A Graph-Theoretic Approach to Multitasking

Noga Alon* (Tel Aviv), **Daniel Reichman*** (Berkeley), **Igor Shinkar*** (Berkeley), **Tal Wagner*** (MIT), **Sebastian Musslick** (Princeton), **Jonathan D. Cohen** (Princeton), **Thomas L. Griffiths** (Berkeley), **Biswadip Dey** (Princeton), **Kayhan Ozcimder** (Princeton)

(These authors contributed equally)*

Multitasking Control Demanding Tasks

- **Severely limited ability** *(Posner & Snyder; Shiffrin & Schneider)*
- **Reason? Still unclear**

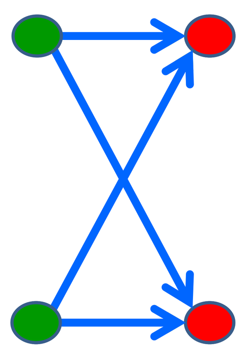
Our Approach

- Builds on **connectionist models** *(Cohen, Dunbar & Mcclelland)*
- Presumes multitasking limitations emerge from localized processes **accessing the same representation at the same time**
- Graph theoretical

Our Formalism

Bipartite graph G=(A,B,E), |A|=|B|=n:

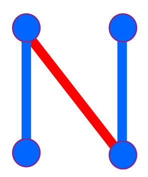
- Side **A**: Inputs (colors, words, features)
- Side **B**: Outputs (simple actions like naming, pointing)
- **Edges are tasks**
  - **Task**: (input) ➡ (output), e.g. color naming



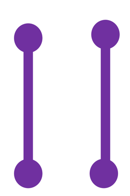
Assumptions

Which sets of tasks (edges) can be multitasked?

- Necessary condition: **Edges form a matching**
  - i.e., have no mutual endpoints
  - Extensive **empirical support** from Cognitive Psychology
  - Exclusive-Read-Exclusive-Write (**EREW**) in Computer Science



- Sufficient condition: **Edges form an induced matching**
  - i.e., no other edges between endpoints
  - *(Feng et al; Musslick et al)*
  - **PDP:** Nodes **propagate signals** to **all** neighbors
  - Arises in **communication models** *(Birk, Linial & Meshulam)*

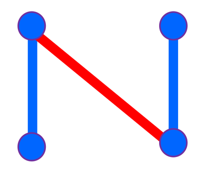


Our Measure of Multitasking Capacity

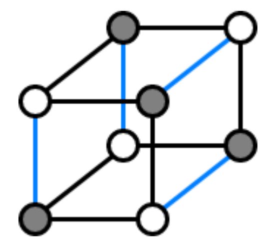
Given a graph $G=(A,B,E)$, $|A|=|B|=n$, and $\alpha \in (0,1]$,

> G is an **α-multitasker** if **every matching** M contains an **induced matching** $M' \subset M$ of size $|M'| \geq \alpha|M|$.

Examples:



3-path
$\alpha=1/2$



3D hypercube
$\alpha=1/4$

Generally: Every d-regular graph satisfies $\alpha \geq 1/(2d)$.

Can we do better?

Questions

- Which graphs have **good** multitasking properties?
- What are the **limitations** of multitasking?
- Does **average degree constrain** multitasking?

Applications

- Choosing **architectures** of interconnected **neural networks** working **in parallel**
- Relationship between **over-connectivity** and **signal interference** *(Navlakha, Bar-Joseph & Barth)*

Main Result

Theorem: Let G be a d-regular bipartite graph.

> **If G is an α-multitasker, then $\alpha \leq 9/\sqrt{d}$.**

- **Nearly tight for perfect matchings**: There are d-regular graphs s.t. every perfect matching contains induced matching of relative size $\Omega(1/\sqrt{d \log d})$.

Proof of Main Result

- Set up auxiliary bipartite graph (S,T,F):
  - S: **perfect matchings in** G
  - T: **induced matchings** of size αn in G
  - **Edges:** $(M,M') \in F \Leftrightarrow M' \subset M$
- We have:
  - $|S| \geq (d/e)^n$ *(Schrijver; Bregman)*
  - $|T| \leq \binom{n}{\alpha n}^2 \leq (e/\alpha)^{2\alpha n}$ (counting)
  - $\deg(M') \leq (d!)^{(1-\alpha)n/d}$ for $M' \in T$ *(Alon & Friedland)*
- If $\alpha=9/d^{1/2}$ then average degree of side S is **less than 1**
- **Therefore**: $\exists$ perfect matching M containing no induced matching of size αn.

Deeper Networks

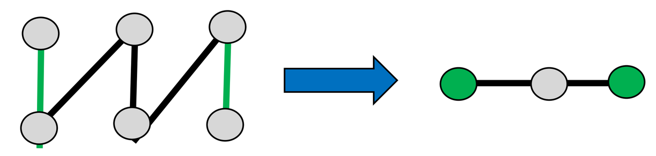
Theorem: Every d-regular r-layered graph satisfies:

$$\alpha \leq r / (d^{1-1/r} \log r)$$

- **Separation of shallow vs. deep networks**

Good Multitaskers

- Technique: Given matching M, **contract** its edges.
  Large **independent sets** within contraction correspond to **induced matchings** before contraction.



- Immediate consequences:
  → **Forests**: $\alpha \geq 1/2$
  → **Planar** graphs: $\alpha \geq 1/4$
- However, these graph families have *constant* average degree
- Question: Under what conditions can we get $\alpha=\Omega(1)$ for arbitrary d?

Locally Sparse Graphs are Good Multitaskers

- **Idea: Restrict task set size.** Require multitasking only up to k tasks.

> G is a **(k,α)-multitasker** if **every matching** M, $|M| \leq k$ contains an **induced matching** $M' \subset M$ of size $|M'| \geq \alpha|M|$.

- **Theorem**: $\exists$ $(k,1/2)$-multitasker for every $k=\Omega(\log_{d-1} n)$
  Proof: d-regular graphs of **high girth**.
- **Theorem**: $\exists$ (k,α)-multitasker for every $k=\Omega(n/d^{1+\Omega(\alpha)})$ and $0<\alpha<1/5$
  Proof: Graphs in which every subset of size up to $s=\Omega(n/d^{1+\Omega(\alpha)})$ spans $O(\alpha^{-1}s)$ edges *(Feige & Wagner)* + **Turan Theorem**.

Future Directions

- Prove/disprove: $\alpha=o(1/d^{1/2})$ for **every** d-regular $(k=99n/100, \alpha)$-multitasker
- **Empirical examination** of α in **parallel architectures**
- **Tight lower bound** for α for **networks of depth > 2**

References

- L. M. Bregman, Soviet Math. Dokl., (1973).
- N. Alon and S. Freidland. Electronic J of Combinatorics, (2008).
- Y. Birk, N. Linial and R. Meshulam, IEEE transactions of information theory (1993).
- J. D. Cohen, K. Dunbar, and J. L. McClelland, Psychological Review (1990).
- S. F. Feng, M. Schwemmer, S. J. Gershman, and J. D. Cohen, Cognitive, Affective, & Behavioral Neuroscience (2014).
- S. Musslick, B. Dey, K.Ozcimder, M. M. A. Patwary, T. L. Willke, and J. D. Cohen, CogSci (2016).
- S. Navalka J. Bar Joseph and A. Barth, Trends in Cognitive Science (2017).
- M. Posner and C. Snyder, In Information processing and cognition: The Loyola symposium, pp. 55–85 (1975).
- A. Schrijver, Journal of Combinatorial Theory, Series B (1998).
- R. M. Shiffrin and W. Schneider, Psychological review (1977).