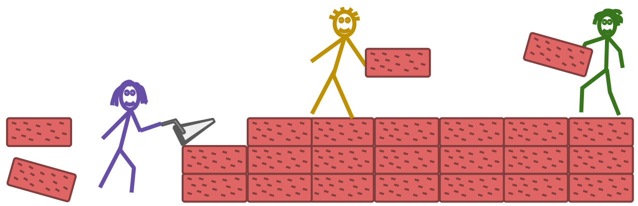
Scoring Rule Markets as Machine Learning Contests



Rafael Frongillo — Colorado, Boulder

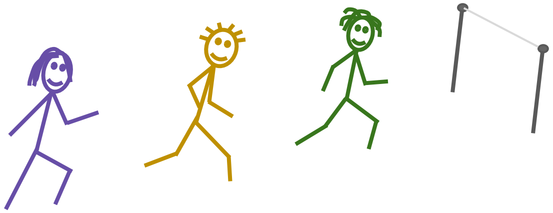
Bo Waggoner — **University of Pennsylvania**

and in part

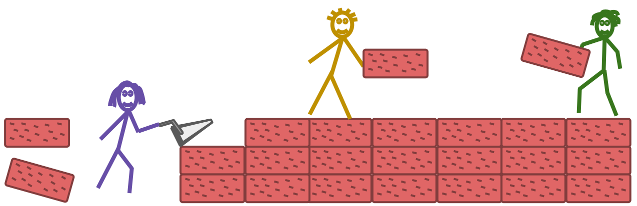
Jacob Abernethy — Georgia Tech

CiML

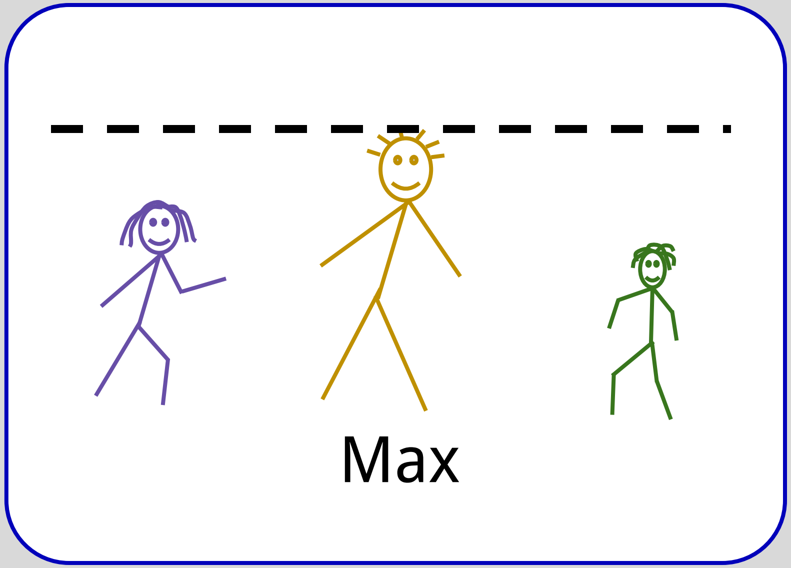
December 2017

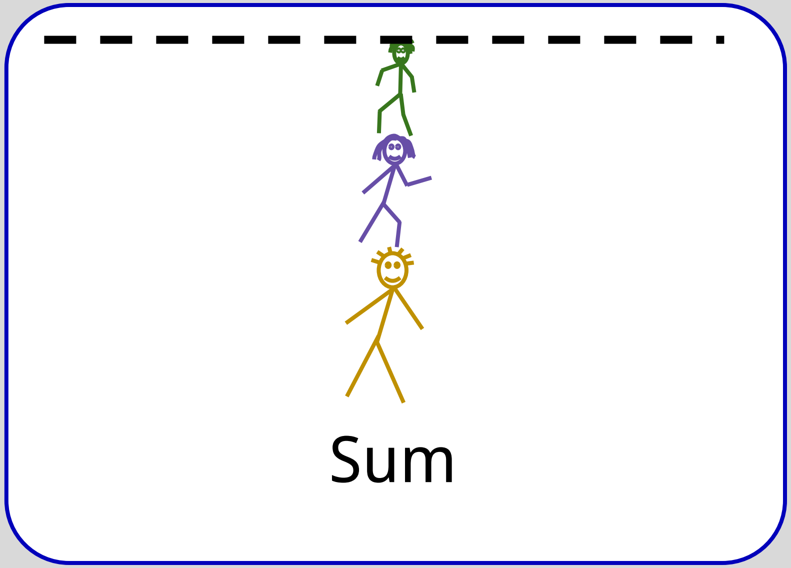


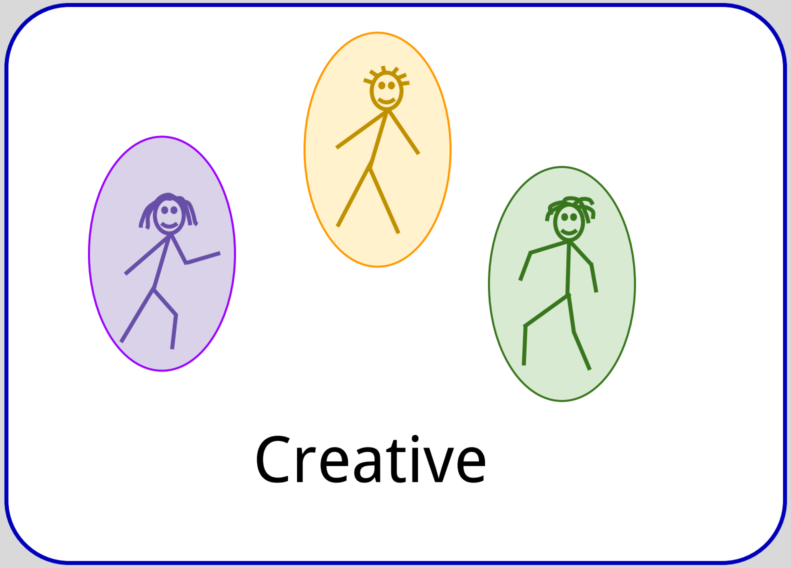
Competition

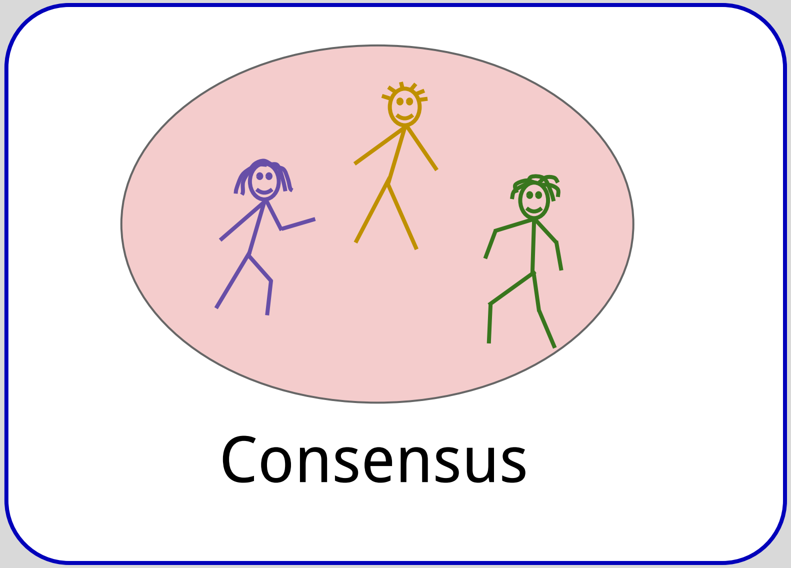


Collaboration


Max


Sum


Creative


Consensus

Outline

1. Basic collaborative framework [AF 2011]
2. Some useful extensions [WFA 2015]
3. Axiomatic investigations [FW 2018]

Outline

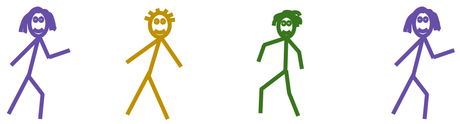
1. Basic collaborative framework [AF 2011]

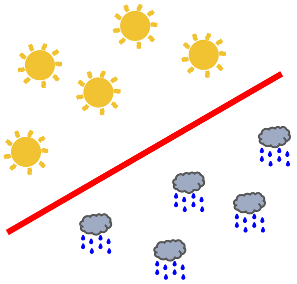
2. Some useful extensions [WFA 2015]

3. Axiomatic investigations [FW 2018]

Basic collaborative framework

A collaborative mechanism for crowdsourcing prediction problems, Abernethy & Frongillo, NIPS 2011

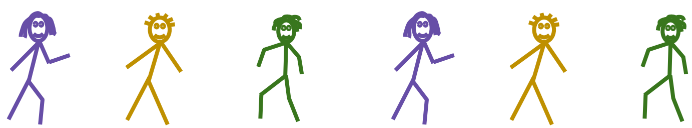


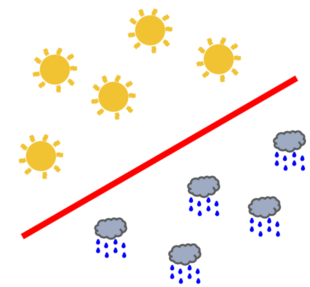


“Scoring Rule Market (SRM)”

Basic collaborative framework

A collaborative mechanism for crowdsourcing prediction problems, Abernethy & Frongillo, NIPS 2011



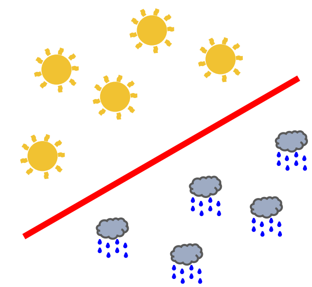


“Scoring Rule Market (SRM)”:

1. Designer chooses initial **public** hypothesis h^0

Basic collaborative framework

A collaborative mechanism for crowdsourcing prediction problems, Abernethy & Frongillo, NIPS 2011

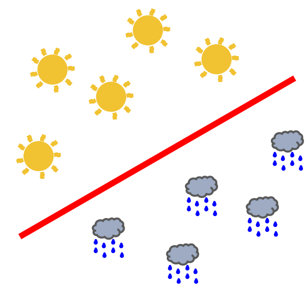


“Scoring Rule Market (SRM)”:

1. Designer chooses initial **public** hypothesis h^0
2. Participant t=1,..., proposes **public update** $h^{t-1} \rightarrow h^t$

Basic collaborative framework

A collaborative mechanism for crowdsourcing prediction problems, Abernethy & Frongillo, NIPS 2011

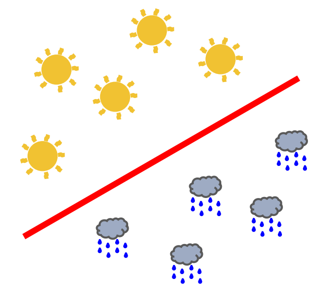


“Scoring Rule Market (SRM)”:

1. Designer chooses initial **public** hypothesis h^0
2. Participant t=1,..., proposes **public update** $h^{t-1} \rightarrow h^t$
3. Validation set D revealed

Basic collaborative framework

A collaborative mechanism for crowdsourcing prediction problems, Abernethy & Frongillo, NIPS 2011



"Scoring Rule Market (SRM)":

1. Designer chooses initial **public** hypothesis h^0
2. Participant t=1,..., proposes **public update** $h^{t-1} \rightarrow h^t$
3. Validation set D revealed
4. Reward for t is $L(h^{t-1}, D) - L(h^t, D)$

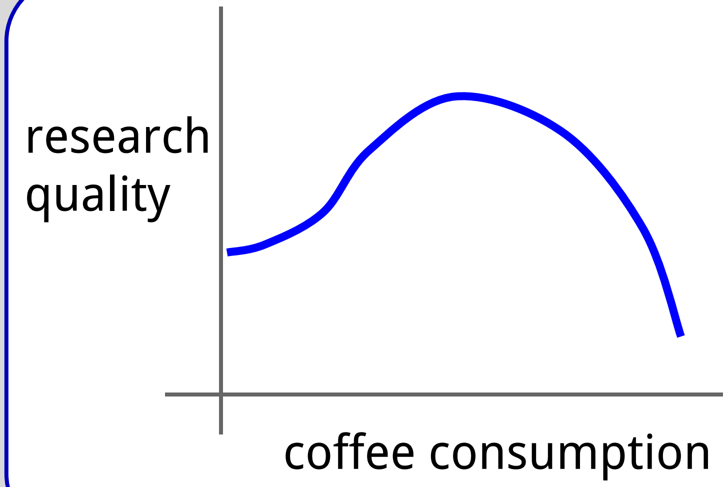
Outline

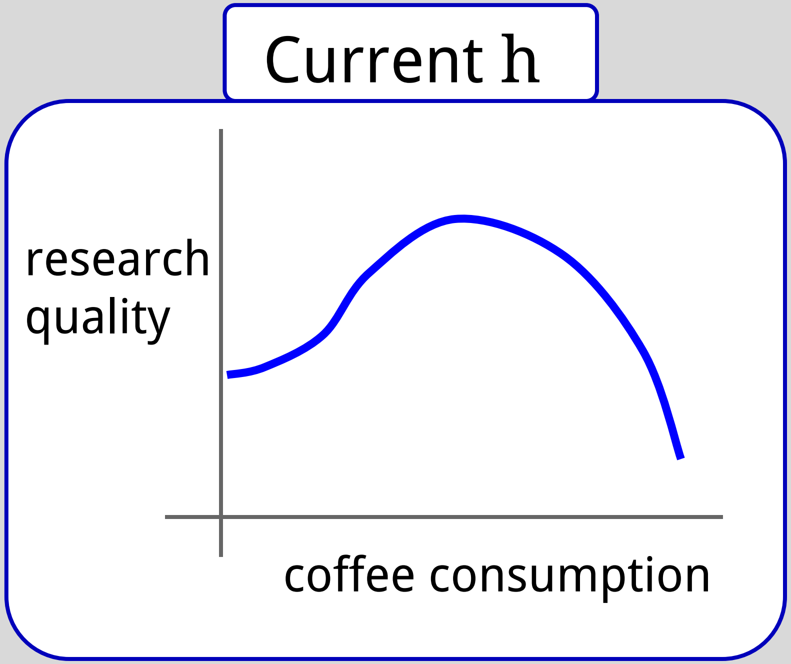
✓ 1. Basic collaborative framework [AF 2011]

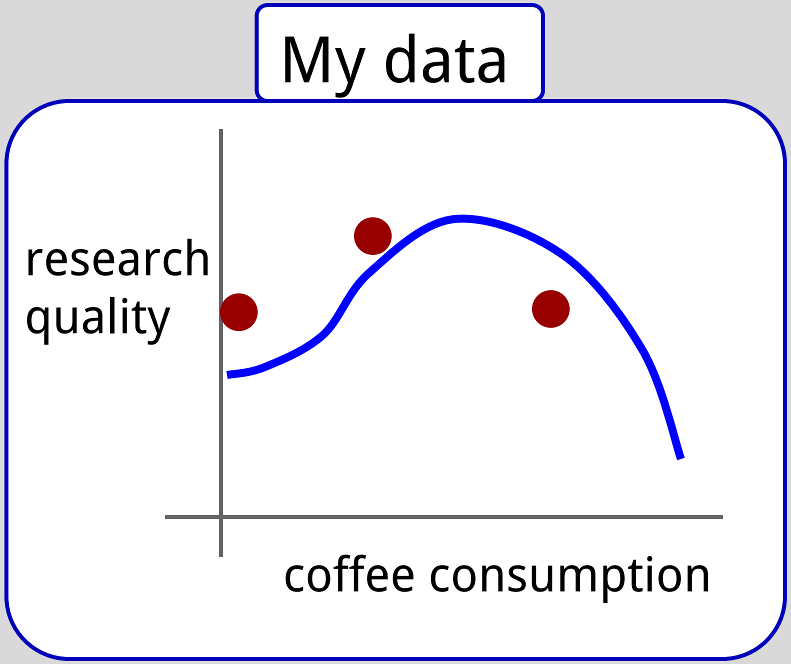
2. Some useful extensions [WFA 2015]

3. Axiomatic investigations [FW 2018]

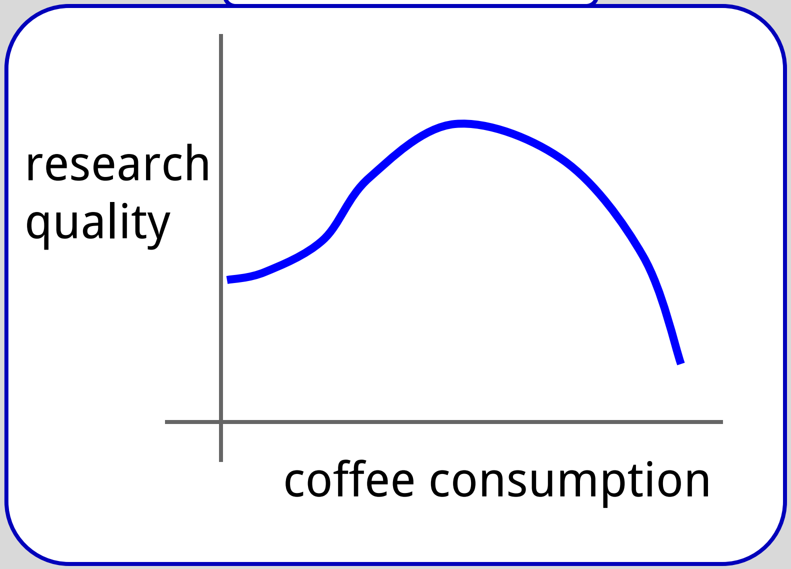
Current h



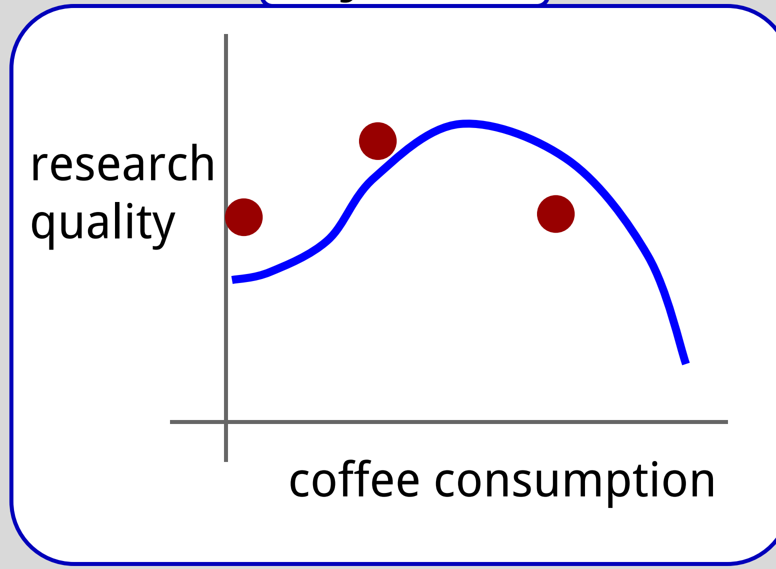

Current h
research quality
coffee consumption

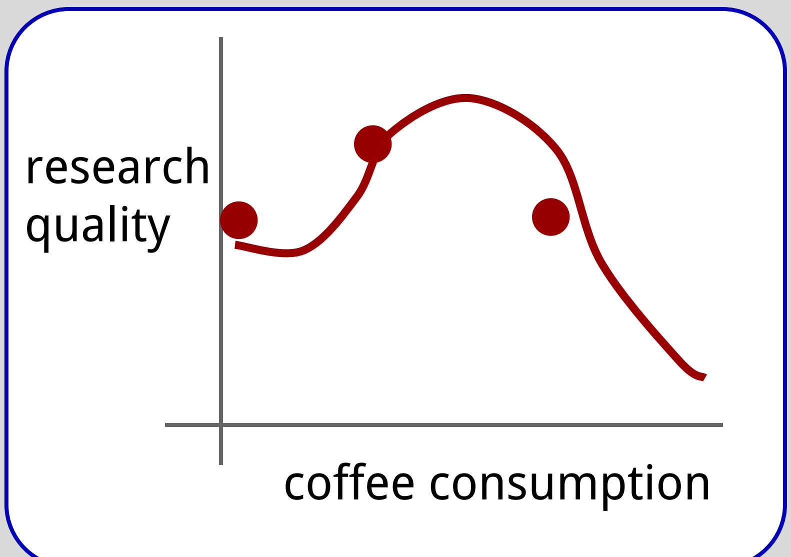

My data
research quality
coffee consumption

Current h

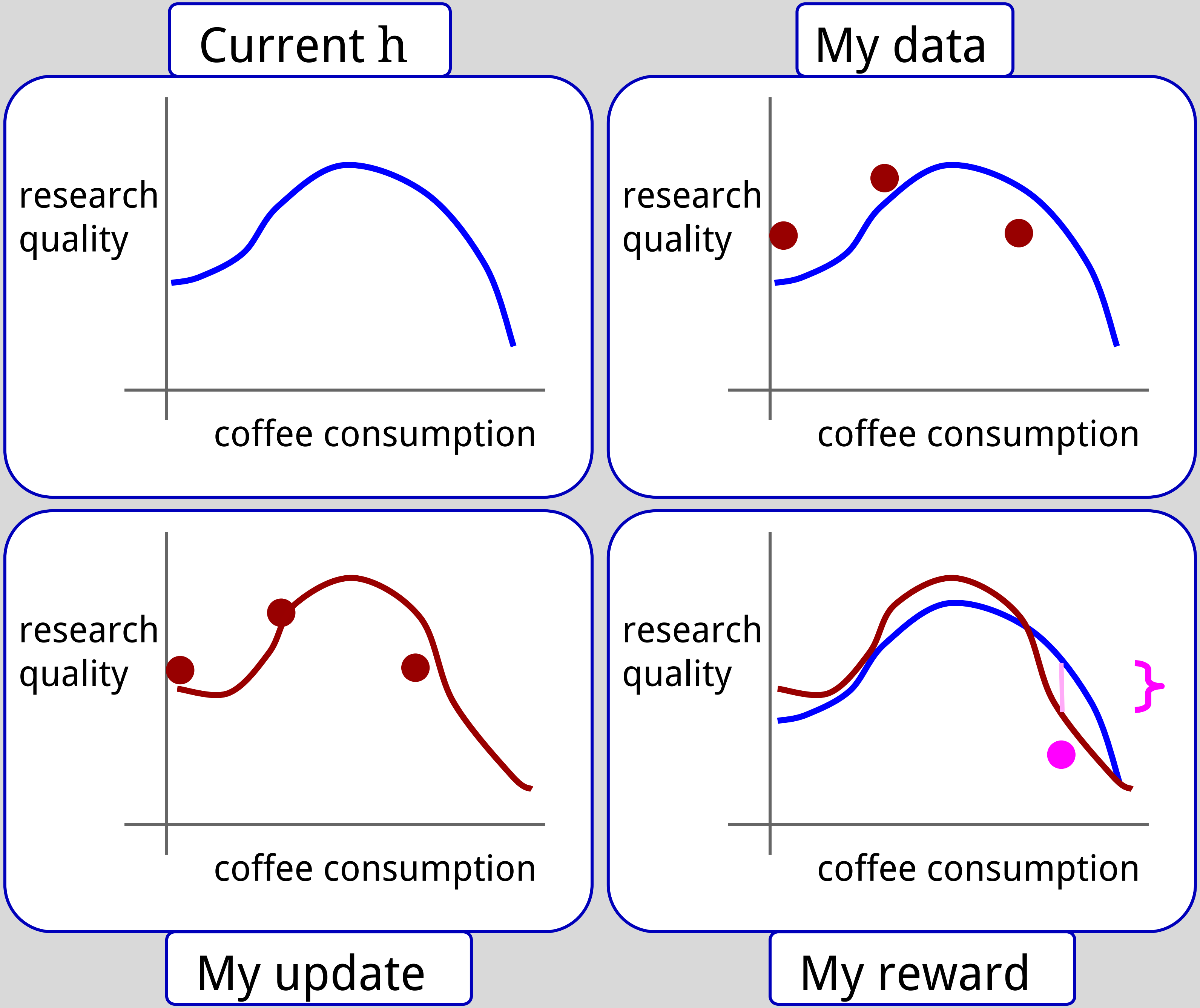


My data





My update



Current h
My data
research quality
coffee consumption
research quality
coffee consumption
research quality
coffee consumption
research quality
coffee consumption
My update
My reward

Some useful extensions

A market framework for eliciting private data, Waggoner, Frongillo, and Abernethy, NIPS 2015.

Cost function based markets:

1. Designer chooses "feature function" f^0

Some useful extensions

A market framework for eliciting private data, Waggoner, Frongillo, and Abernethy, NIPS 2015.

Cost function based markets:

1. Designer chooses “feature function” f^0
2. Each t=1,..., updates $f^{t-1} \rightarrow f^t$ and pays $C(f^t)$ - $C(f^{t-1})$

Some useful extensions

A market framework for eliciting private data, Waggoner, Frongillo, and Abernethy, NIPS 2015.

Cost function based markets:

1. Designer chooses "feature function" f^0
2. Each $t=1,\ldots$, updates $f^{t-1} \rightarrow f^t$ and pays $C(f^t) - C(f^{t-1})$
3. Validation set revealed

Some useful extensions

A market framework for eliciting private data, Waggoner, Frongillo, and Abernethy, NIPS 2015.

Cost function based markets:

1. Designer chooses "feature function" f^0
2. Each t=1,..., updates $f^{t-1} \rightarrow f^t$ and pays $C(f^t) - C(f^{t-1})$
3. Validation set revealed
4. Reward for t is $\sum_{x,y \in D} df^t(x,y)$

Some useful extensions

A market framework for eliciting private data, Waggoner, Frongillo, and Abernethy, NIPS 2015.

Cost function based markets:

1. Designer chooses "feature function" f^0
2. Each t=1,..., updates $f^{t-1} \rightarrow f^t$ and pays $C(f^t) - C(f^{t-1})$
3. Validation set revealed
4. Reward for t is $\sum_{x,y \in D} df^t(x,y)$

Fact (extension of prior results):
Cost function based with RKHS F is equivalent to SRM with a Bregman divergence-based loss function.

Outline

✓ 1. Basic collaborative framework [AF 2011]

✓ 2. Some useful extensions [WFA 2015]

➡ 3. Axiomatic investigations [FW 2018]

Axiomatic investigations

An axiomatic study of scoring rule markets. Frongillo and Waggoner, ITCS 2018.

When/why are SRMs (collaborative contests) effective?

Plan:

- Introduce axioms
- Show examples where they are violated
  $\longrightarrow$ demonstrate why they're desirable
- Characterize satisfaction of the axioms

Axioms

Define **liability:** participant's worst-case payment.

Axioms

Define **liability:** participant's worst-case payment.

Bounded trader budget:
Can make updates with arbitrarily small liability.
$\rightarrow$ *market can scale relative to participants*

Axioms

Define **liability:** participant's worst-case payment.

Bounded trader budget:
Can make updates with arbitrarily small liability.
$\longrightarrow$ *market can scale relative to participants*

Trade neutralization:
Given a previous update yielding liability d, there exists an update that yields constant net liability $< d$.
$\longrightarrow$ *can sell the previous contract back for a nontrivial price*

Axioms

Define **liability:** participant's worst-case payment.

Bounded trader budget:
Can make updates with arbitrarily small liability.
$\longrightarrow$ *market can scale relative to participants*

Trade neutralization:
Given a previous update yielding liability d, there exists an update that yields <u>constant</u> net liability $< d$.
$\longrightarrow$ *can sell the previous contract back for a nontrivial price*

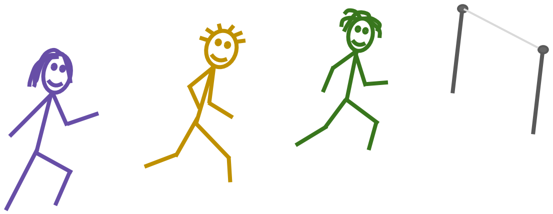
Weak neutralization:
Given a previous update yielding liability d, there exists an update that yields net liability $< d$.

Axioms cannot be satisfied for this loss

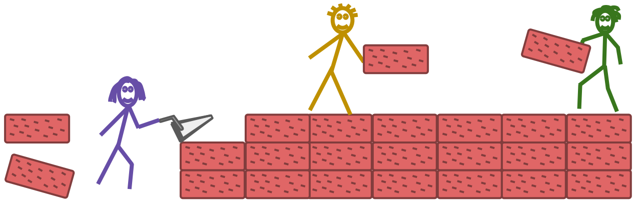
⇒ collaborative mechanism <u>in</u>effective

Axioms can be satisfied

⇒ collaborative mechanism effective



Find the best
domain expert?



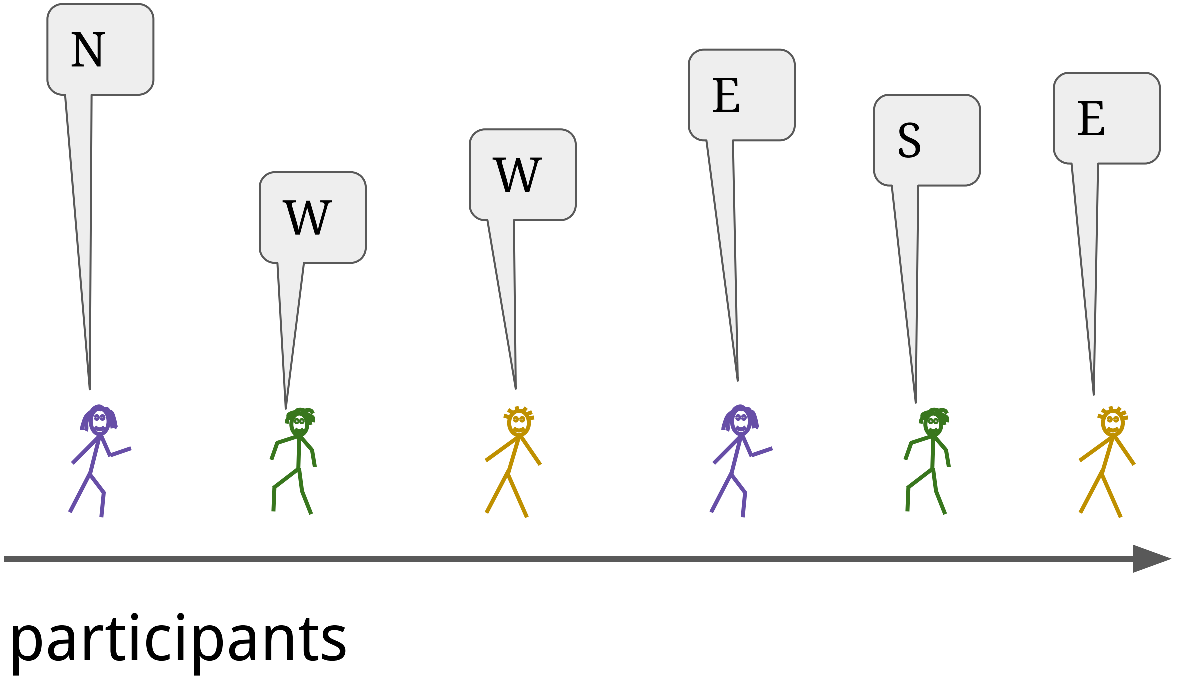
…or aggregate
from “the crowd”?

Example: categorical classification

The wind tomorrow will most likely blow from the:

- North?
- East?
- South?
- West?
- Calm?

Using: 0-1 loss.



Example: categorical classification

Conjecture (B. Dylan, 1965):
"You don't need a weatherman to know which way the wind blows."

Example: categorical classification

Conjecture (B. Dylan, 1965):
"You don't need a weatherman to know which way the wind blows."

Theorem (Frongillo, Waggoner 2018):
Actually, you kinda do.

Example: categorical classification

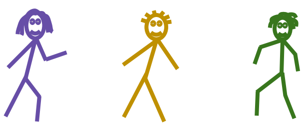
Conjecture (B. Dylan, 1965):
"You don't need a weatherman to know which way the wind blows."

Theorem (Frongillo, Waggoner 2018):
No "scoring-rule market" for categorical classification can satisfy:

- **"Bounded trader budget"**
  ⇒ cannot reach consensus
- **nor "(weak) neutralization".**
  ⇒ participants cannot improve or "cash out"

It’s not all bad

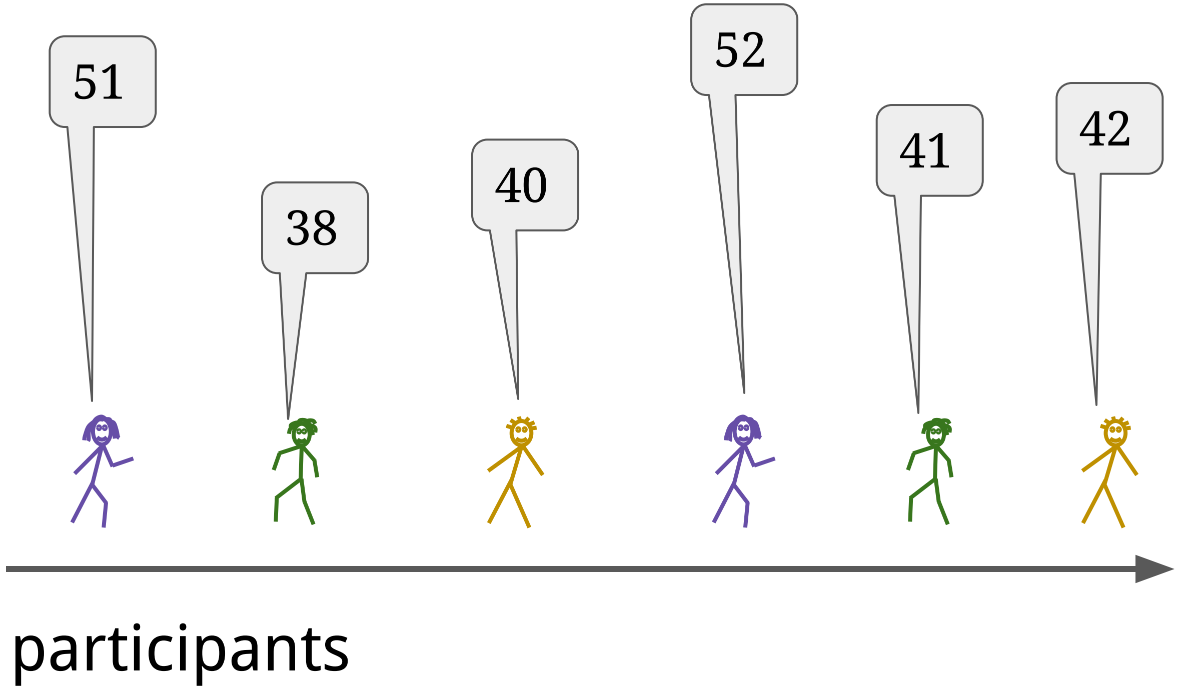
Corrected conjecture:
You don’t need a weatherman to know the wind’s *velocity* via a *surrogate loss*.



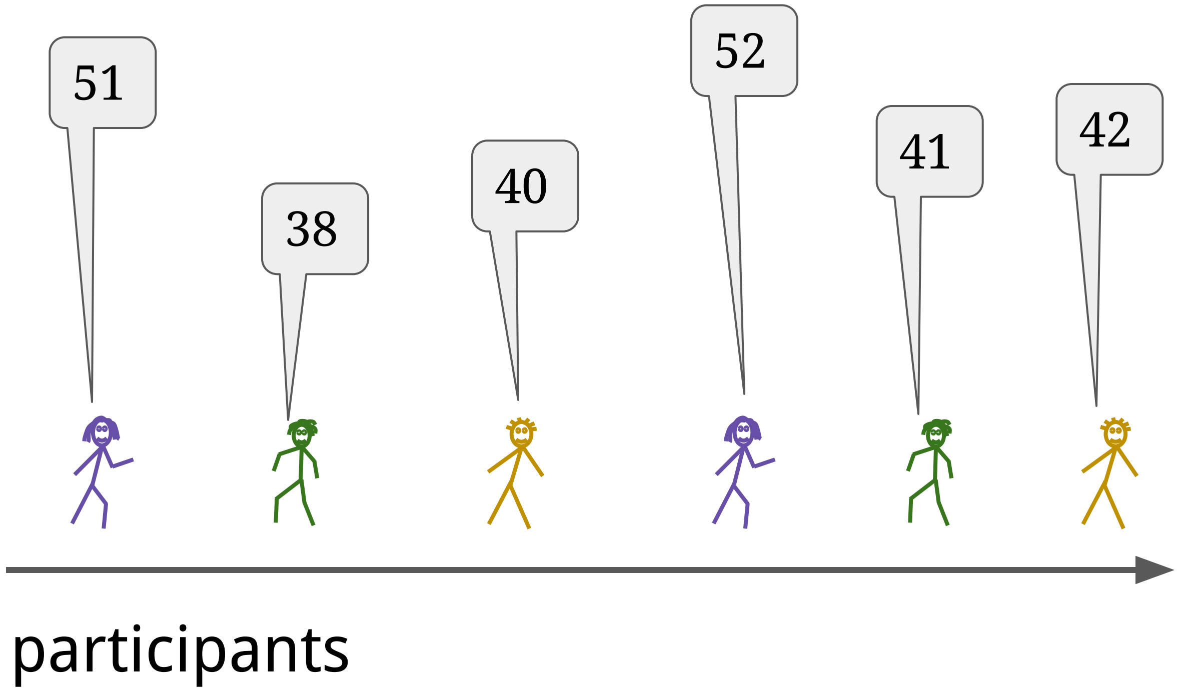
$\theta \in \mathbf{R}^2$

{N, E, S, W, Calm}

Example: Median or quantiles



Example: Median or quantiles



Theorem:
All "scoring-rule markets" for quantiles:

- **satisfy "bounded trader budget"**
- **but not "(weak) neutralization".**

Satisfying trade neutralization

Theorem:
If a scoring-rule market satisfies “trade neutralization”:

- it can be written as a **cost-function based** market
- it elicits a (discretized) **expectation**
  i.e. minimizes a Bregman-divergence loss function.

Satisfying trade neutralization

Theorem:
If a scoring-rule market satisfies "trade neutralization":

- it can be written as a **cost-function based** market
- it elicits a (discretized) **expectation**
  i.e. minimizes a Bregman-divergence loss function.

Theorem:
For any Bregman-divergence loss function (mean), there exists a cost-function based market satisfying all axioms.

Other possibilities

Some markets satisfy weak but not strong neutralization!
$\rightarrow$ Exciting direction for investigation.

Example: **ratio of expectations**, e.g. $\mathbf{E}\, X / \mathbf{E}\, Y$

- Not cost-function based (no trade neutralization)
- But can be written "almost" as cost function…
  … and satisfies weak neutralization!

"Pay" $(Y)(C(f^t) - C(f^{t-1}))$

"Reward" $\sum_{x,y \in D} f^t(X) - f^{t-1}(X)$

Outline

✓ 1. Basic collaborative framework [AF 2011]

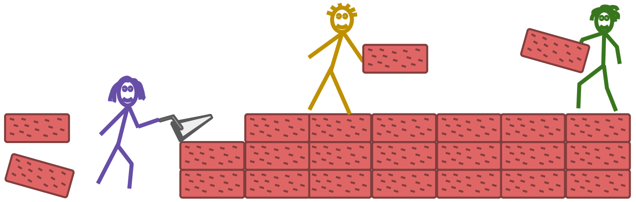
✓ 2. Some useful extensions [WFA 2015]

✓ 3. Axiomatic investigations [FW 2018]

Takeaways

When is the collaborative framework good?

- Parametric form chosen, just need to “buy data”
- Participants with diverse knowledge; non-experts
- Divergence-based losses and means
- (e.g. surrogate losses)



Thanks!

Other Axioms

Incentive compatibility:
Update at each time defines a valid hypothesis;
optimal update is to minimize (some) loss function.

Path independence:
Agents cannot gain by making multiple reports in a row.

Theorem:
IC + PI $\Leftrightarrow$ "scoring rule markets" (collaborative contests).

cf Abernethy, Chen, Wortman-Vaughan 2013