Algorithms, Incentives, and Autonomy

Improving Fairness while Respecting Free Will

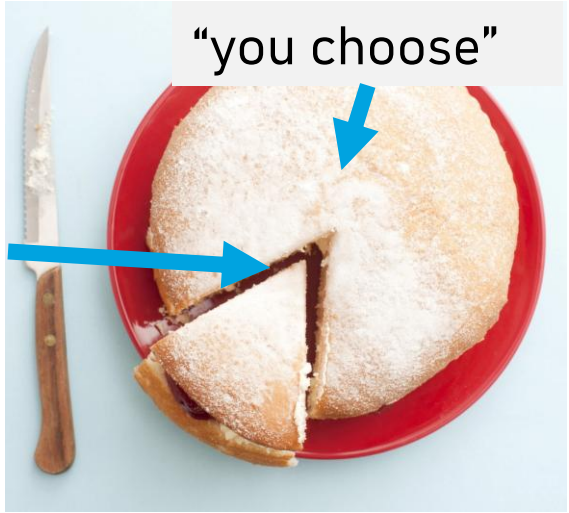
Ian Ludden

Franciscan University of Steubenville

March 6, 2025

How should siblings split the last slice?

- You and your sibling need to share the last slice of cake. But how?
- Simple protocol: I cut, you choose
- Is this fair? Why?
  - Proportional
  - And envy-free
  - … even if players have different valuation functions (e.g., one likes icing more)
- Keeps siblings' autonomy, but ensures fair end result



Genesis Ch. 13, [Steinhaus '48]

Algorithmic Game Theory: Definition

"Game theory is the study of **mathematical models** of **strategic interactions** among **rational agents**."
– Roger Myerson (1991),
Game Theory: Analysis of Conflict

Agents try to maximize their **utility**
- 1 for win, ½ for draw, 0 for loss in Chess

Agents select one **strategy** from an available set
- Fold, call, or raise in poker

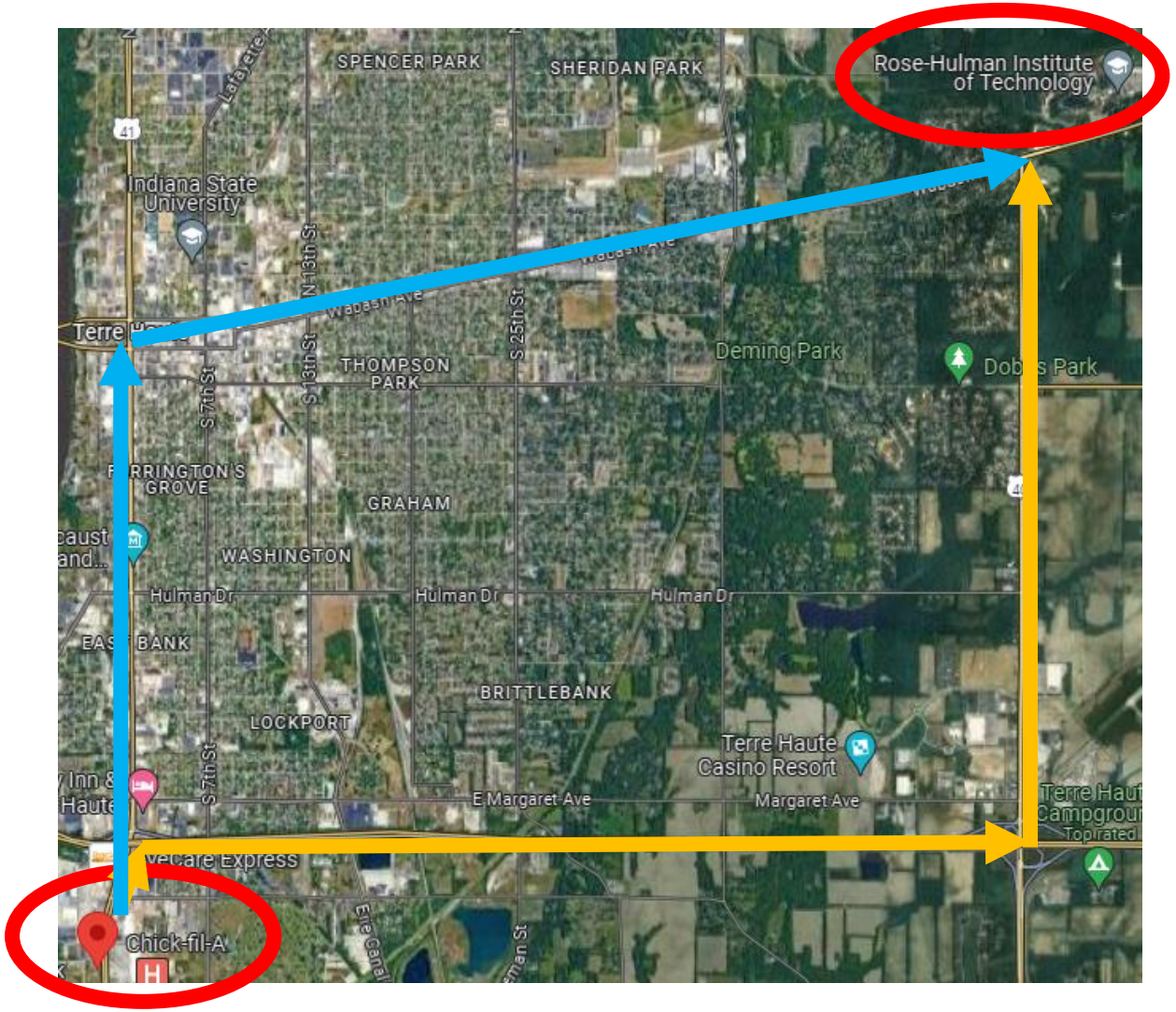
Agents play **optimally, not emotionally**
- Calculated risks only

"An algorithm is a sequence of unambiguous instructions for solving a problem." – Anany Levitin (2011), *Introduction to the Design and Analysis of Algorithms*

"A fascinating fusion of ideas from both fields [...] came into being and was used productively in the effort **to illuminate the mysteries of the Internet**. It has come to be called algorithmic game theory." – Nisan, Tardos, Roughgarden, and Vazirani (2007), *Algorithmic Game Theory*

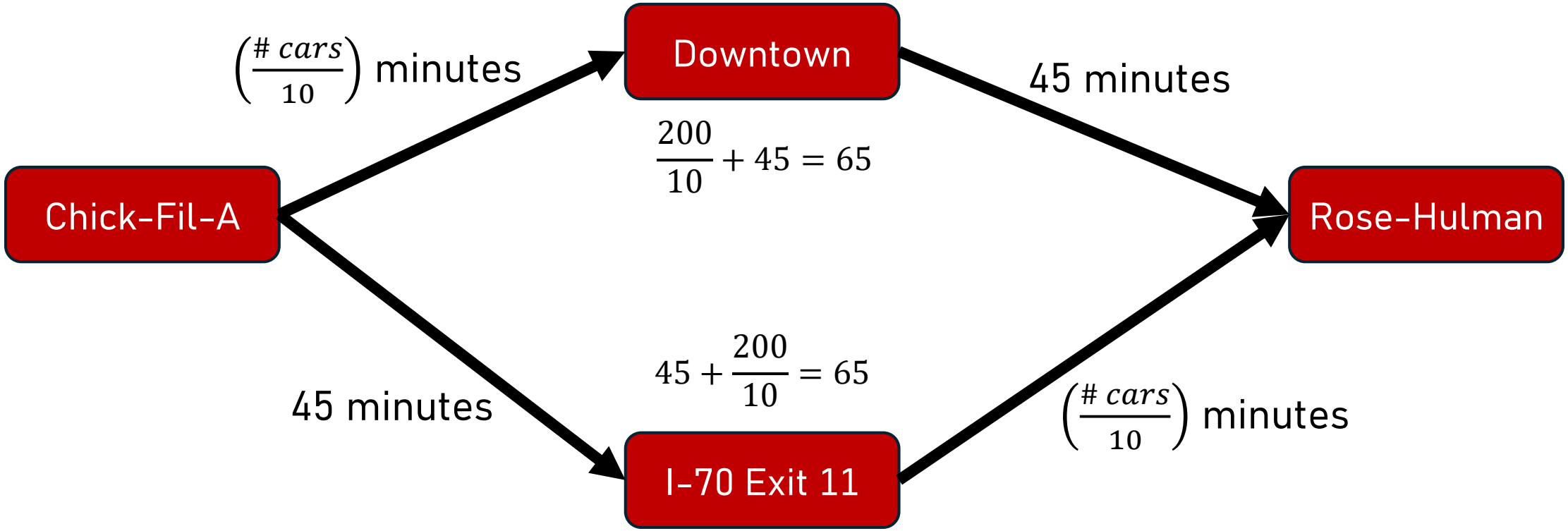
Application: Traffic Networks

- Two roads diverged
- Chick-Fil-A to Rose-Hulman:
  - 400 commuters, all at once
- Two candidate routes:
  - Max speeds: I-70 to US 40
  - Min distance: US 41 to Wabash Ave



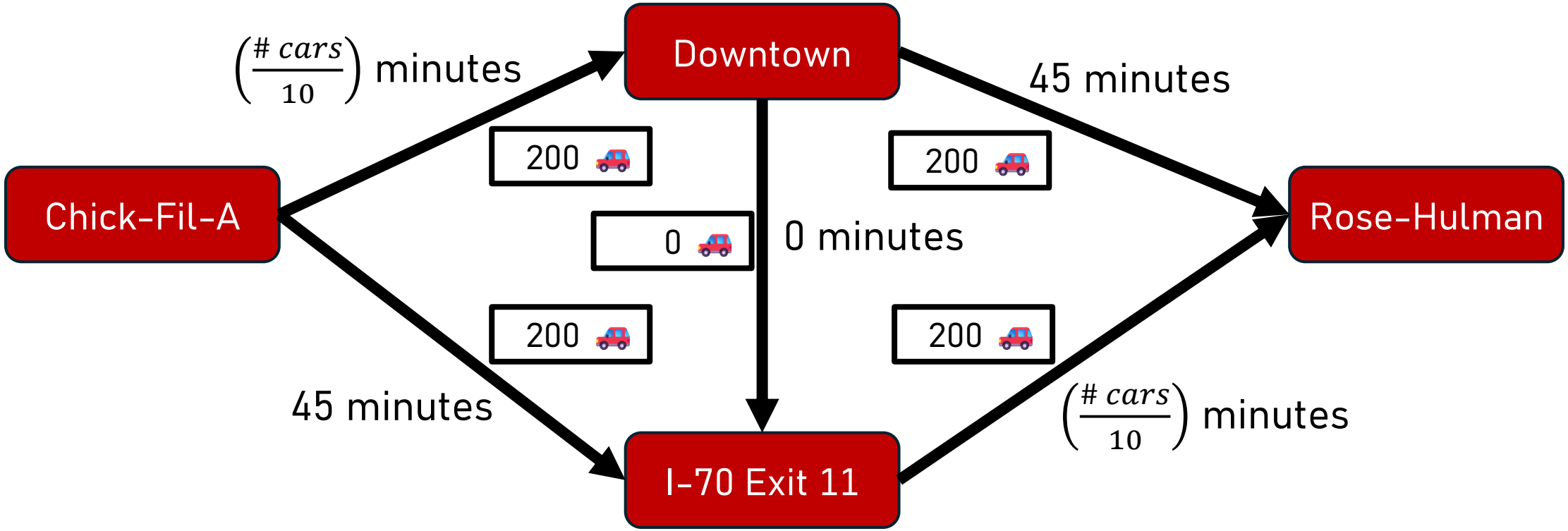
Simplified model of road network

- Assume travel time per road is (a) fixed, or (b) depends on # cars
- Equilibrium: 200 cars take each route, 65 minutes each



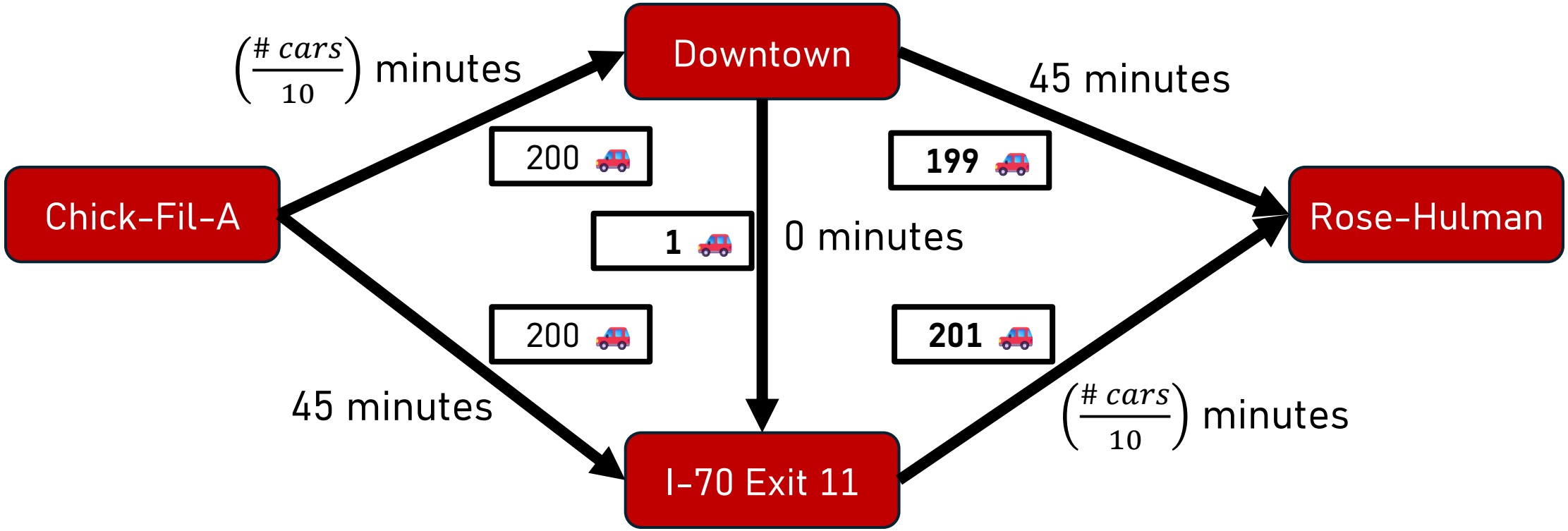
Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- One person takes the wormhole: $200/10 + 201/10 = 40.1 < 65$



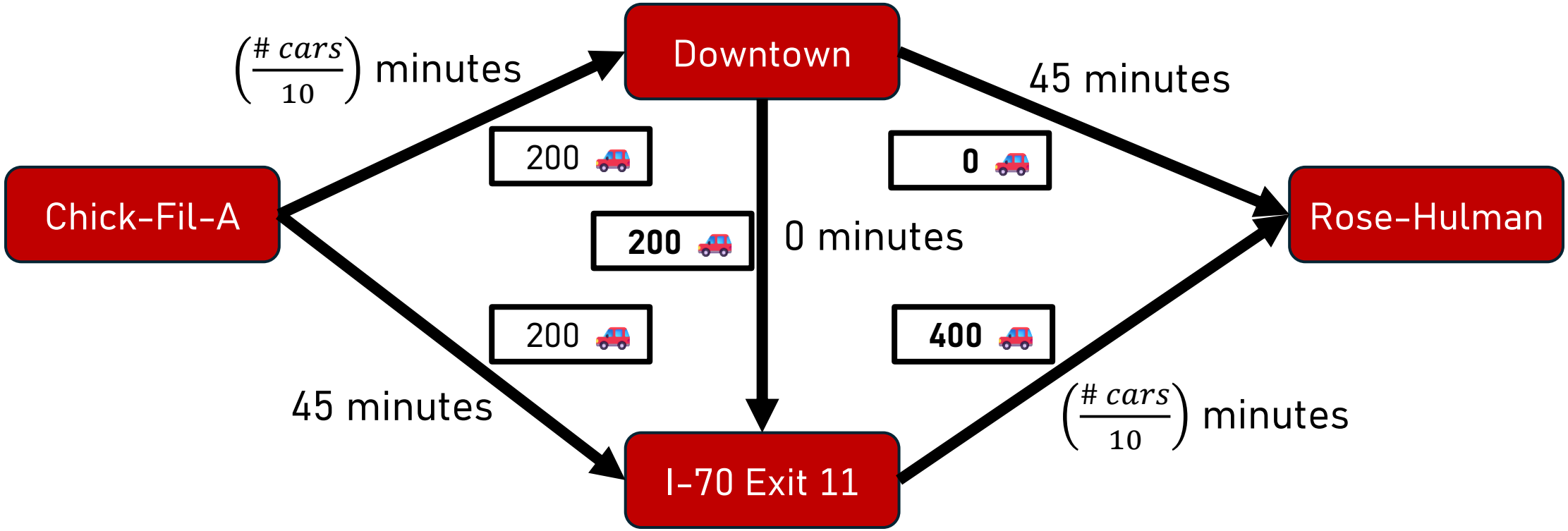
Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- One person takes the wormhole: $200/10 + 201/10 = 40.1 < 65$



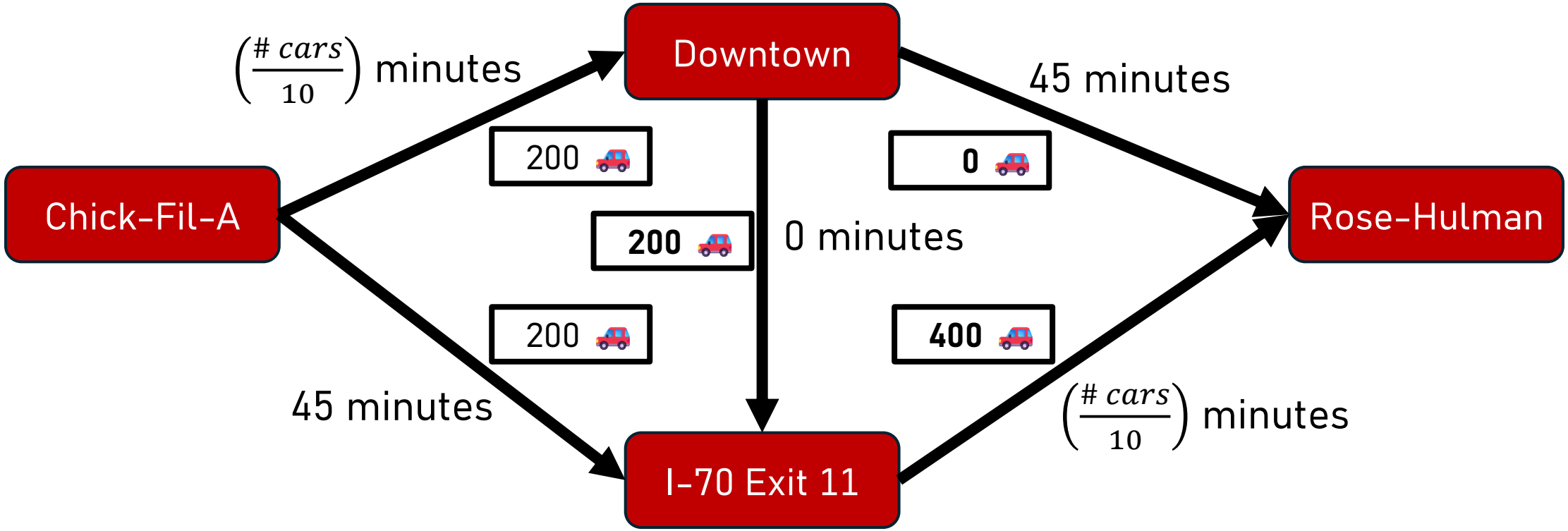
Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- **ALL** downtown cars take the wormhole: $200/10 + 400/10 = 60 < 65$



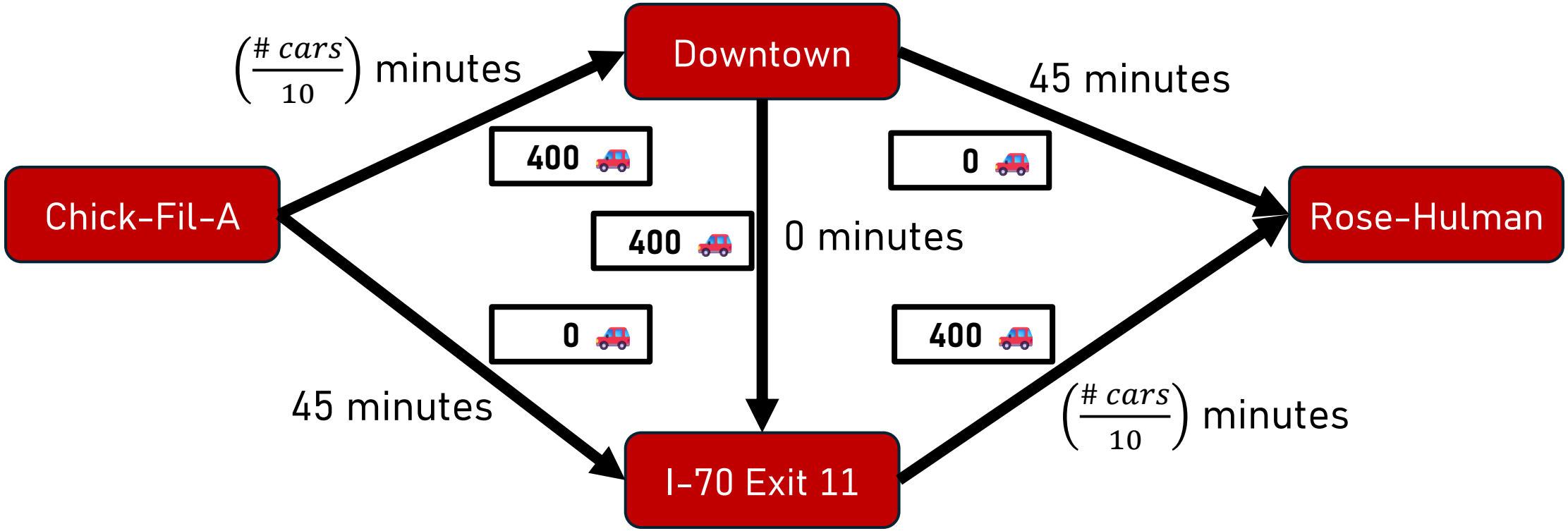
Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- But now, **<u>I-70 cars have incentive to switch</u>**: $201/10 + 400/10 \approx 60 < 85$



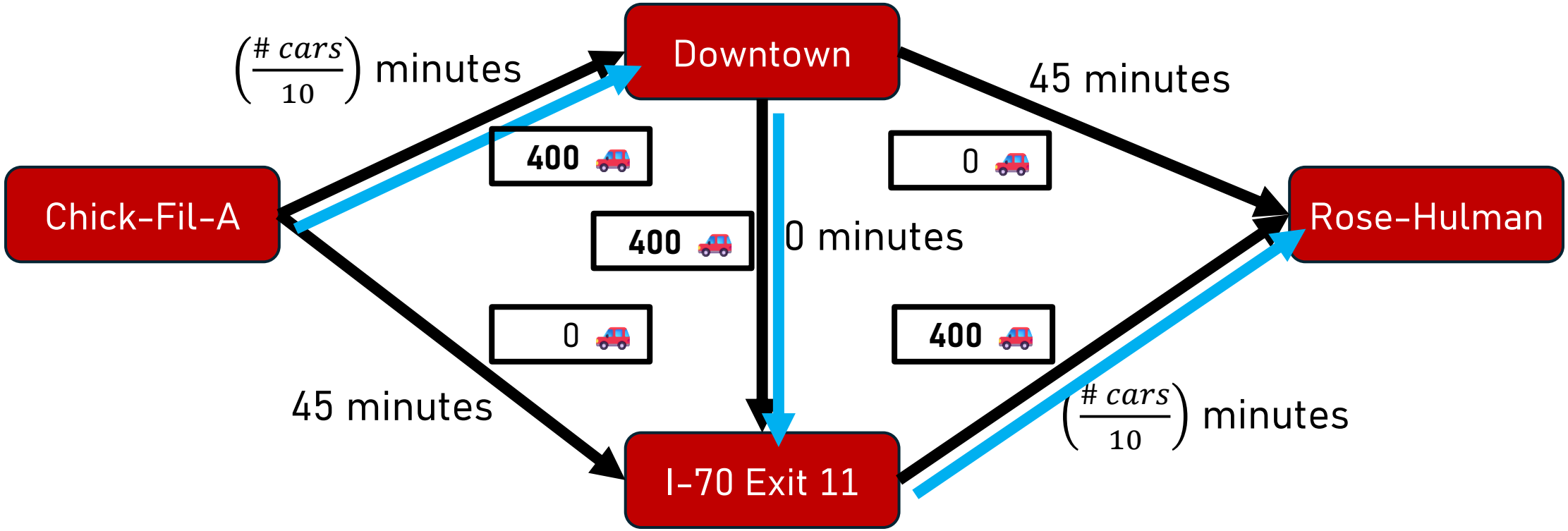
Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- But now, **<u>I-70 cars have incentive to switch</u>**: $201/10 + 400/10 \approx 60 < 85$



Now, there's a wormhole

- Instantaneous travel from Downtown to Exit 11
- **Equilibrium**: *everyone* goes downtown → wormhole → Exit 11, 80 min.



Adding wormhole slowed everyone down!

- Example of ***Braess's Paradox*** [Pigou 1920, Braess 1968]
  - Adding a road may slow down overall traffic flow
- Easy fix for traffic jam: remove the wormhole (close a road)
- Not just hypothetical!

Seoul, South Korea

"the destruction of a six-lane highway to build a public park actually improved travel time into and out of the city"
[Easley & Kleinberg, 2010]

Stuttgart, Germany

New roads added, but traffic didn't improve until a new section ***closed***
[Knödel, 1969]

- Springs and strings: physical demo of Braess's paradox

Understanding Braess's Paradox

- Adding the wormhole/fast road changed the ***Nash equilibrium*** (status quo where no one wants to change strategies) for the worse.

- ***Price of Anarchy (PoA)***: how bad is the Nash equilibrium relative to a global/system optimal solution?
  - Our example: 80 minutes for Nash eq., 65 minutes for global opt., so PoA = 80/65 ≈ 1.23
  - Alternate name: Price of Autonomy

- Takeaway: to respect free will but still have efficient systems, we need designs that ***align individual & group incentives***
  - … as opposed to, say, the government routing your car for you

> Man tends towards good, but he is also capable of evil. […] The social order will be all the more **stable**, the more it takes this fact into account and **does not place in opposition** personal interest and the interests of society as a whole, but rather **seeks ways to bring them into fruitful harmony**. […] In fact, where self-interest is violently suppressed, it is replaced by a **burdensome system of bureaucratic control** which dries up the wellsprings of initiative and creativity.

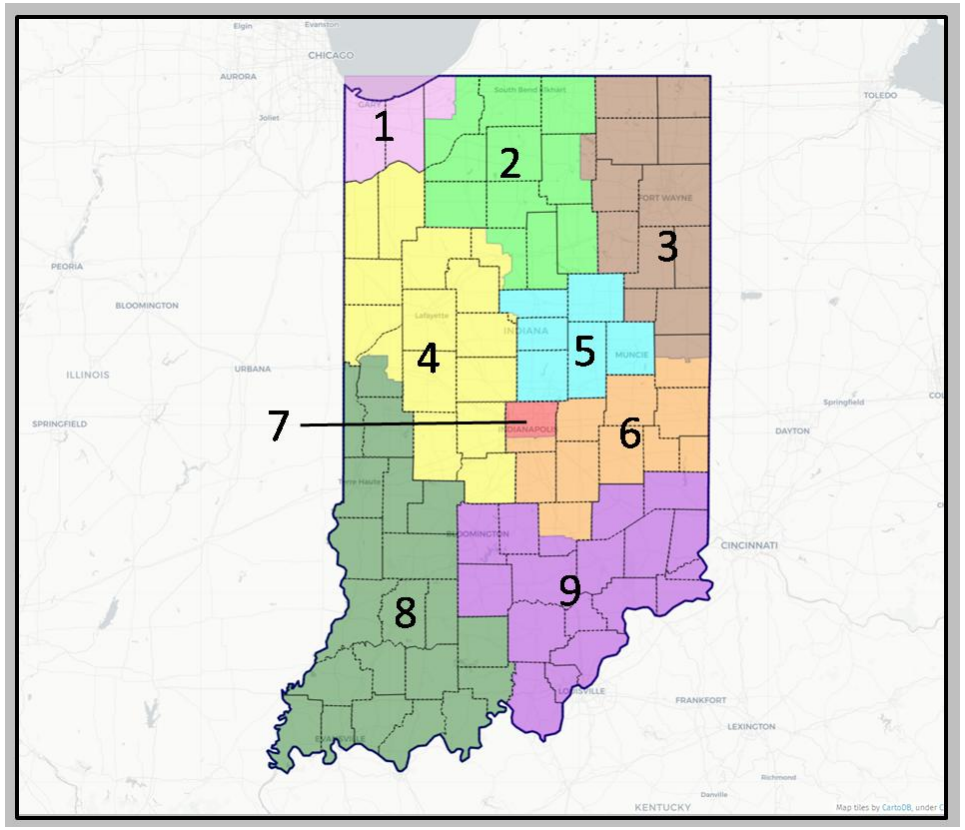
— Pope St. John Paul II, *Centesimus Annus* (May 1991), 25 (emph. added)

Context: Solidarity trade union/social movement in Poland (1980s), fall of the Berlin Wall (Nov. 1989), dissolution of USSR (Dec 1991)

My Research: AGT for Political Redistricting

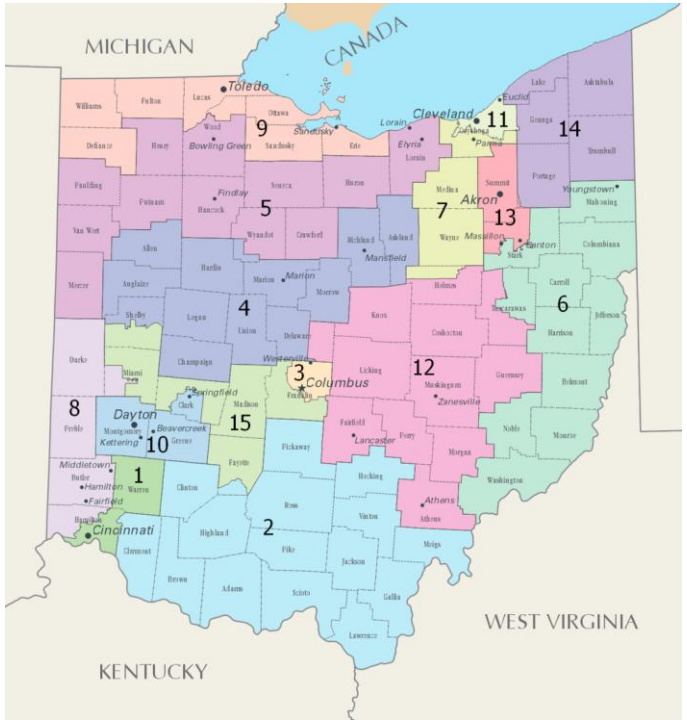
Political redistricting: the process of updating voting district boundaries after each census to rebalance populations

Indiana Congressional Districts



By CX Zoom, CC BY-SA 4.0

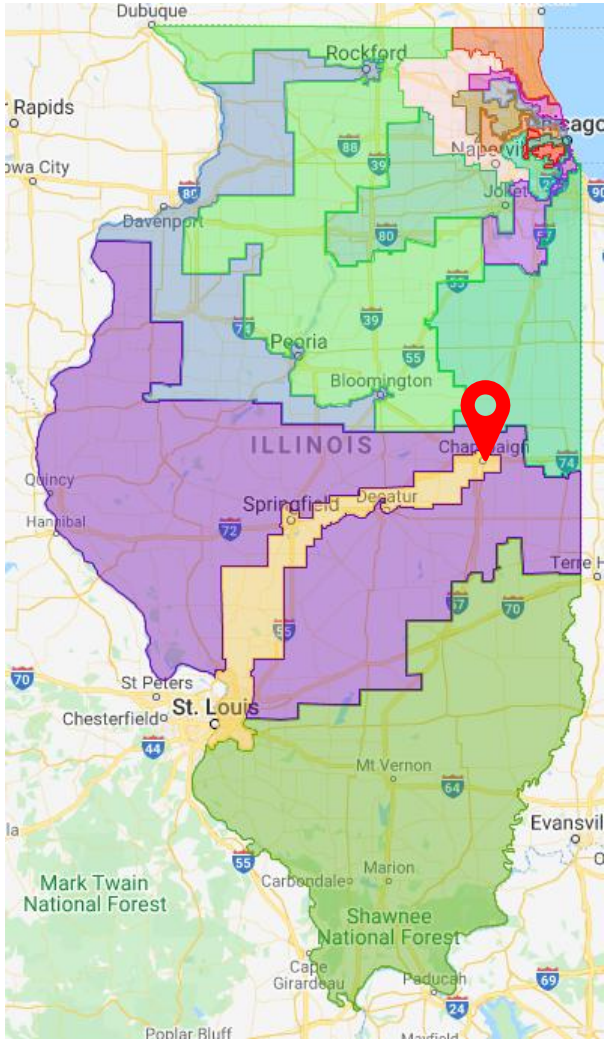
Ohio Congressional Districts

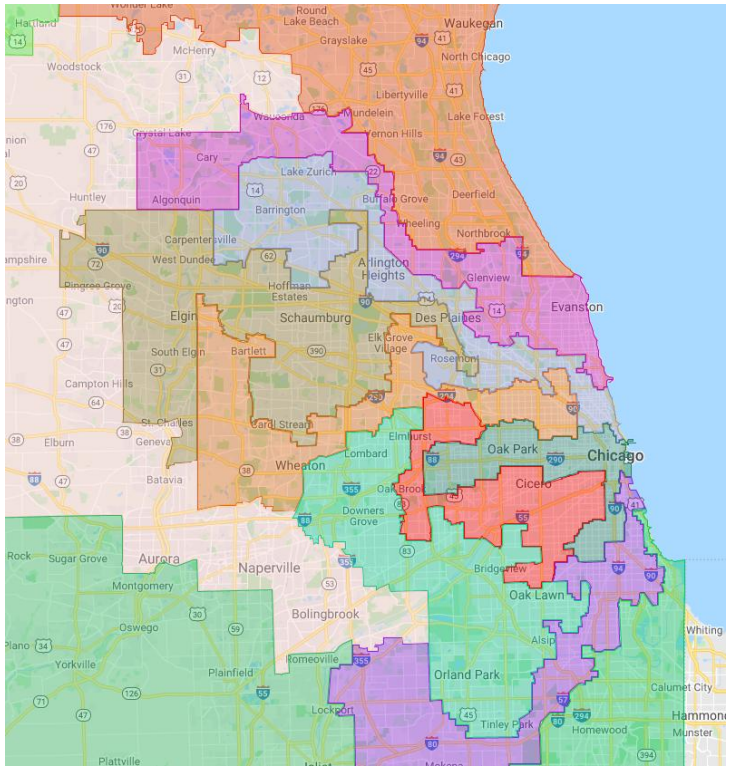


By Nebraskan fellow, CC BY-SA 4.0

The Problem: Gerrymandering

"The Gerrymander: a New Species of Monster" *Boston Gazette*, March 26, 1812, page 2, Library of Congress.

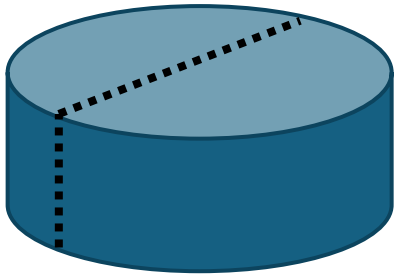




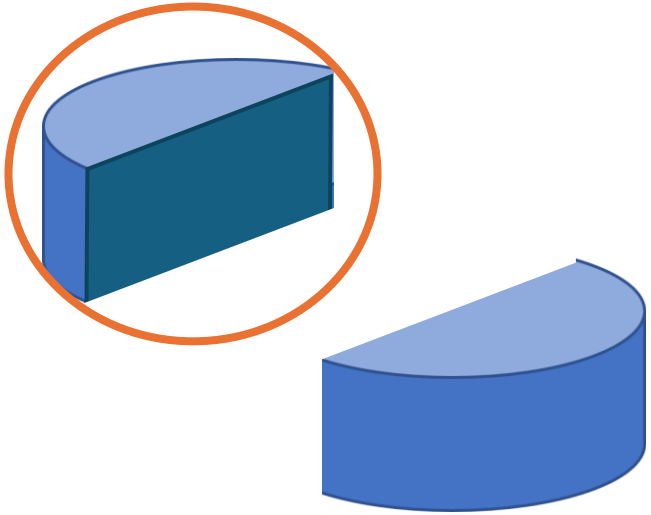
New Illinois Congressional Districts
https://bit.ly/3rODoWl

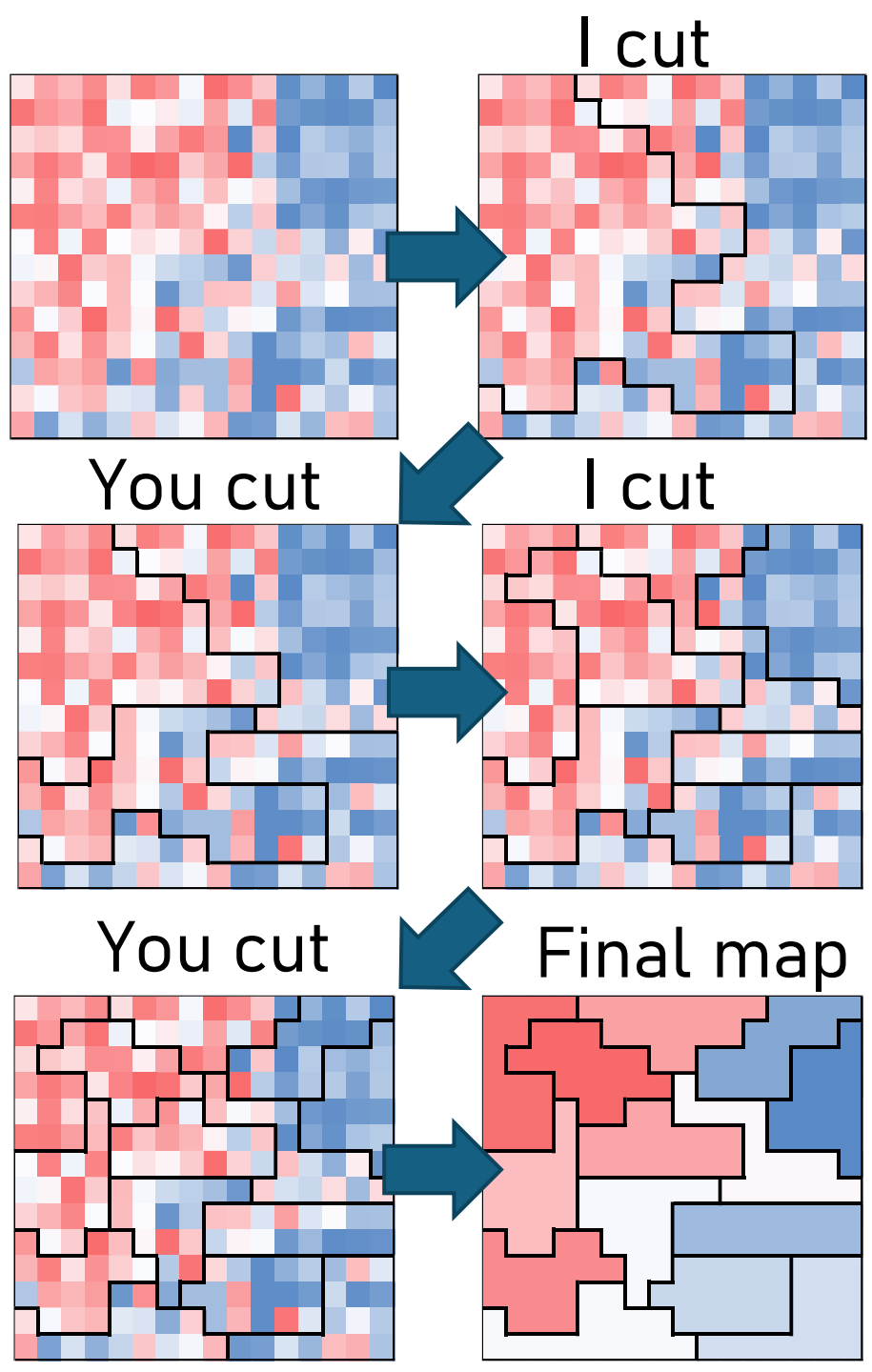
One Possible Solution: The Bisection Protocol

I cut

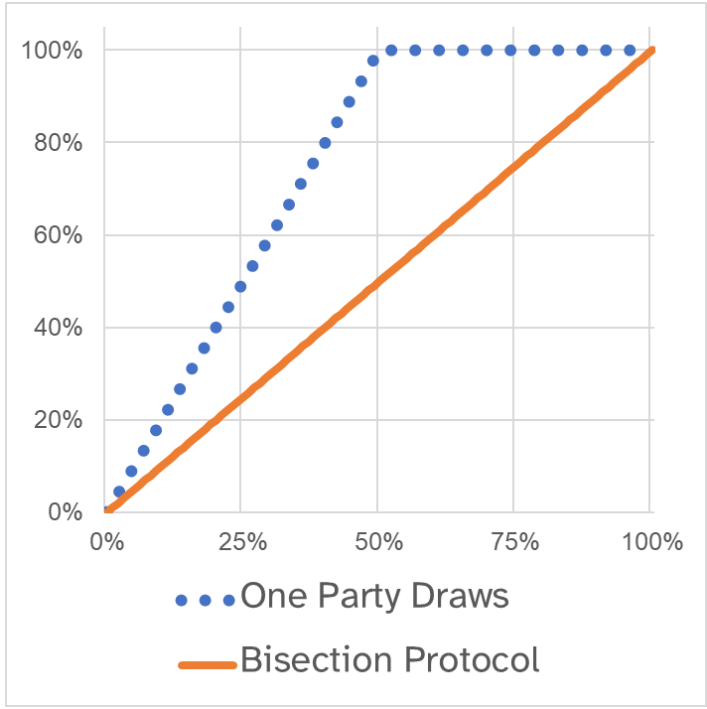


You choose





More proportional than status quo

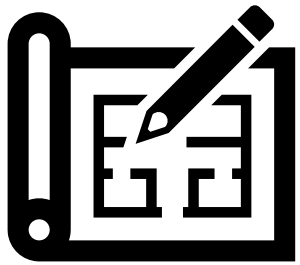


x = vote-share (%)

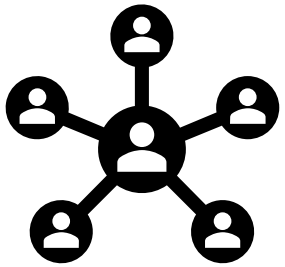
y = seat-share (%)

Why should you care?

- As a system designer/analyst:
  - Keep strategic interactions in mind, or risk inefficiencies & inequities



- As a system participant/user:
  - Resist pressures to act in a self-interested way that harms others



Thank you for coming!

Contact info

Ian Ludden

ian.ludden@rhit.edu

Department of Computer Science and Software Engineering

Rose-Hulman Institute of Technology, Terre Haute, IN

Slides and more

https://ianludden.com/presentations