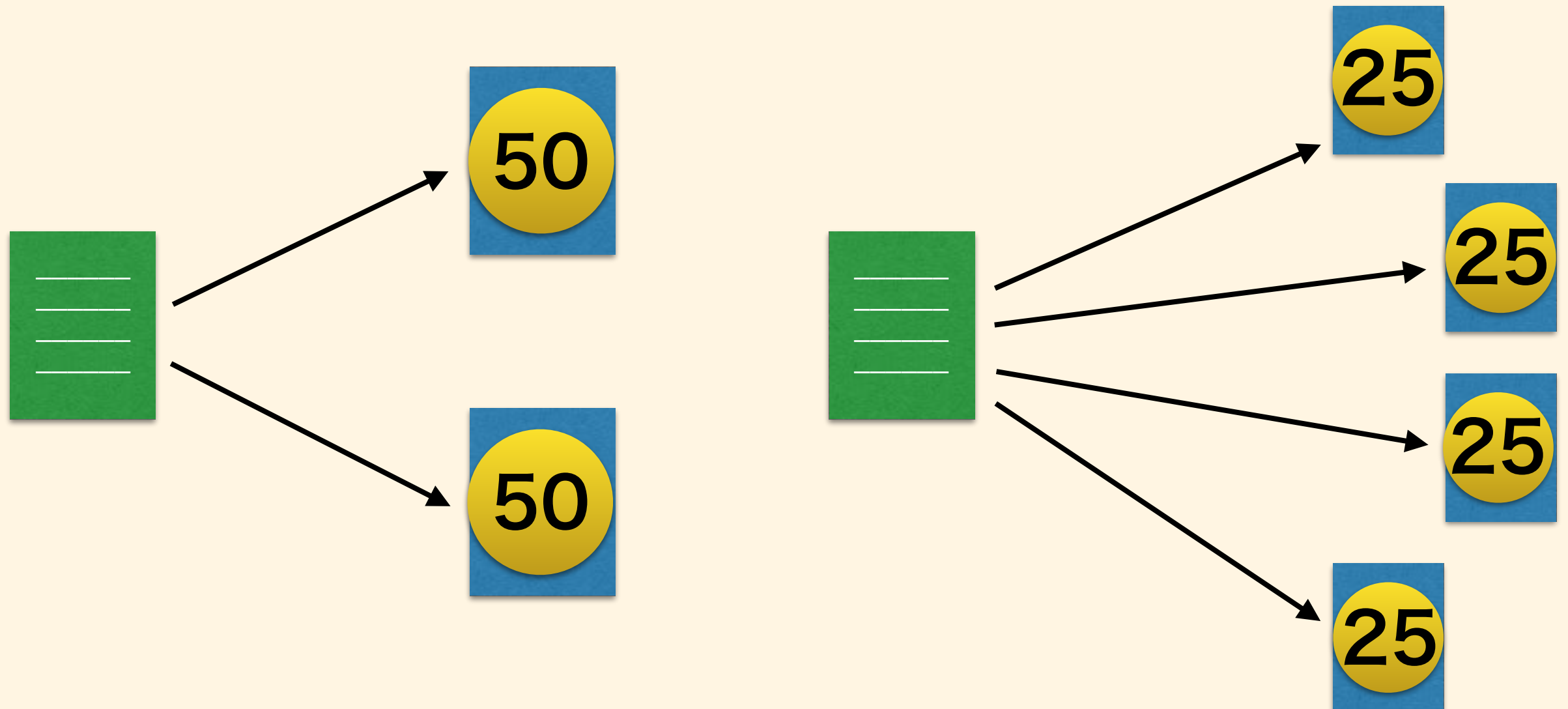
- ・ 自分のスコアをリンク先のページに等しく分配



PageRank アルゴリズム

スコアの計算

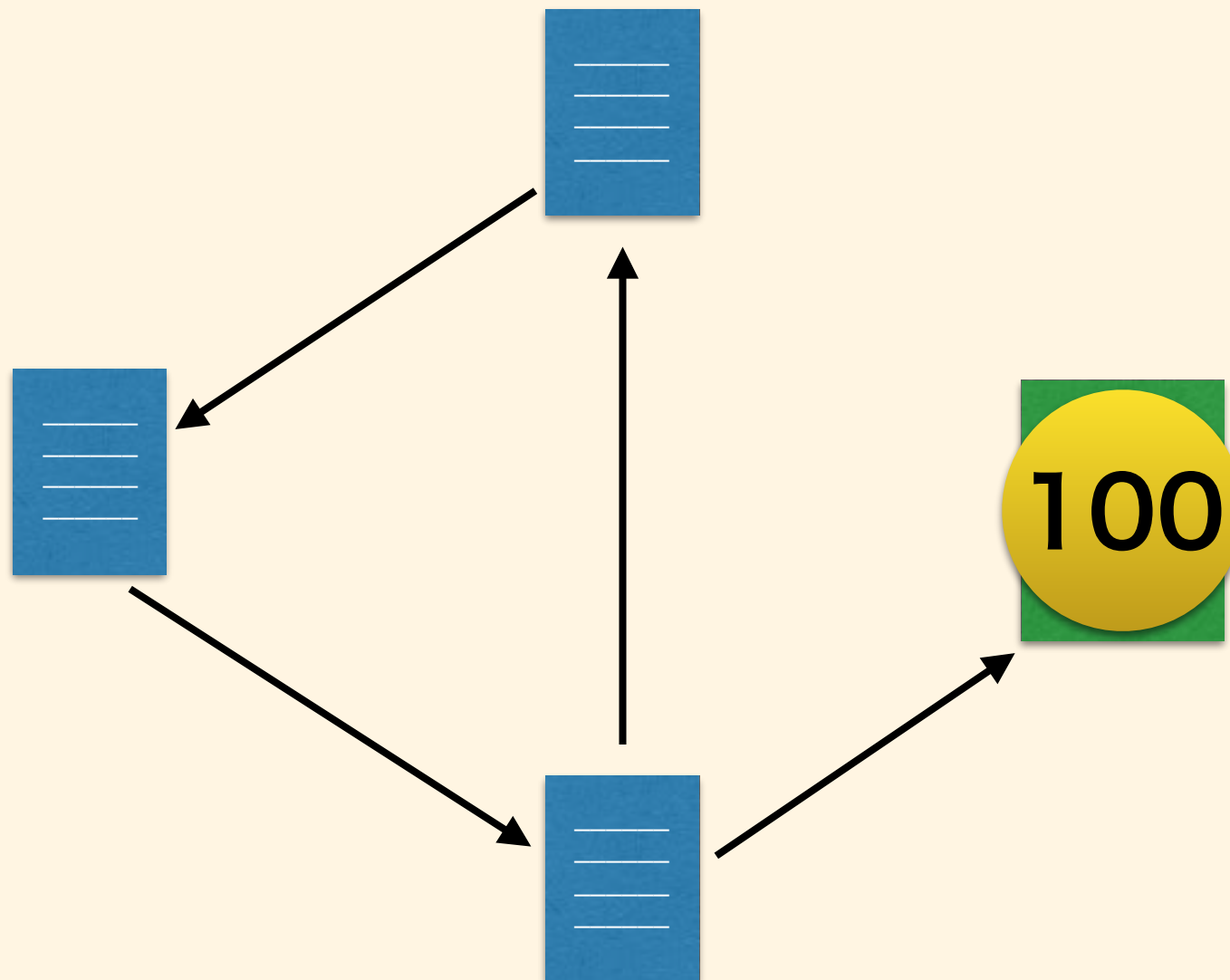
- ・ 自分のスコアをリンク先のページに等しく分配



PageRank アルゴリズム

スコアの計算

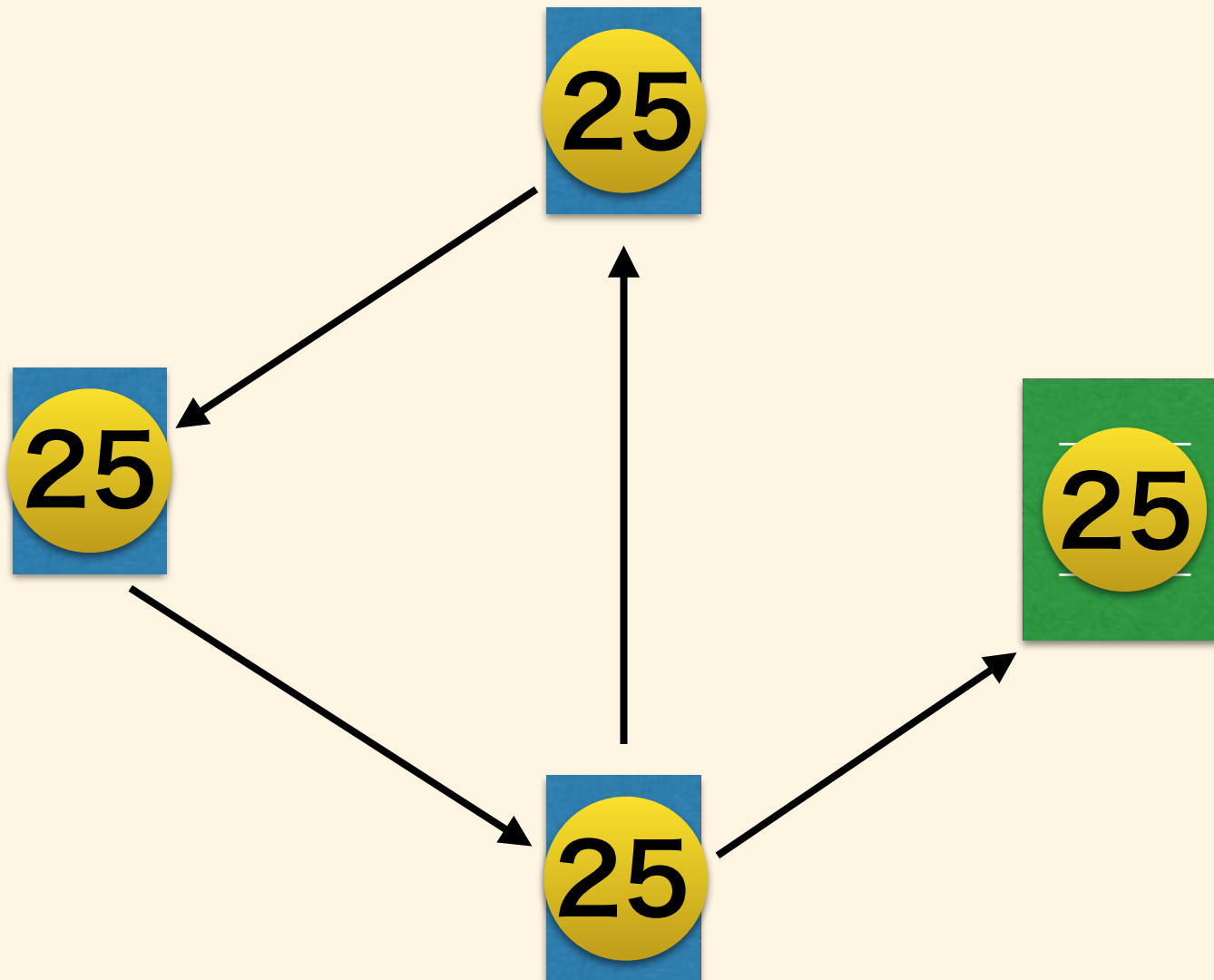
- ・リンク先が無い場合は、全ページに等しく分配



PageRank アルゴリズム

スコアの計算

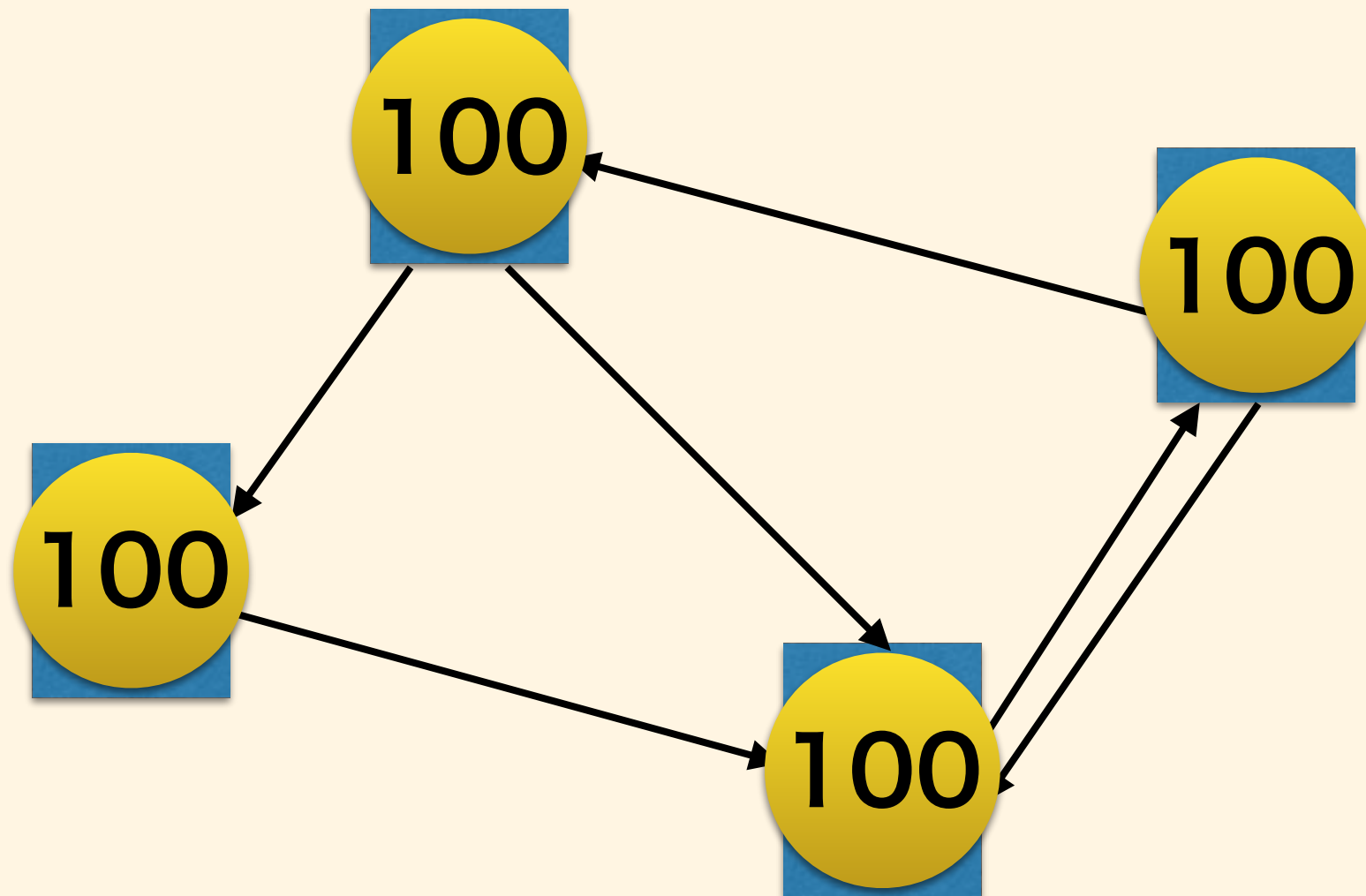
- ・リンク先が無い場合は、全ページに等しく分配



PageRank アルゴリズム

スコアの計算

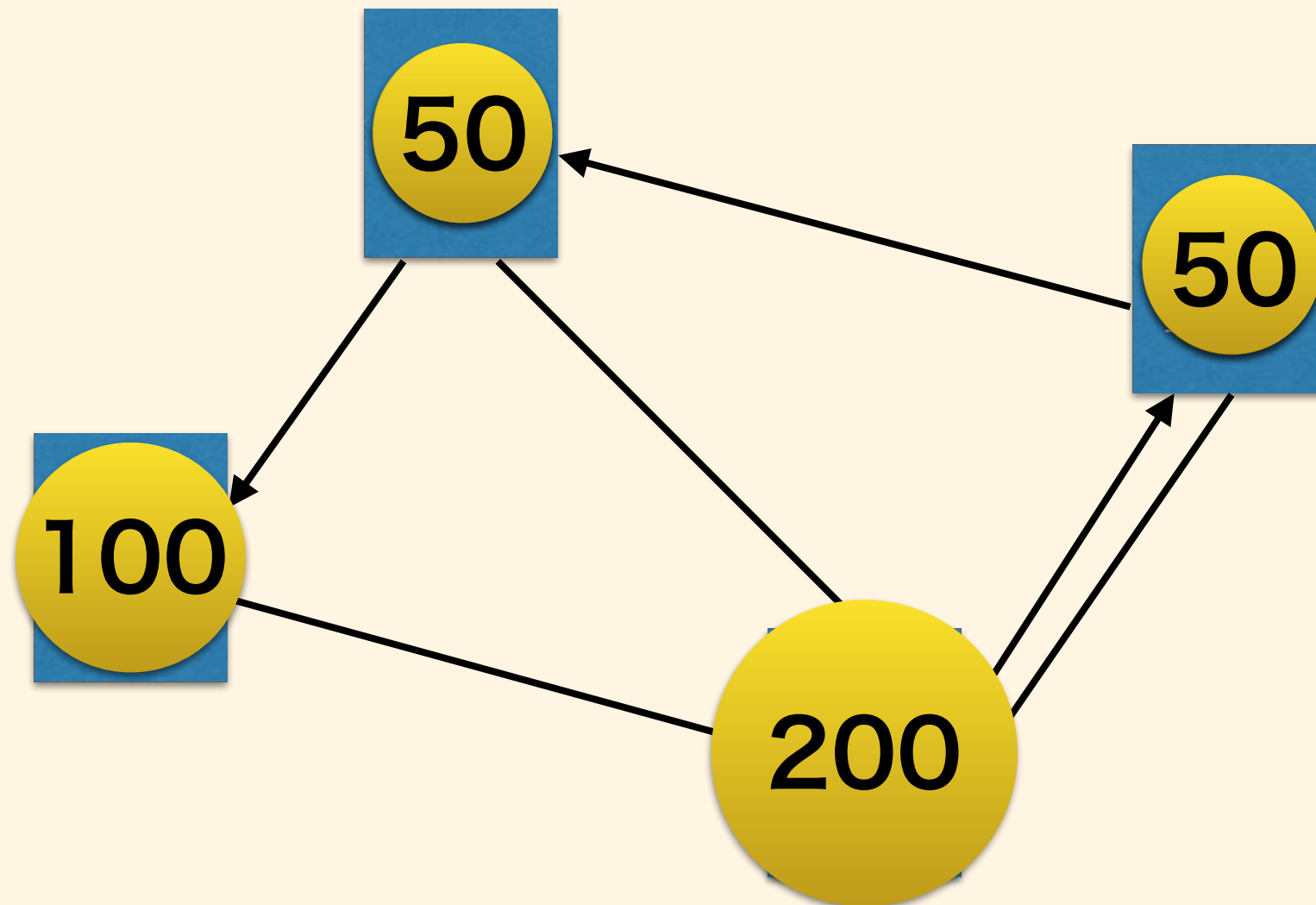
- ・ これらを全ページが同時に行う



PageRank アルゴリズム

スコアの計算

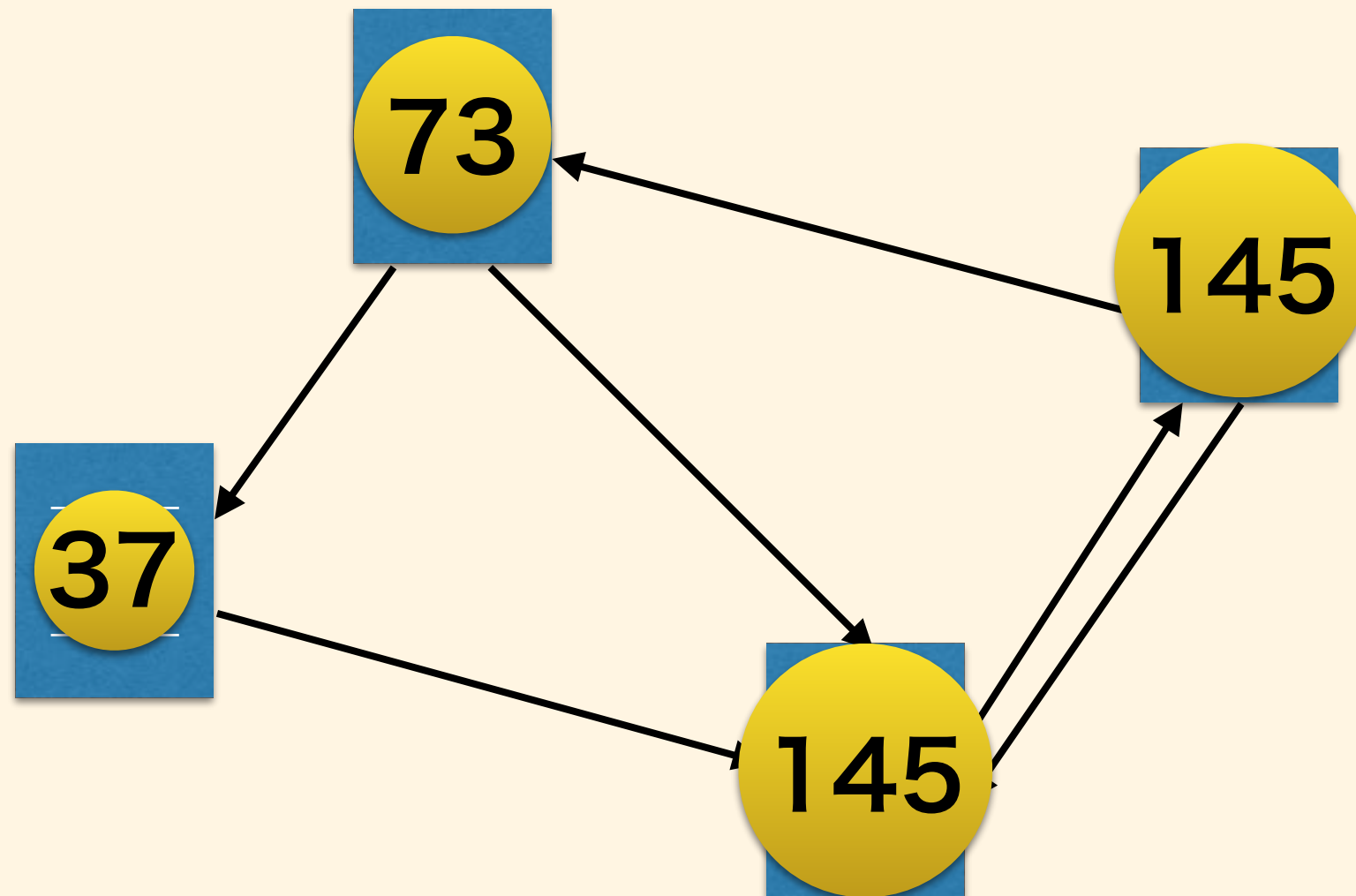
- ・ これらを全ページが同時に行う



PageRank アルゴリズム

スコアの計算

- ・ この計算を何度も繰り返し、値が大きいほど重要



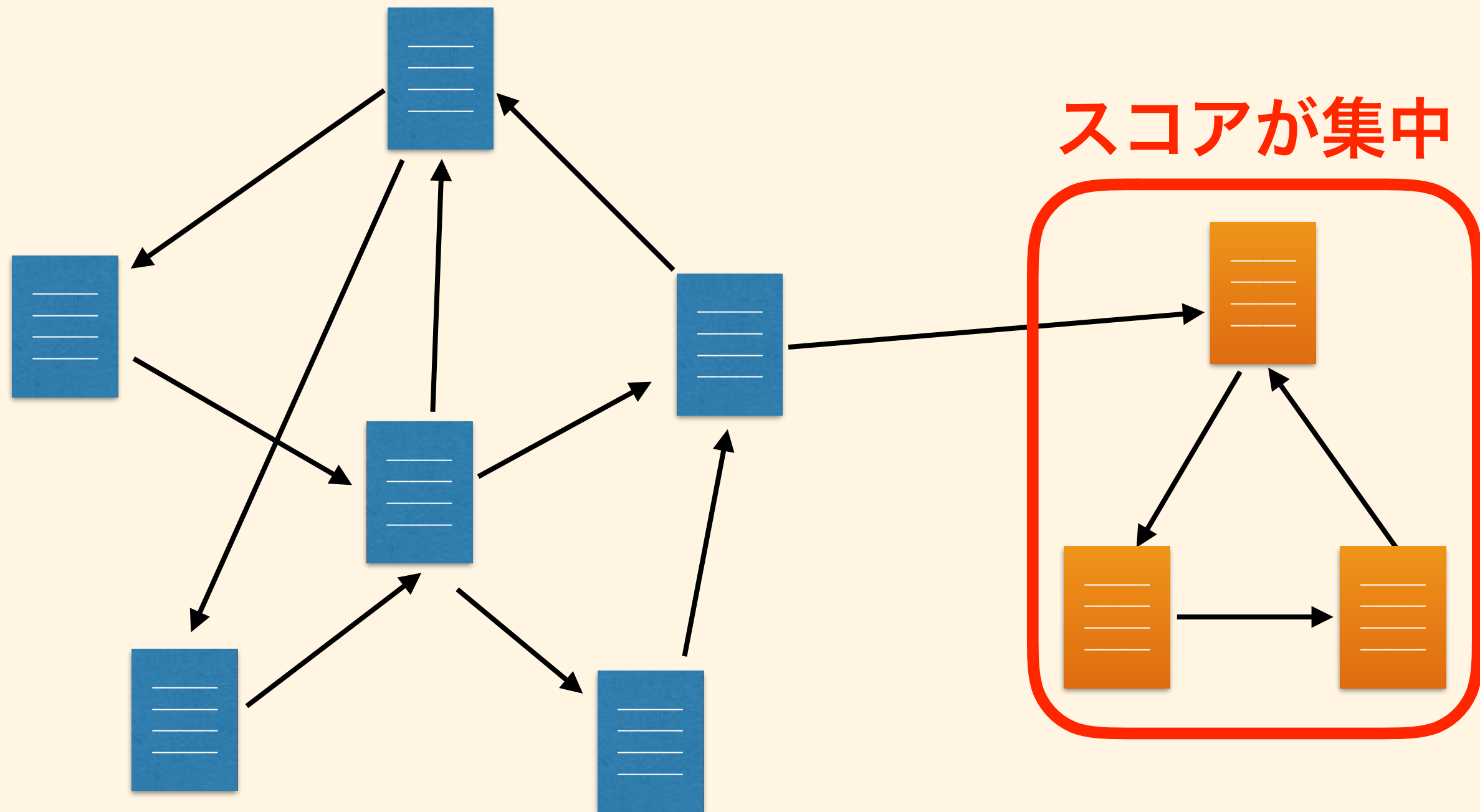
Processingで作ってみた

<http://bit.ly/PRAdemo>

問題点

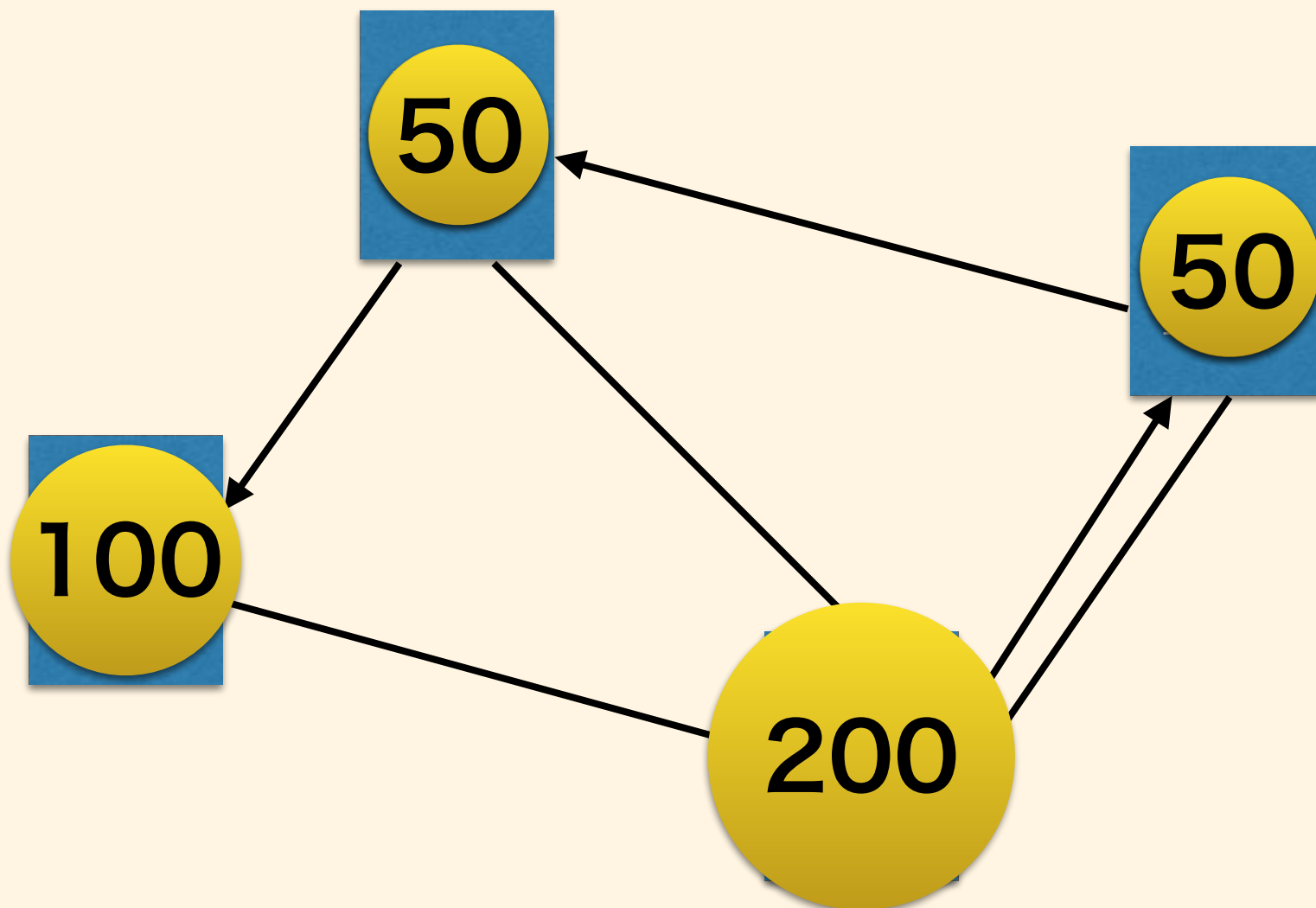
問題点

- 以下のようなループがあると、スコアがそこに集中



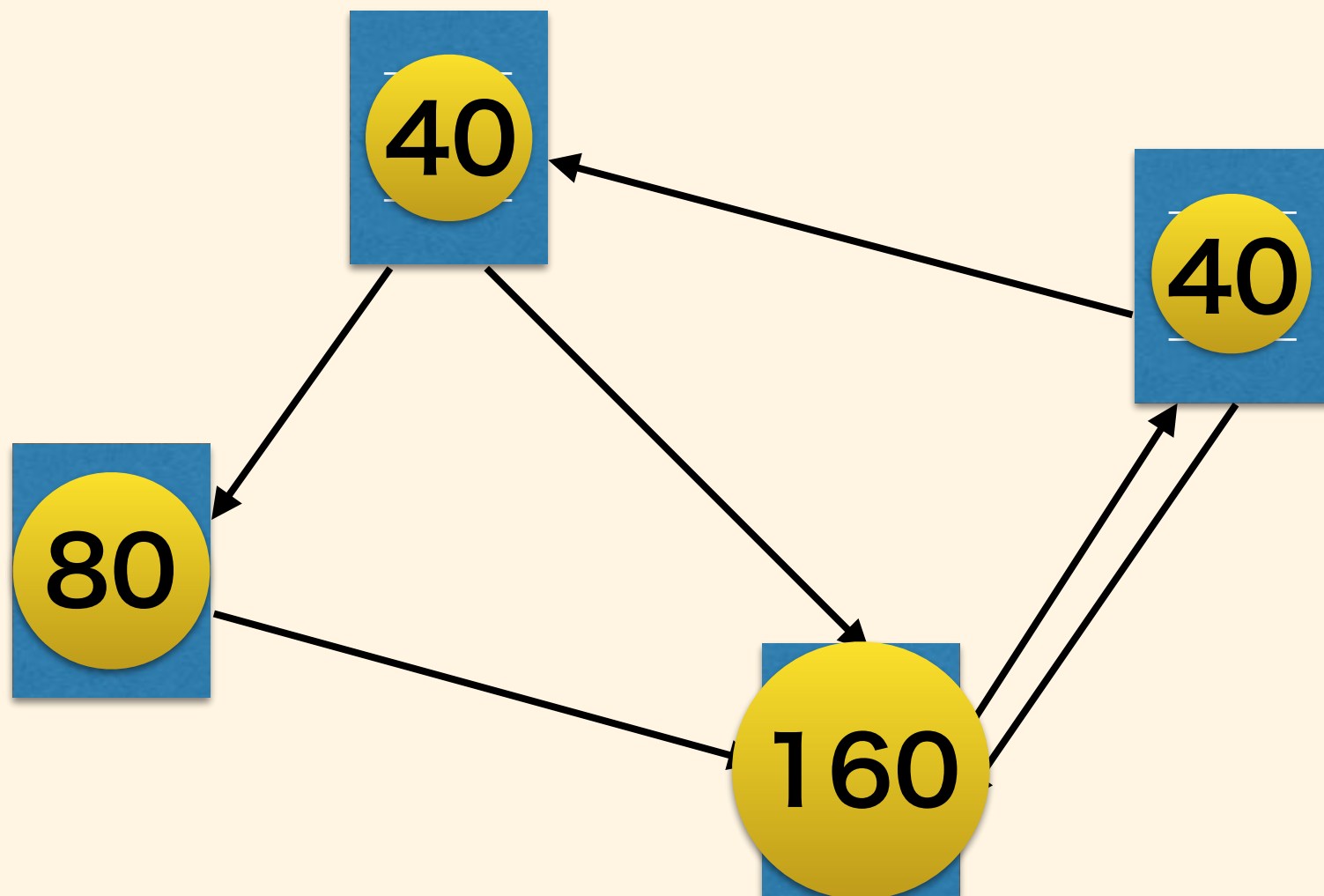
解決策

- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配

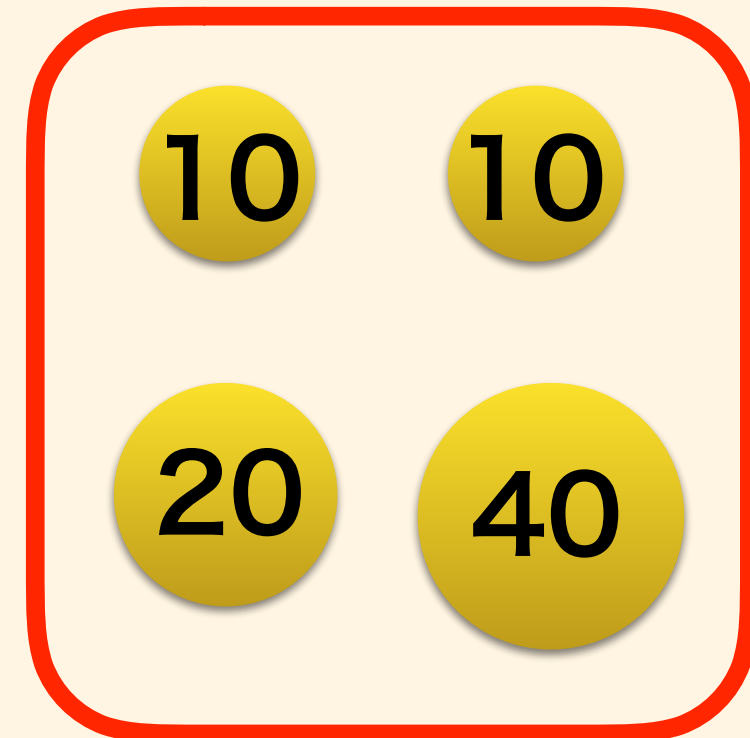


解決策

- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配

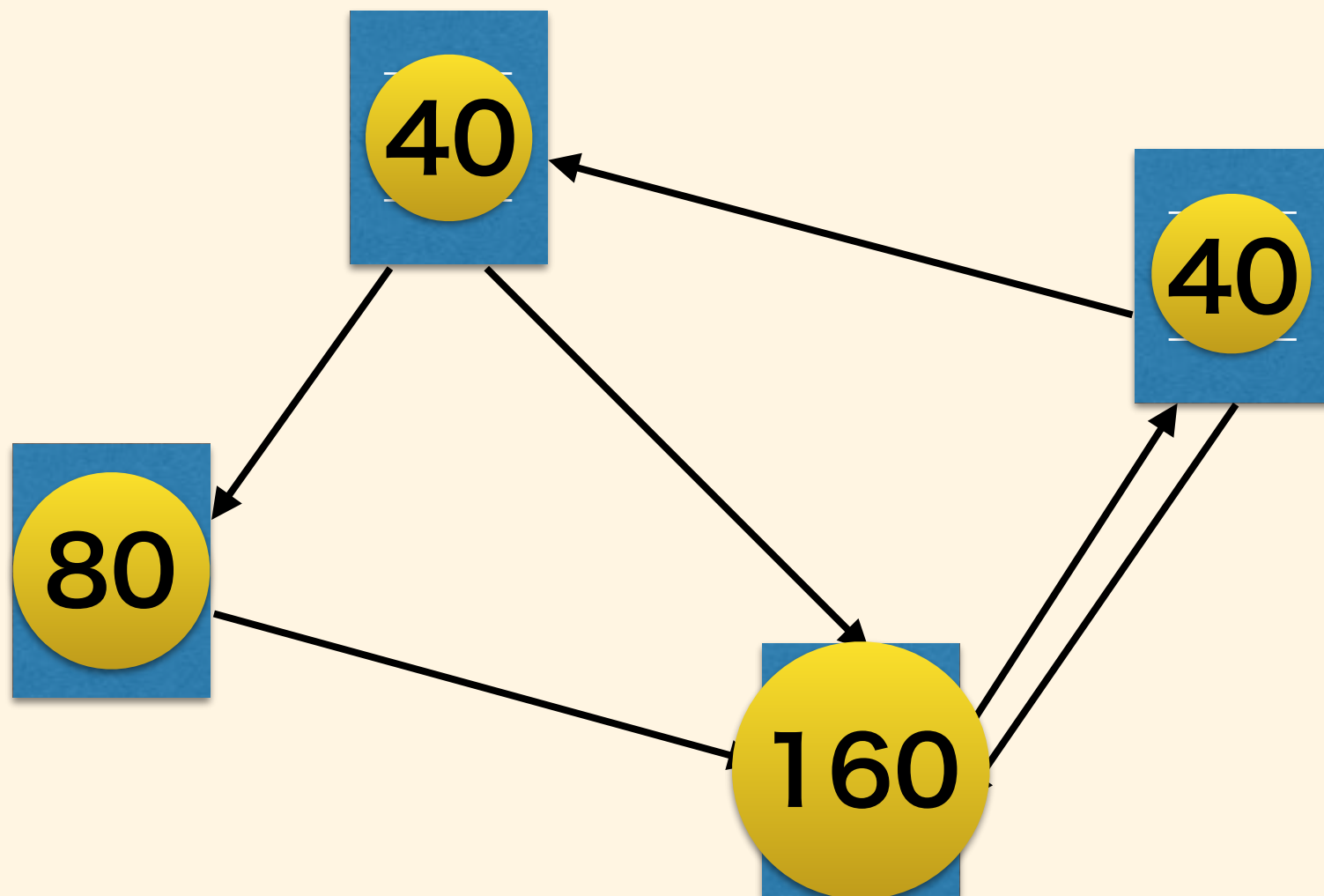


それぞれから
20%ずつ徴収

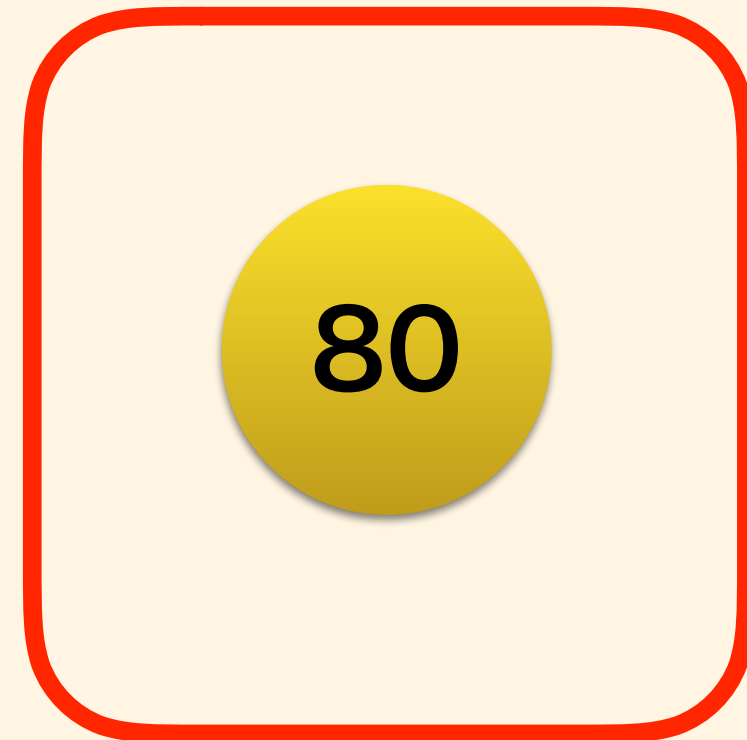


解決策

- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配



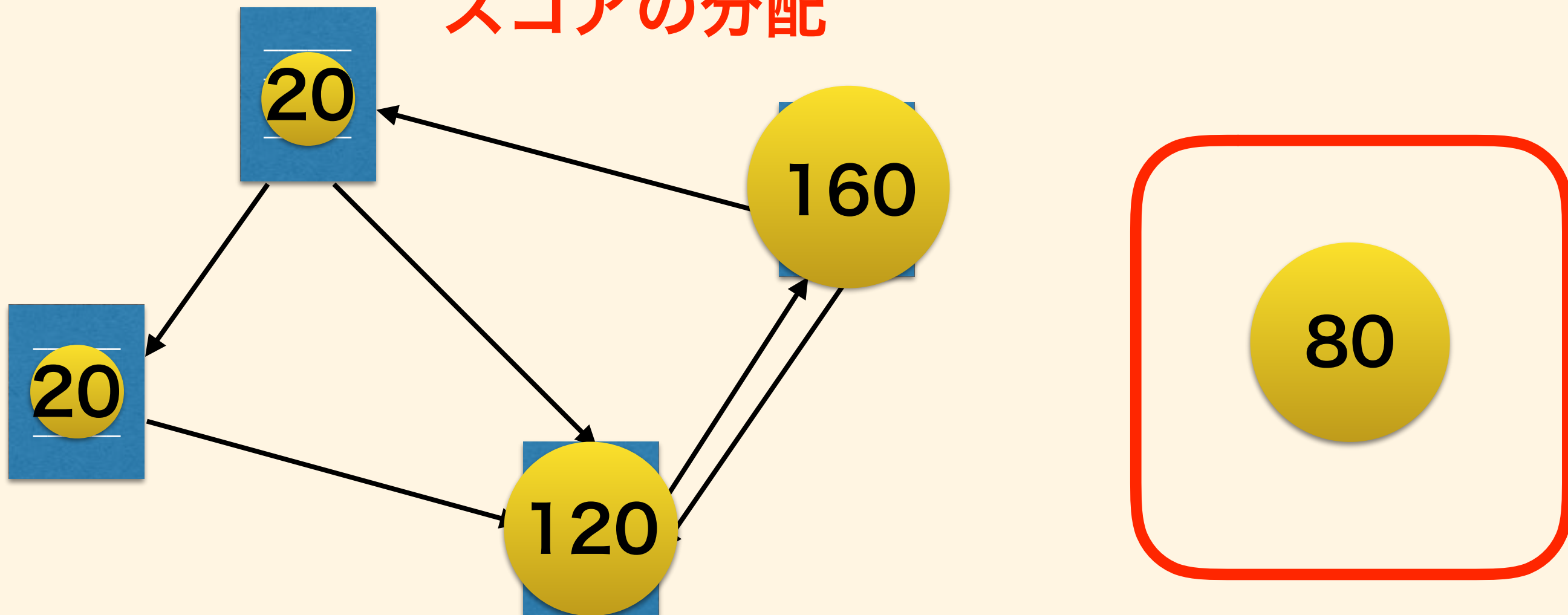
それぞれから
20%ずつ徴収



解決策

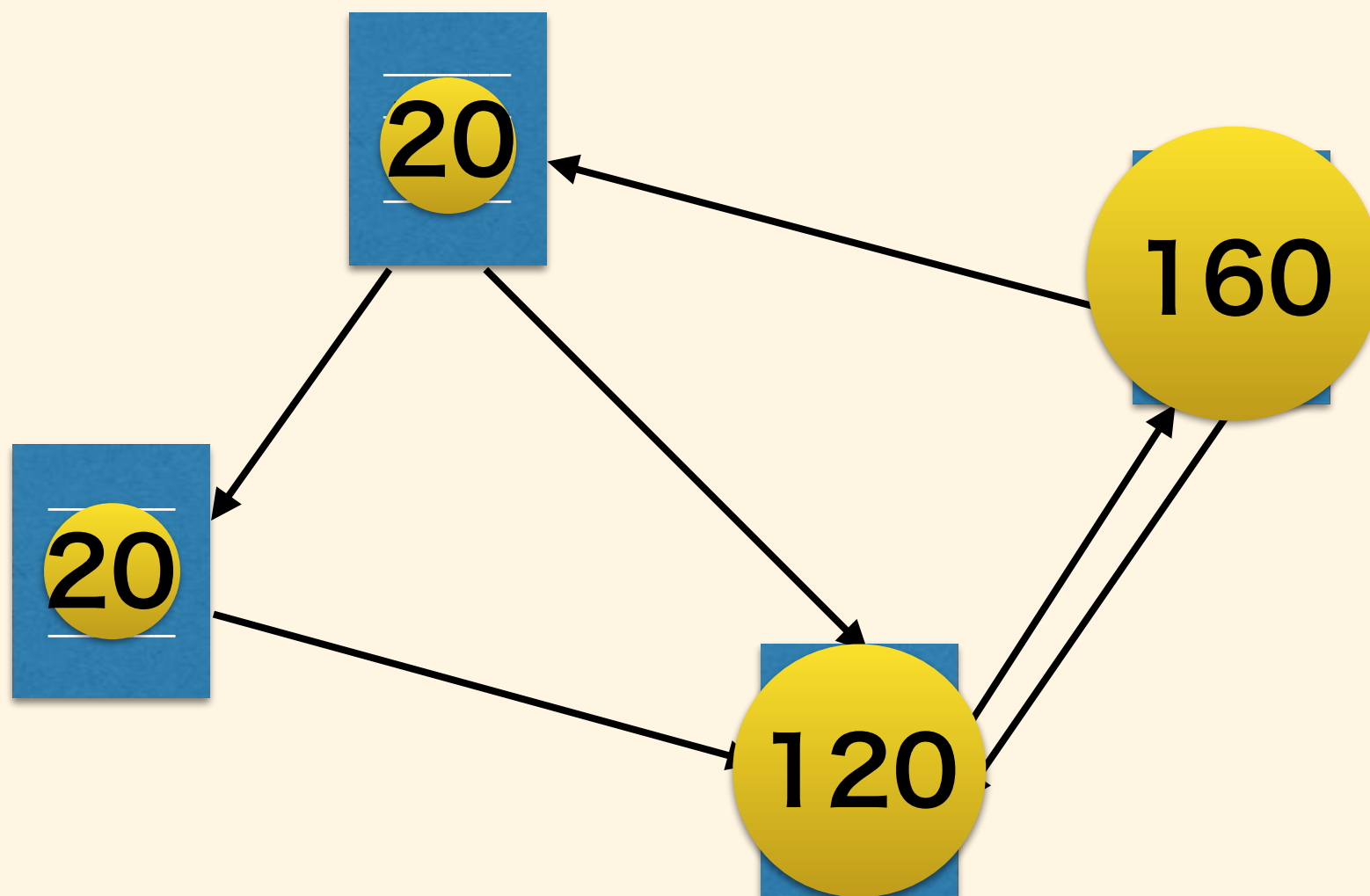
- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配

スコアの分配

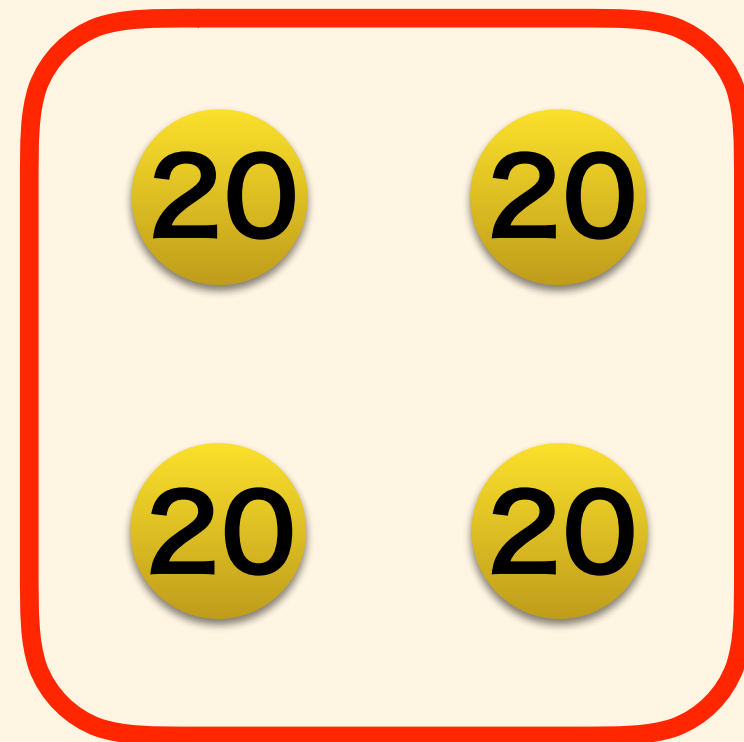


解決策

- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配

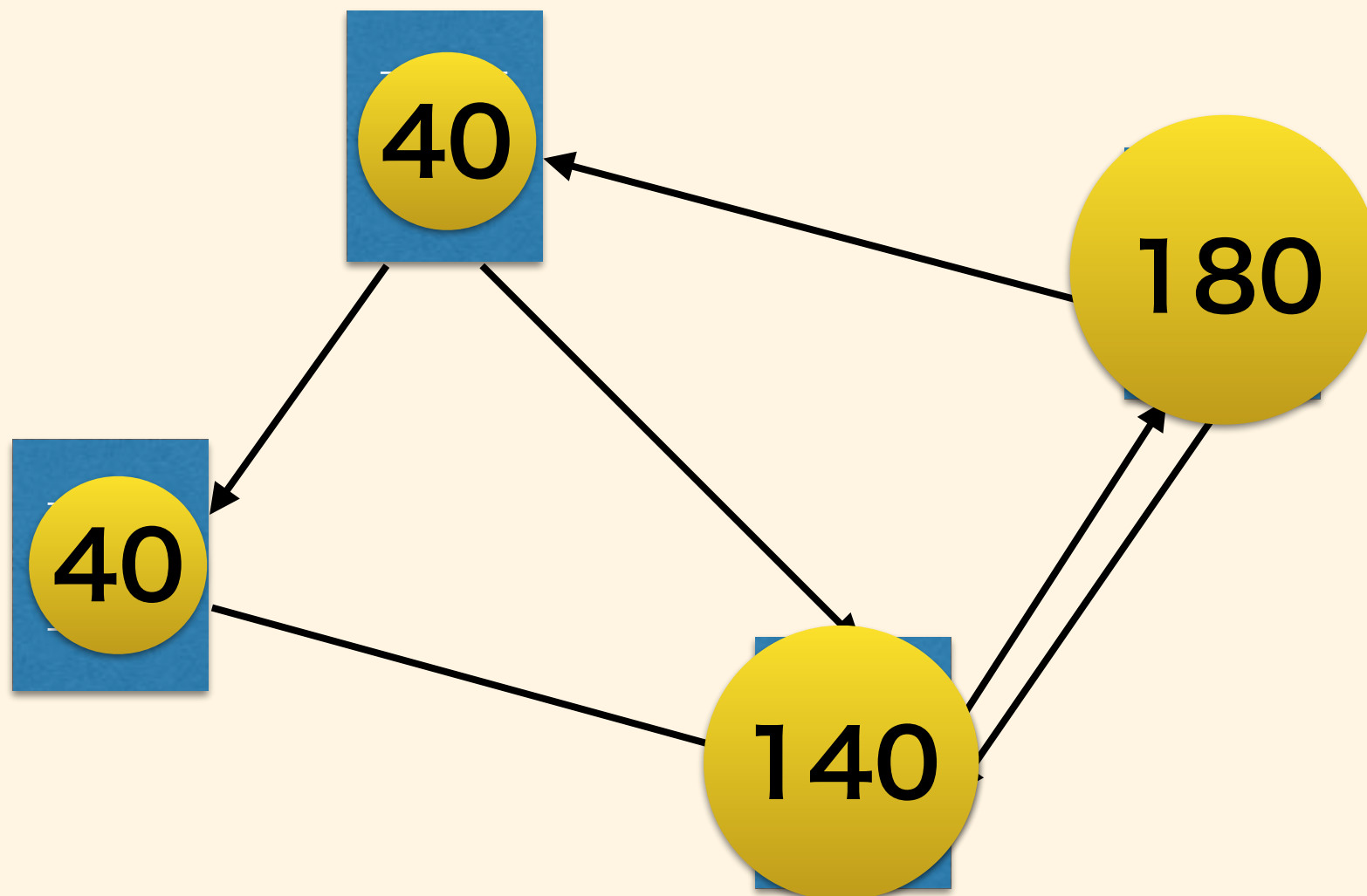


均等に分配

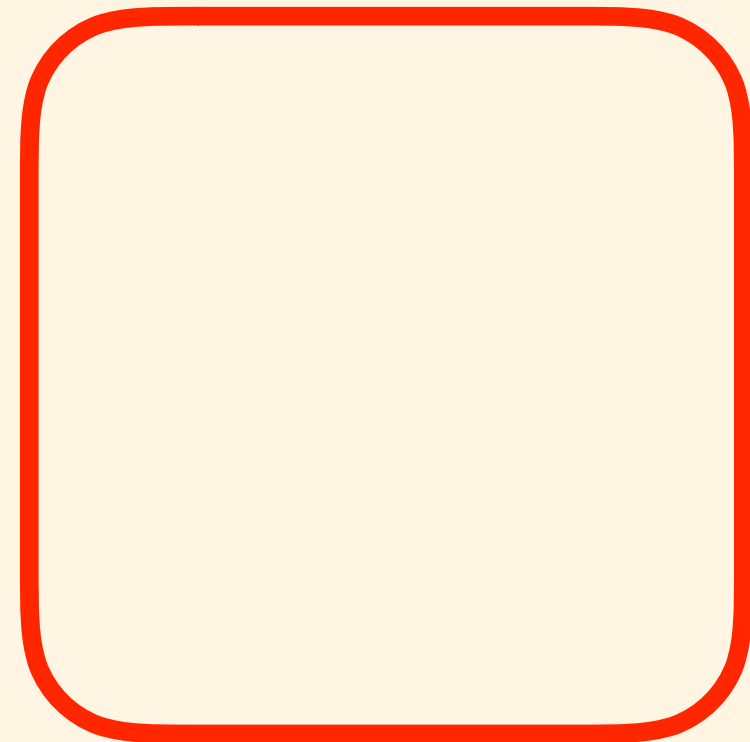


解決策

- それぞれのページから一定の割合を徴収
- スコア分配後に、徴収分を均等に分配



均等に分配



Processingで作ってみた（再掲）

<http://bit.ly/PRAdemo>

PageRankアルゴリズムの式

$$r_u = d \sum_{v \in B_u} \frac{r_v}{n_v} + (1 - d)$$

r_u : ページuのPageRankスコア

B_u : ページuへのリンクを持つページの集合

n_v : ページvからのリンクの総数

d : ダンピングファクター ($0 \leq d < 1$)

PageRankアルゴリズムの特徴

- ページが増えると計算量が爆増
→行列演算で計算を高速化
- 計算結果が必ず収束する
→どんなリンク構造でもランキングが決まる
- このアルゴリズムの説明として
“ランダムサーファーマodel”という考え方もある
→結局同じ式になる。詳しくはGoogleで検索！

Google検索アルゴリズムの実際

- この考え方をベースに様々な改良が加えられている
→ 計算量削減、スパム判定、etc...
- **PageRankアルゴリズム以外の指標も活用**
→ 様々な指標からページの重要度を総合的に判断

参考文献

PageRank アルゴリズム およびそれに関連する研究について

<http://www.kentmiyajima.com/document/pagerank.pdf>

PageRankについて調べてみた

https://serihiro.hatenablog.com/entry/2016/12/02/000000_1

GoogleのPageRankアルゴリズムにみるイノベーションのパターン

<http://d.hatena.ne.jp/s-feng/20070127/1169913763>