CS246 Exam 2025 — Question 6: PageRank and TrustRank (20 points)

Parts 1–3: Three Graphs

Graph A (6 nodes): A → B → C → D → E ↔ F (chain with bidirectional end) **Graph B** (4 nodes): Fully connected (all pairs linked in both directions) **Graph C** (4 nodes): Directed cycle A → B → C → D → A

Parameters: even initialization (1/n), teleportation $1 - \beta = 0.15$.

Part 1 (2 points) — Fastest Convergence

Answer: (B) Graph B and (C) Graph C

Both graphs B and C converge in **one iteration** because all nodes are perfectly symmetric. With even initialization (each node starts at 1/n), the initial PageRank vector is already the stationary distribution — every node has identical in-degree and out-degree structure, so one iteration confirms the uniform distribution.

Part 2 (2 points) — Slowest Convergence

Answer: (A) Graph A

Graph A is a non-symmetric chain (A→B→C→D→E↔F) with a spider-trap-like structure at the end (the E↔F bidirectional link accumulates rank). The asymmetry means the uniform initialization is far from the stationary distribution, requiring many iterations for rank to flow down the chain and stabilize.

Part 3 (2 points) — Adding Back-Links Instead of Teleportation

Question: Add a link from each dead-end node to its parent(s). Does this solve dead ends?

Answer: No. Two problems: (1) **Artificial bias** — nodes pointing to dead ends get disproportionate PageRank boost. (2) **Spider traps** — if A links to dead-end B and B links back to A, the pair {A, B} forms a trap that absorbs rank. This "fix" introduces a new pathology.

Part 4: TrustRank on a 5-Node Graph

Graph Structure

- Node 1 → 2 (out-degree 1)
- Node 2 → 3 (out-degree 1)
- Node 3 → 1, 4 (out-degree 2)
- Node 4 → 1, 5 (out-degree 2)

- Node 5 → 1 (out-degree 1)

Node 1 = Trusted, Nodes 3 & 4 = Spam, Nodes 2 & 5 = Unknown.
Parameters: β = 0.8, 1−β = 0.2, n = 5, initial r = 1/5 for all, spam threshold = 0.5.

Part 4a (2 points) — Column-Stochastic Matrix M

M_{i,j} = fraction of j's out-links going to i:

$$M = \begin{pmatrix} 0 & 0 & 1/2 & 1/2 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 \end{pmatrix}$$

Every column sums to 1 (no dead ends). ✓

Part 4b (4 points) — Standard PageRank (One Iteration)

With no dead ends: $r'_i = \beta \sum_j M_{ij} \cdot r_j + \frac{1-\beta}{n}$

Every node receives the teleportation base of (1−β)/n = 0.2/5 = **1/25**. The link contributions (β × M_{ij} × 1/5) are:

Node	Incoming links (× β/5)	Link contrib	+ 1/25	= r'
1	from 3 (1/2), 4 (1/2), 5 (1) → sum = 2	0.32	+ 0.04	**9/25**
2	from 1 (1) → sum = 1	0.16	+ 0.04	**5/25**
3	from 2 (1) → sum = 1	0.16	+ 0.04	**5/25**
4	from 3 (1/2) → sum = 0.5	0.08	+ 0.04	**3/25**
5	from 4 (1/2) → sum = 0.5	0.08	+ 0.04	**3/25**

$$\boxed{r_1 = \frac{9}{25},\ r_2 = \frac{5}{25},\ r_3 = \frac{5}{25},\ r_4 = \frac{3}{25},\ r_5 = \frac{3}{25}}$$

Verification: 9 + 5 + 5 + 3 + 3 = 25 → sums to 1. ✓

Part 4c (4 points) — TrustRank (One Iteration)

Teleportation goes **only to node 1.** Non-trusted nodes receive zero teleportation.

Node 1 (trusted): $r'^{+}_1 = \beta \sum_j M_{1j} \cdot r_j + (1 - \beta)$

$$r_1'^{+} = 0.8 \times \left(\frac{1}{2}\cdot\frac{1}{5} + \frac{1}{2}\cdot\frac{1}{5} + 1\cdot\frac{1}{5}\right) + 0.2 = 0.8 \times \frac{2}{5} + 0.2 = 0.32 + 0.20 = \frac{13}{25}$$

Nodes 2–5 (not trusted): $r_i'^{+} = \beta \sum_j M_{ij} \cdot r_j$ (link contributions only)

Node	Calculation	$\mathbf{r^+}$
2	$0.8 \times 1 \times 1/5$	**4/25**
3	$0.8 \times 1 \times 1/5$	**4/25**
4	$0.8 \times 1/2 \times 1/5$	**2/25**
5	$0.8 \times 1/2 \times 1/5$	**2/25**

$$\boxed{r_1^+ = \frac{13}{25},\ r_2^+ = \frac{4}{25},\ r_3^+ = \frac{4}{25},\ r_4^+ = \frac{2}{25},\ r_5^+ = \frac{2}{25}}$$

Verification: 13 + 4 + 4 + 2 + 2 = 25 → sums to 1. ✓

Part 4d (4 points) — Spam Mass Classification

$$\text{spam mass}(p) = \frac{r_p - r_p^+}{r_p}$$

Node 2:

$$\text{spam mass}(2) = \frac{5/25 - 4/25}{5/25} = \frac{1}{5} = 0.2 < 0.5 \quad \Rightarrow \quad \textbf{Not Spam}$$

Node 5:

$$\text{spam mass}(5) = \frac{3/25 - 2/25}{3/25} = \frac{1}{3} \approx 0.333 < 0.5 \quad \Rightarrow \quad \textbf{Not Spam}$$

Both nodes have low spam mass — most of their PageRank comes from trusted sources, since trust from node 1 propagates well through the graph's link structure.