Current State of Affairs

- Better known CS Dept ratings in US:
  - US News and World Report
  - National Research Council (1995, 2010)

- Uses/abuses of ratings:
  - Efficient way to inform decisions
    - Choosing a PhD program (especially foreign students)
    - Applying for an academic position (PhD graduates)
  - Imposes structure on the field
    - Eg CRA salary comparisons of "like" institutions.
  - Used in discussions between Dept and Univ administration
    - Rewards for ratings improvements
    - Funding for remedial action when ratings fall
    - Reality check on claims

Problems with Rating Schemes

- Trailing indicator
- Imposes a value system
  - Different people have different needs and will flourish in different environments

Nb: *Horror vacui*   -Parmenides 485BC

 = "Nature abhors a vacuum"

Infeasible for our community to decide:

  There should be no rating system.

Ideal Rating Scheme Properties

Accuracy: Ratings correlate with what they purport to represent. E.g.

- Quality of PhD program
- Quality of research
- Quality of undergrad program

Accuracy implies: Hard to "game" the scheme.

Assurance: People have reason to believe the ratings are accurate.

- Transparency of process is crucial.

Temporal continuity with prior rankings.

3

Ranking Methodology: Inputs

- **Inputs:**
  - Objective data
    - Publications, grants, size, graduates, graduation rate, …
  - Subjective data
    - Who are the authorities?
      - Per-dept "expert" (=US News and World Report)
      - Designated "experts" (NRC, UK RAE/REF)
    - What data do the authorities use?
      - Gut feeling (=US News and World Report)
        - Often formed indirectly (grants, newspaper, awards)
      - Objective documentation provided by subjects (=UK RAE/REF)

Ranking Methodology: Outputs

- Outputs:
  - Predefined summary statistics:
    - One or several? E.g.,
      - PhD program quality
      - Research impact
      - Undergraduate programs

    Note: Certain statistics could create or imply a sub-field structure.

  - Allow users to define "formula"?

    "If everybody is special, then nobody is." [*The Incredibles, Pixar Films*]

  - Numeric (w/wo error bars) vs equivalence classes?
    - Non-ordinal helps prevent spurious inferences.

When outsiders rank us …

- *Prima fascia* "objective"

- Lower cost to the field, but there are costs.
  - Depts still must provide data
  - Individuals still must provide opinions (1x or 2x rounds)

- Outsider defines what is important.
  Likely choice: something that works across many disciplines.
  - Dimensions of concern:
    - what are <u>our</u> sub-areas
    - value systems differ from discipline to discipline:
      - publications vs "software artifacts"
      - conference publications vs journal publications
    - ranking formula(s) coeffecients

Questions to Ponder…

Are current ranking schemes bad enough to warrant replacement?

- How can we determine this?
- Can CRA help improve current schemes?
- Can CRA develop a scheme that is good enough to be useful?
  - How might the community agree on what that scheme is?
  - Can the community afford to implement that scheme?