Urban dashboardistics

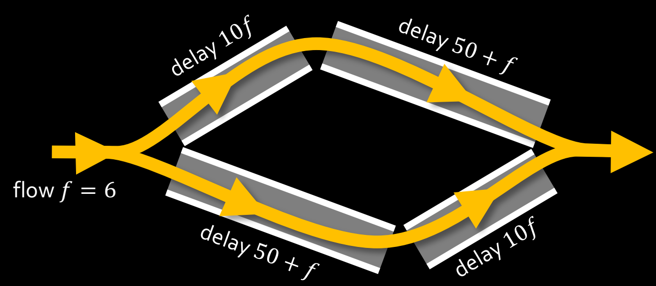
Damon Wischik
Dept. of Computer Science and Technology

UNIVERSITY OF CAMBRIDGE

Wardrop modelled route choice by drivers in a traffic network.

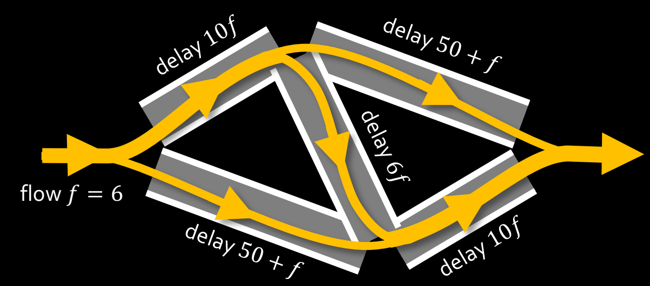
Braess discovered paradoxical outcomes
(which can be overcome by setting appropriate tolls).

DRIVERS	CITY
▪ can choose which route to take ▪ want the fastest travel time	▪ provides infrastructure ▪ wants to maximize total utility over all drivers



travel time 83

city builds a new road



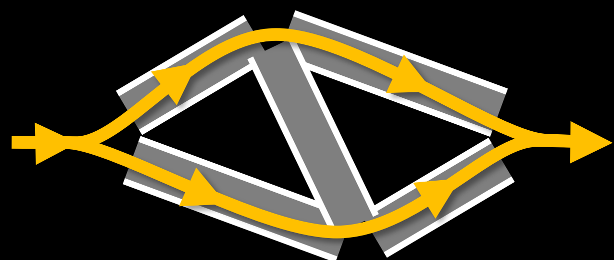
travel time 92

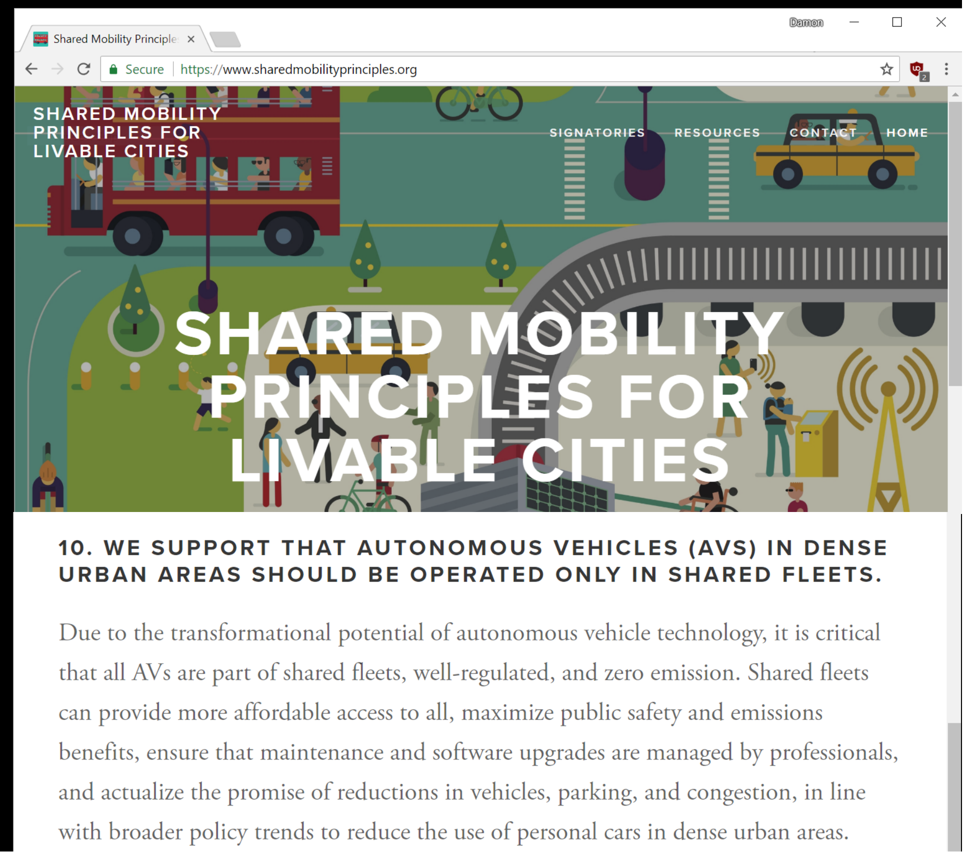
Towards a multimodal Wardrop equilibrium model:
Let's consider three main agents, each with their own objectives.

USERS	RIDESHARE FLEET	CITY
▪ can choose mode of transport ▪ want low price, low travel time, etc.	▪ can choose its pricing model ▪ needs to rebalance its fleet ▪ wants profit	▪ wants to maximize total utility over all users ▪ provides public transport, and sets prices
▪ prompt, cheap deliveries	DELIVERY FLEETS ▪ and assorted other fleets, in competition	▪ wants universal service ▪ responsible for environmental externalities etc.
many uncoordinated individuals	a few large operators, with direct user contact and real-time backends	REGULATORY POWER

How should cities relate to fleets?

1. A fleet operator, if it's a dominant player, can avoid Braess's paradox by internalizing the cost of congestion.





How should cities relate to fleets?

2. Fleets have the effect of *coupling* parts of the transport network in new ways. This might lead to paradoxical behaviour.

USERS

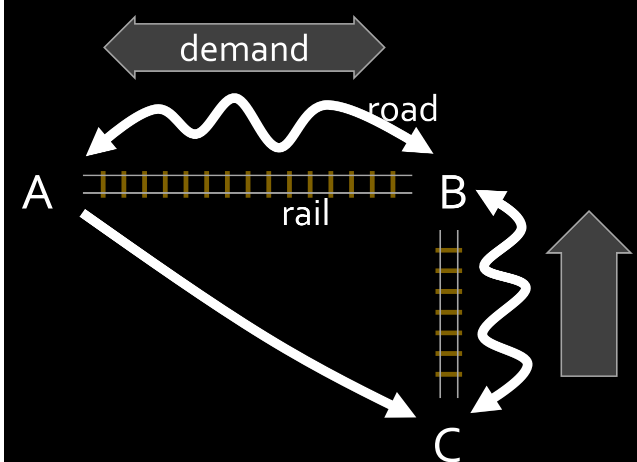
- can choose mode of transport
- wants lowest price

RIDESHARE FLEET

- can set origin-based 'surge' multipliers
- needs to rebalance the fleet
- has to balance cost and revenue

CITY

- sets public transit fares



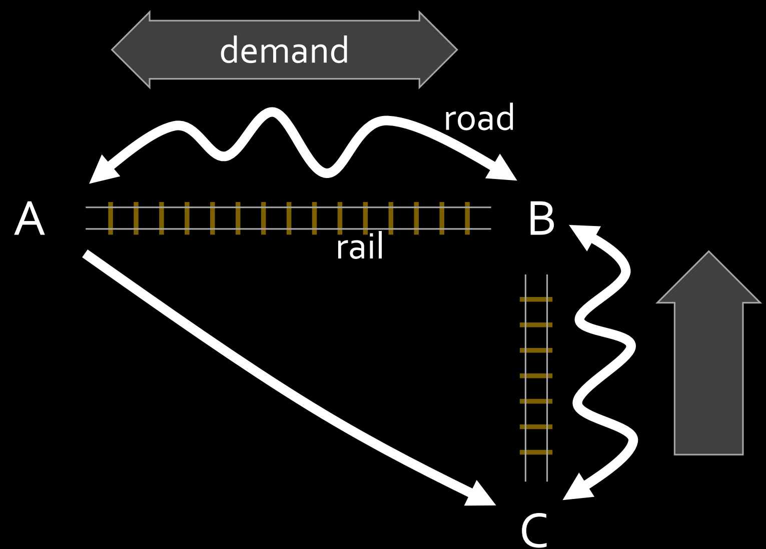
Let c_{ij} be the base cost of a road trip from i to j, and f_{ij} the rail fare
Let μ_i be the surge multiplier at location i, so passenger pays $\mu_i c_{ij}$
Let $\alpha_{ij} = \alpha_{ij}(\mu_i)$ be the demand for rideshare trips
[measured in trips per unit time, and depending on c and f]

The fleet picks rebalancing rates β_{ij} such that for all i, $\sum_j(\alpha_{ij} + \beta_{ij}) = 0$ and adjusts μ to achieve budget balance:

$$\sum_{i,j}(\alpha_{ij} + \beta_{ij})c_{ij} = \sum_{i,j} \alpha_{ij}\, \mu_i\, c_{ij}$$

How should cities relate to fleets?

2. Fleets have the effect of *coupling* parts of the transport network in new ways. This might lead to paradoxical behaviour.



If the city reduces the rail fare A→B

- some A→B passengers will shift to rail
- the fleet operator still wants to get cars to B, maybe via C
- it might as well discount rides C→B
- this takes C→B passengers off rail
- there might even be a net decrease in rail trips [Sid's paradox]

How should cities relate to fleets?

3. Cooperate with fleets, so that we don’t use streets as parking lots.

How should cities relate to fleets?

3. Cooperate with fleets, so that we don’t use streets as parking lots.

- The city simulates a *virtual road network* whose capacity is 95% of what’s really there, and measures congestion
- Fleet operators agree to set routes and prices according to virtual congestion
- The city sends real-time virtual congestion signals, and the fleets send enough data that the city can verify compliance
- The streets are kept free-flowing
- In return, the fleets are permitted special access to restricted zones

Can't this all be done with congestion charges?

Singapore's
Electronic Road Pricing

DIGRESSION

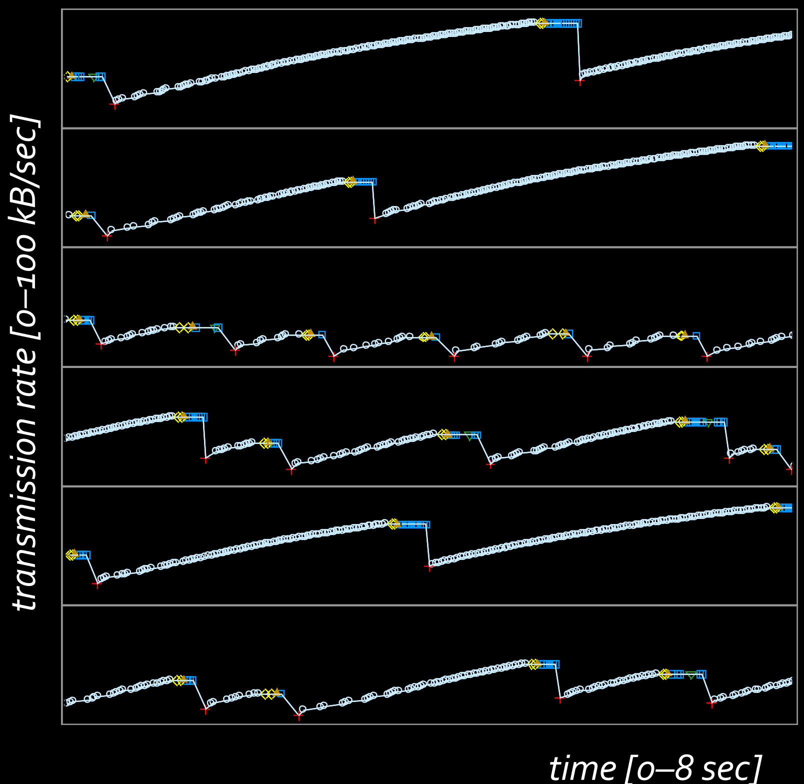
The full story of Internet congestion control theory

Lawrence Berkeley National Lab

"In October of '86, the Internet had the first of what became a series of 'congestion collapses'. During this period, the data throughput from LBL to UC Berkeley (sites separated by 400 yards and two IMP hops) dropped from 32 Kbps to 40 bps. We were fascinated by this sudden factor-of-thousand drop in bandwidth and embarked on an investigation of why things had gotten so bad. In particular, we wondered if the 4.3BSD (Berkeley UNIX) TCP was mis-behaving or if it could be tuned to work better under abysmal network conditions."

Van Jacobson, "Congestion avoidance and control", 1988

Jacobson's TCP algorithm




```
if (seqno > _last_acked) {
   if (!_in_fast_recovery) {
      _last_acked = seqno;
      _dupacks = 0;
      inflate_window();
      send_packets(now);
      _last_sent_time = now;
      return;
      }
   if (seqno < _recover) {
      uint32_t new_data = seqno - _last_acked;
      _last_acked = seqno;
      if (new_data < _cwnd) _cwnd -= new_data;
          else _cwnd=0;
      _cwnd += _mss;
      retransmit_packet(now);
      send_packets(now);
      return;
      }
   uint32_t flightsize = _highest_sent - seqno;
   _cwnd = min(_ssthresh, flightsize + _mss);
   _last_acked = seqno;
   _dupacks = 0;
   _in_fast_recovery = false;
   send_packets(now);
   return;
   }
if (_in_fast_recovery) {
   _cwnd += _mss;
   send_packets(now);
   return;
   }
_dupacks++;
if (_dupacks!=3) {
   send_packets(now);
   return;
   }
_ssthresh = max(_cwnd/2, (uint32_t)(2 * _mss));
retransmit_packet(now);
_cwnd = _ssthresh + 3 * _mss;
_in_fast_recovery = true;
_recover = _highest_sent;
}
```

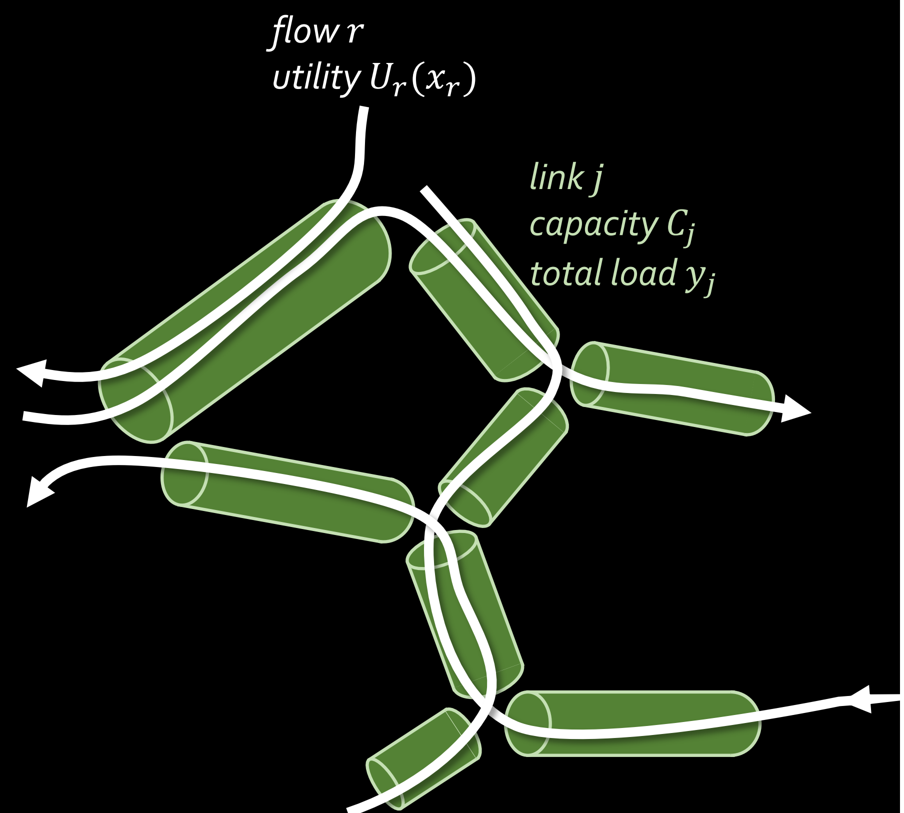
1974	Packet-switched Internet invented (Kahn, Cerf)
1986	congestion collapse
1988	TCP congestion control designed (Jacobson)
1998	mathematical model (Kelly)
2002	control theory solved (Vinnicombe)
2011	standardization of TCP-friendly multipath (Handley)
2013	Apple's Siri is using multipath TCP

Kelly's model for Internet congestion control

x_r = transmission rate of user / route r
$U_r(x_r)$ = utility associated with user r
y_j = total rate on link $j = \sum_r A_{jr} x_r$
C_j = capacity of link j

Consider the optimization problem

maximize	$\sum_r U_r(x_r)$
over	$x_r \geq 0,\ y_j \geq 0$
such that	$y = Ax$ and $y \leq C$



Kelly's model for Internet congestion control

x_r = transmission rate of user / route r
$U_r(x_r)$ = utility associated with user r
y_j = total rate on link $j = \sum_r A_{jr} x_r$
C_j = capacity of link j

Consider the optimization problem:

maximize	$\sum_r U_r(x_r)$
over	$x_r \geq 0,\ y_j \geq 0$
such that	$y = Ax$ and $y \leq C$

A dynamical system for a relaxed version:

$$\frac{dx_r(t)}{dt} = \frac{1}{\mathrm{RTT}_r^2} - p_r(t)\,\frac{x_r(t)^2}{2}$$

$$p_r(t) = \sum_j A_{jr}\left(1 - \frac{C_j}{y_j(t)}\right)^+$$

embodies the utility function implicit in Jacobson's TCP, including the round trip time for route r, RTT_r

a cost / dual variable incurred by the flow on each congested link it uses, corresponding to packet drop rate

Kelly's model for Internet congestion control

	A	B	C	D
1	C	3		
2	RTT	0.1		
3				
4	**t**	**x**	**p**	**dx/dt**
5	0	0.1	=MAX(1-B$1/B5,0)	=1/B$2^2-C5*B5^2/2
6	0.01	=B5+D5*(A6-A5)		
7				

embodies the utility function implicit in Jacobson's TCP, including the round trip time for route r, RTT_r

A dynamical system for a relaxed version:

$$\frac{dx_r(t)}{dt} = \frac{1}{\mathrm{RTT}_r^2} - p_r(t)\,\frac{x_r(t)^2}{2}$$

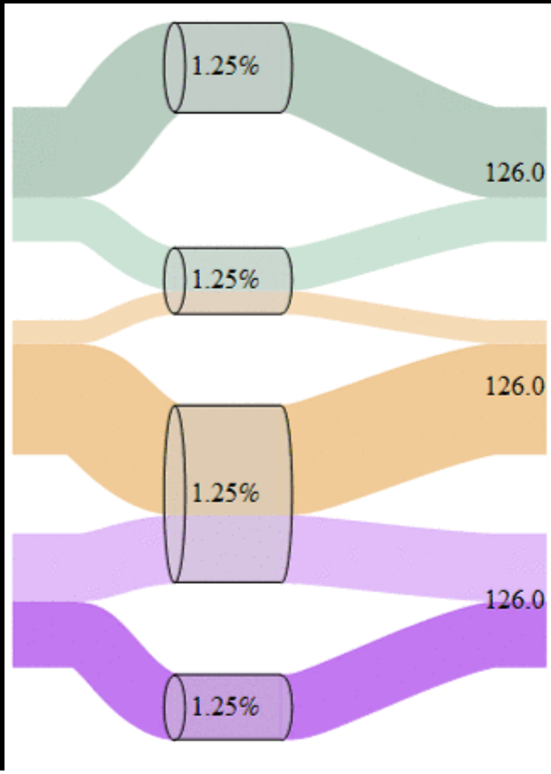
$$p_r(t) = \sum_j A_{jr}\left(1 - \frac{C_j}{y_j(t)}\right)^+$$

a cost / dual variable incurred by the flow on each congested link it uses, corresponding to packet drop rate

The congestion control algorithm extends naturally to multipath settings.

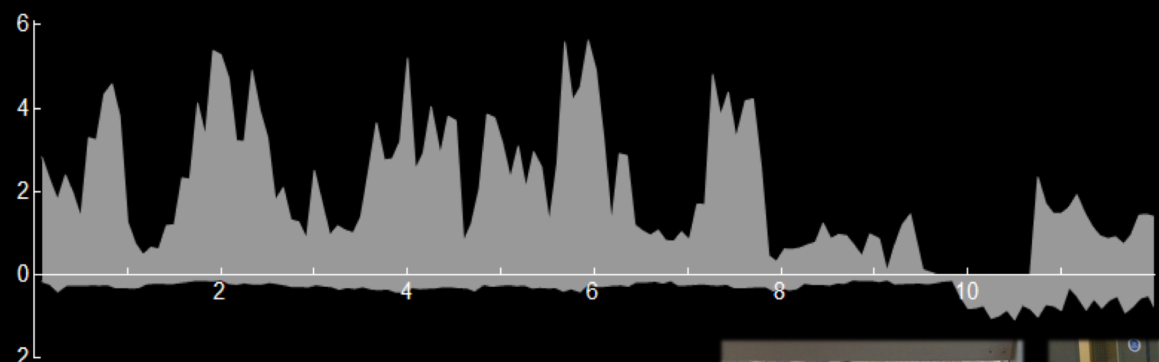
Each end system balances its traffic on the paths it has available, and the overall result is complete resource pooling (when there is sufficient diversity of paths).

Three flows share four resources, as though the network were made up of a single resource.



Mark Handley et al. implemented MPTCP, and ensured it played nicely with middleboxes in the Internet.

```
Internet Engineering Task Force (IETF)                          C. Raiciu
Request for Comments: 6356           Univ. Politehnica of Bucharest
Category: Experimental                         M. Handley, D. Wischik
ISSN: 2070-1721                                 Univ. College London
                                                         Octover 2011
```



User in his office, using wifi and 3G

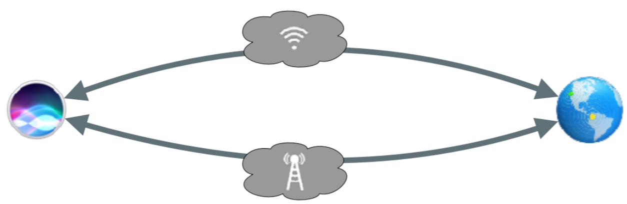
Going downstairs

In the kitchen

Christoph Paasch, Apple
IETF meeting in Prague 2017

MPTCP at Apple

- Implemented since iOS 7 for Siri



User-feedback (Time-to-First-Word) 20% faster in the 95th percentile

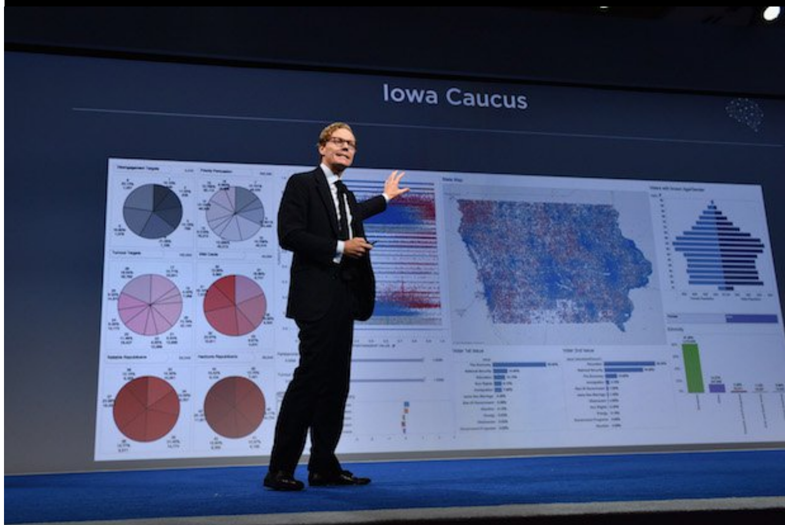
5x reduction of network failures

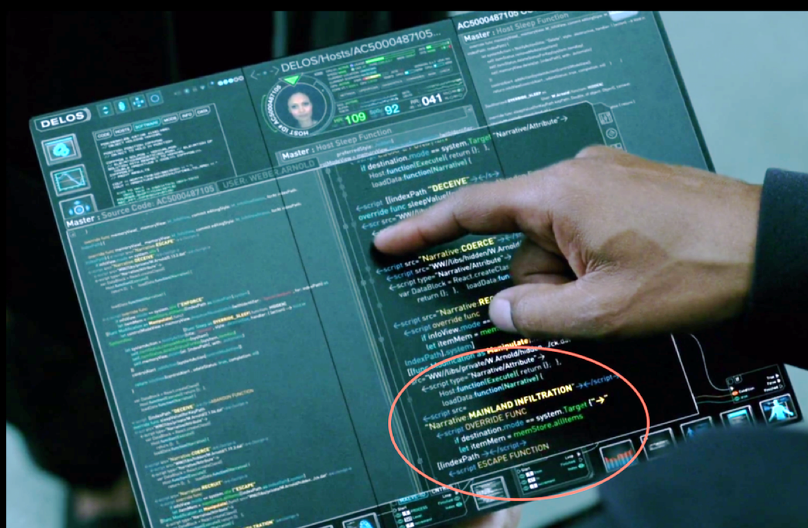
— but Apple discarded the multipath congestion controller, and simply sends each packet on both interfaces!

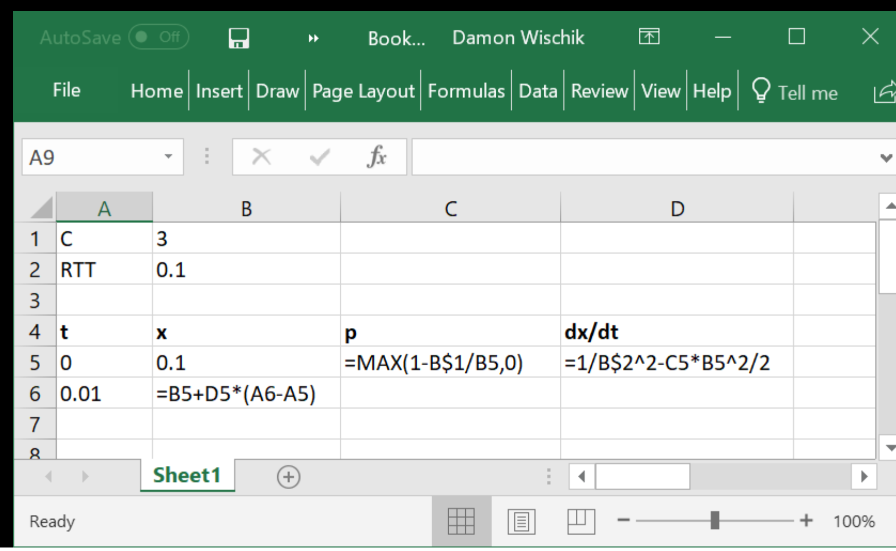
All models are wrong,
but some are useful.

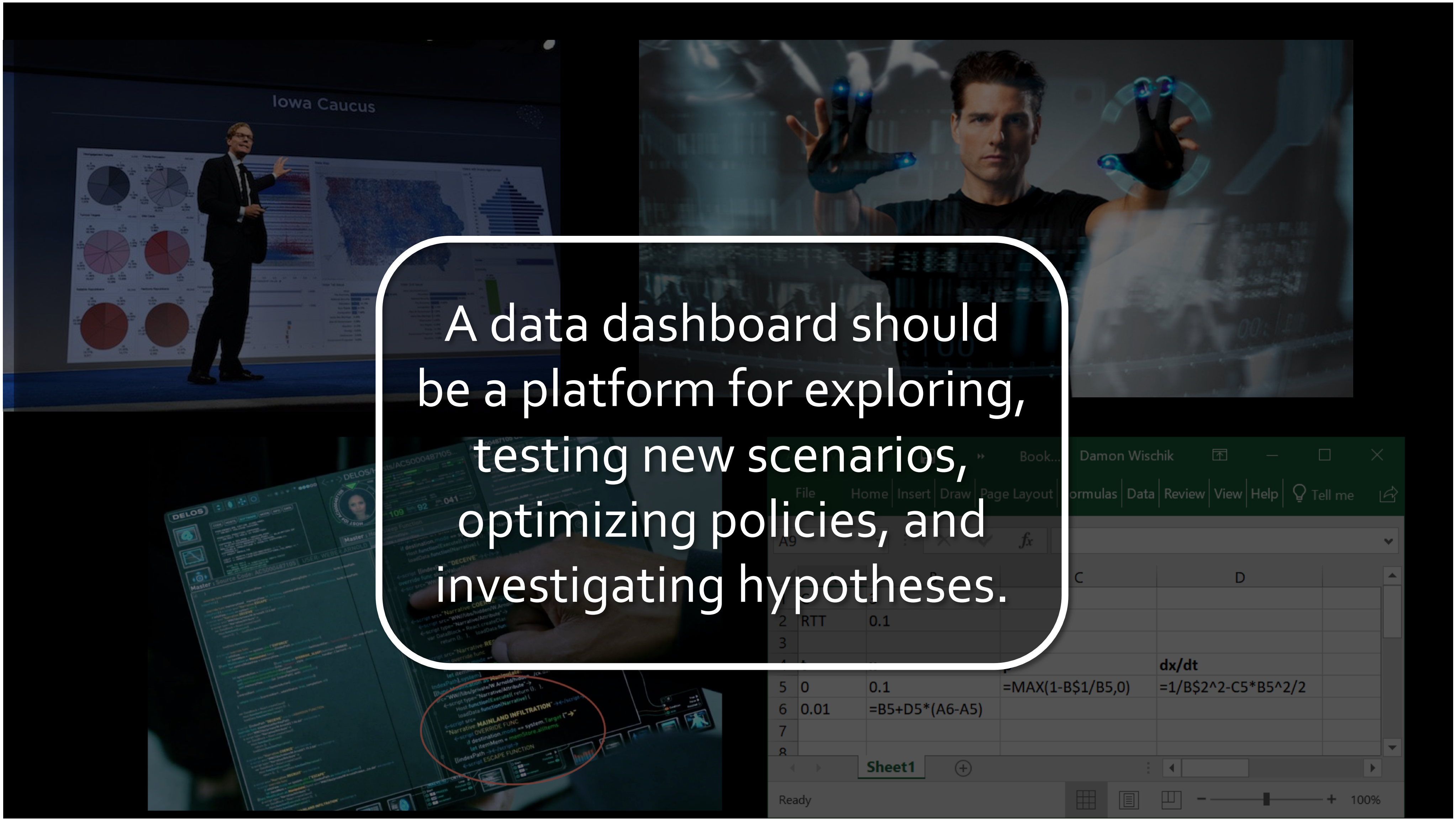
George Box, 1978

What makes a model useful?


Iowa Caucus


DELOS


AutoSave Off
Book... Damon Wischik
File Home Insert Draw Page Layout Formulas Data Review View Help Tell me
A9
A B C D
1 C 3
2 RTT 0.1
3
4 t x p dx/dt
5 0 0.1 =MAX(1-B$1/B5,0) =1/B$2^2-C5*B5^2/2
6 0.01 =B5+D5*(A6-A5)
7
Sheet1
Ready
100%

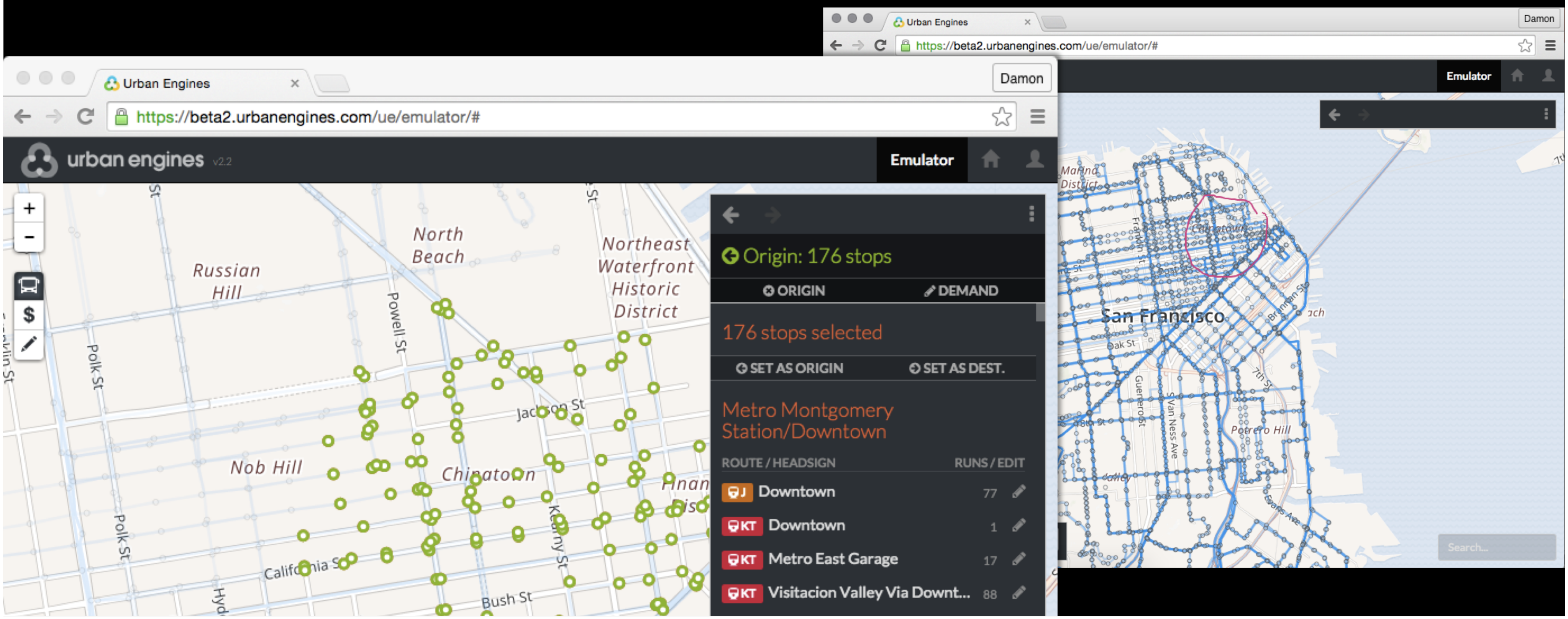


A data dashboard should be a platform for exploring, testing new scenarios, optimizing policies, and investigating hypotheses.

Good design comes from understanding robustly the space of control problems: their data signals, and their data processing operations

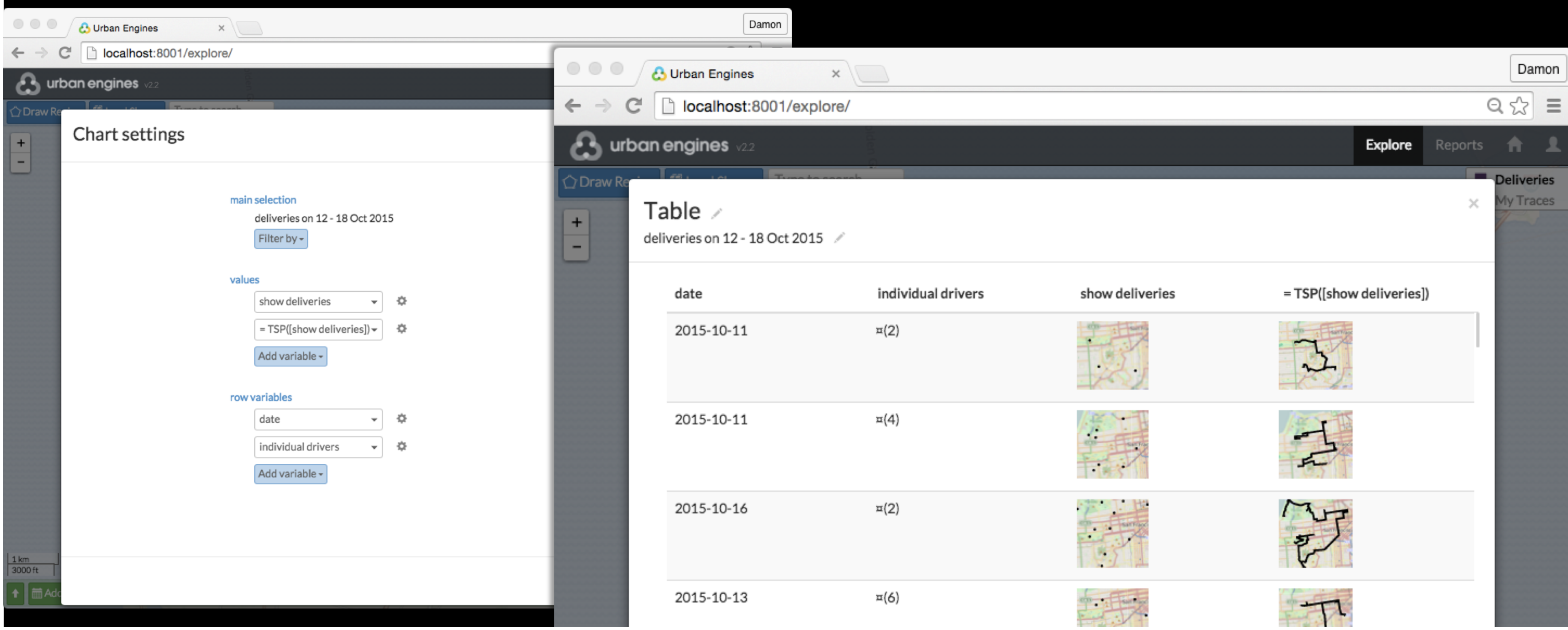
City dashboards we've prototyped:

- GUI interface for modifying trace data, to test out new scenarios
- Novelty: see simulation output in the same environment as your original datasets
- Downside: too clunky to explore the scenario space



City dashboards we've prototyped:

- Like Excel pivot tables, but including set-valued columns / factorized database queries
- Novelty: break down a simulator into functional expressions on data, e.g. =TSP(B3)
- Downside: a confusing way to express simulations



The Alan Turing Institute in London is hiring!
We want a postdoc for two years, to join 3 faculty, 1 PhD, 1 software engineer

- explore user+fleet+city interaction
  → dashboard design → regulatory policy discussion
- study traffic signal control, using reinforcement learning
  → integrate ML with data dashboards