



LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN



ROYAL
HOLLOWAY
UNIVERSITY
OF LONDON

SHARING IS CARING

A Probabilistic Model of Deliberation as
Selective Evidence Disclosure

June 26, 2026

LOFT 16

Rosa Bauer

rosa.bauer@icloud.com

Adrian Haret

adrian.haret@rhul.ac.uk

What is this about?



TWELVE ANGRY MEN
Sydney Lumet
1957



JUROR #8

The wisdom of crowds!



JUROR #8

The wisdom of crowds!



ADRIAN

Ah yes! The Condorcet Jury Theorem.

JUROR #8

The wisdom of crowds!



ADRIAN

Ah yes! The Condorcet Jury Theorem.

CONDORCET

Did someone mention my name?...



JUROR #8

The wisdom of crowds!



ADRIAN

Ah yes! The Condorcet Jury Theorem.

CONDORCET

Did someone mention my name?...



ROSA

But better!

JUROR #8

The wisdom of crowds!



ADRIAN

Ah yes! The Condorcet Jury Theorem.

CONDORCET

Did someone mention my name?...

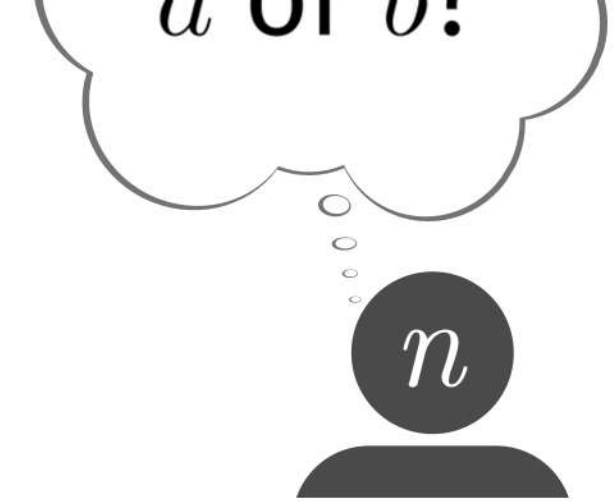


ROSA

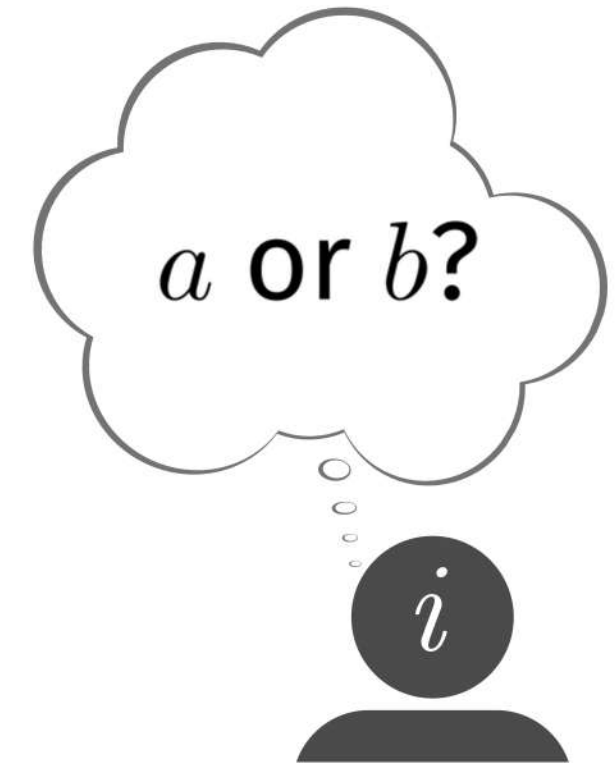
But better! We want to capture the back and forth of deliberation...

MODEL

A group of n agents has to decide between two alternatives, a and b .



...



...

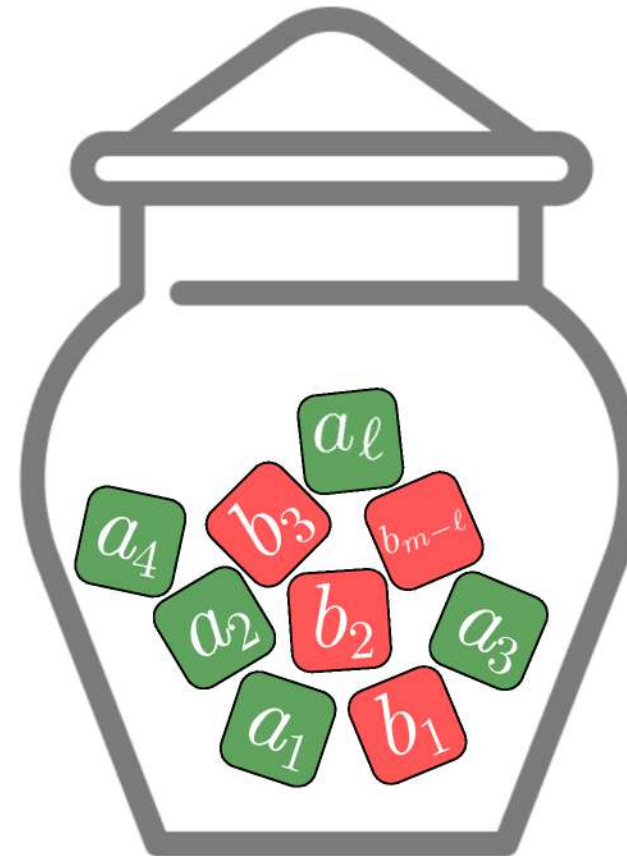


MODEL

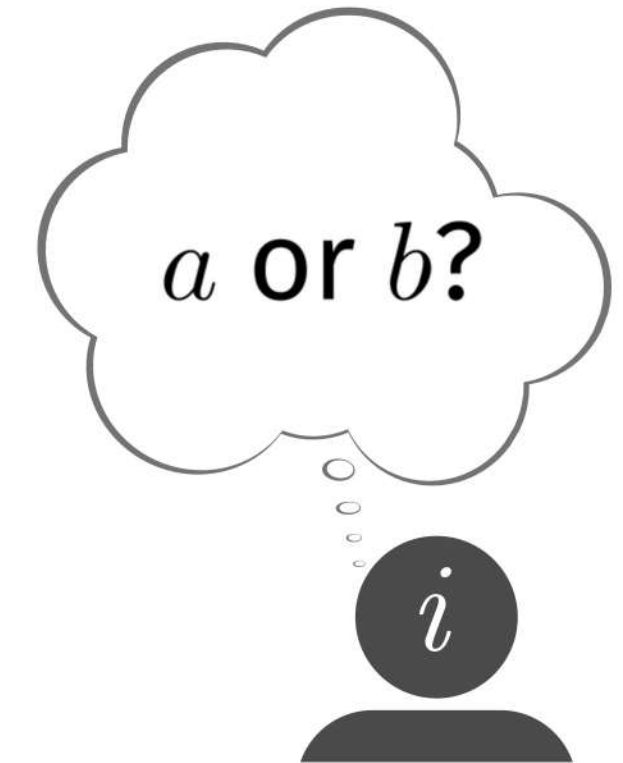
A group of n agents has to decide between two alternatives, a and b .

Each alternative is supported by objective evidence:

$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$



...



...



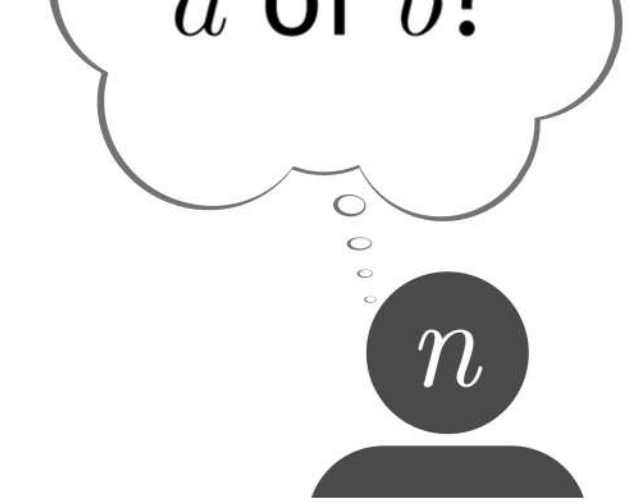
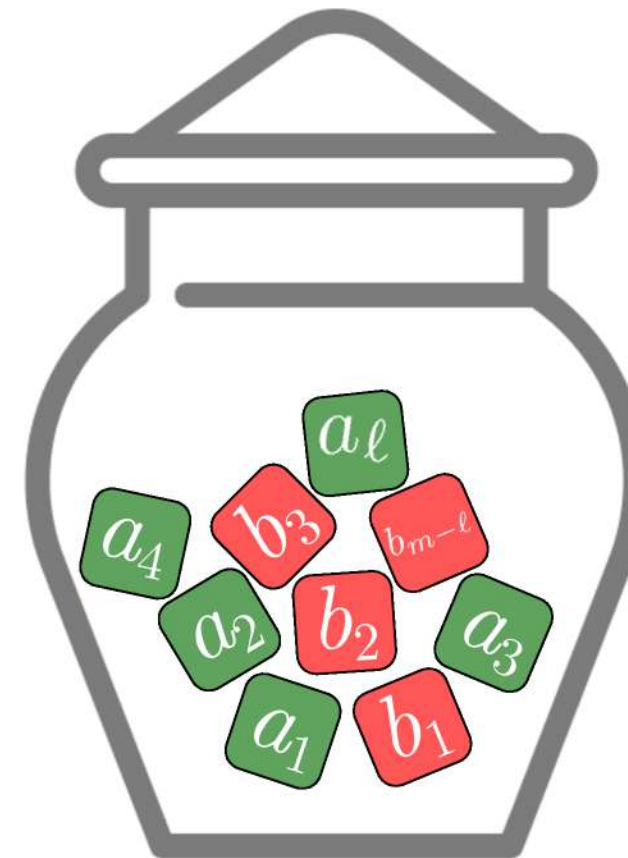
MODEL

A group of n agents has to decide between two alternatives, a and b .

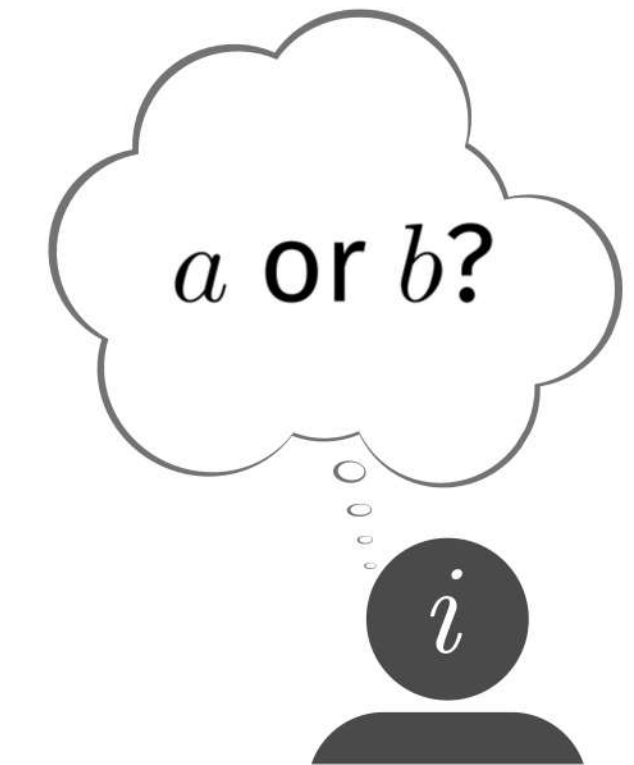
Each alternative is supported by objective evidence:

$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$

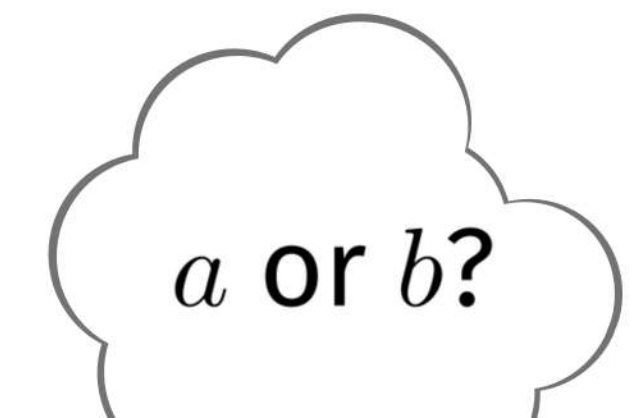
Alternative a is better than b ($\ell > m - \ell$).



...



...



MODEL

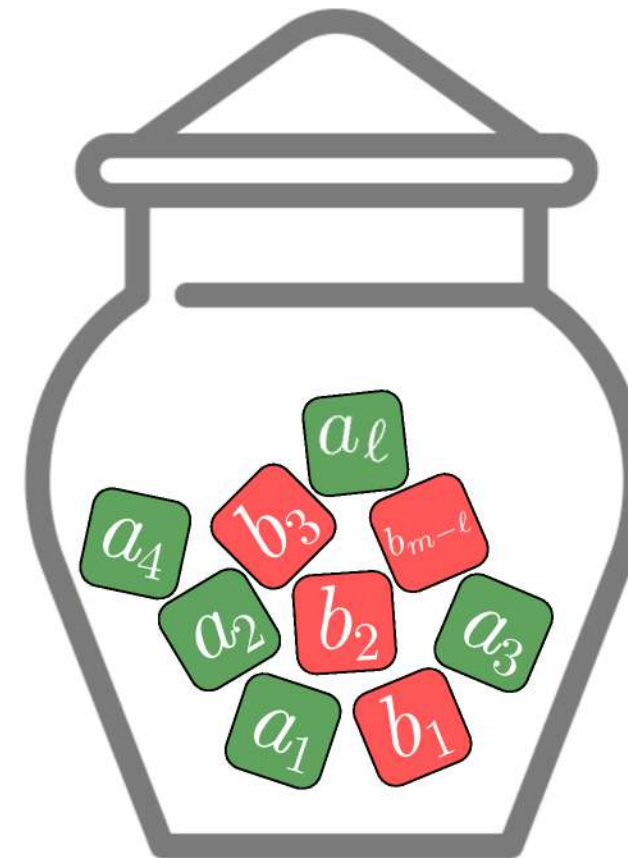
A group of n agents has to decide between two alternatives, a and b .

Each alternative is supported by objective evidence:

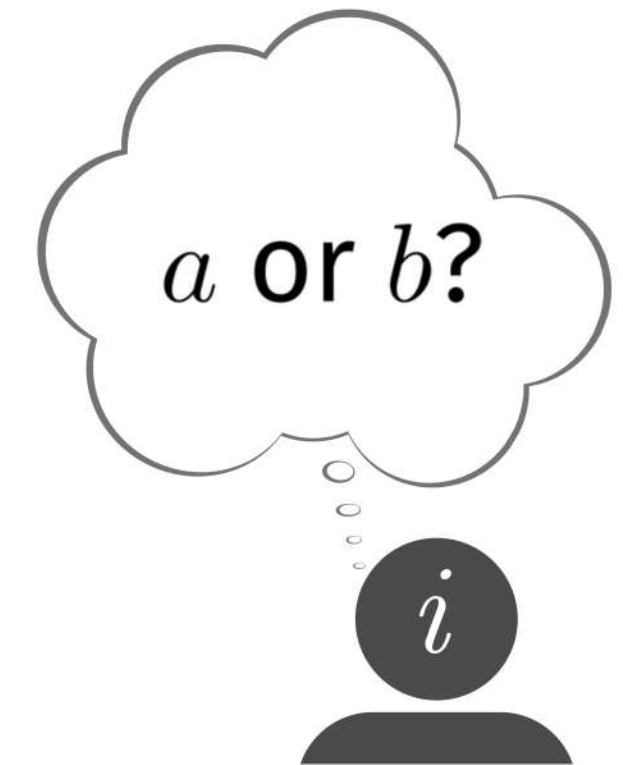
$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$

Alternative a is better than b ($\ell > m - \ell$).

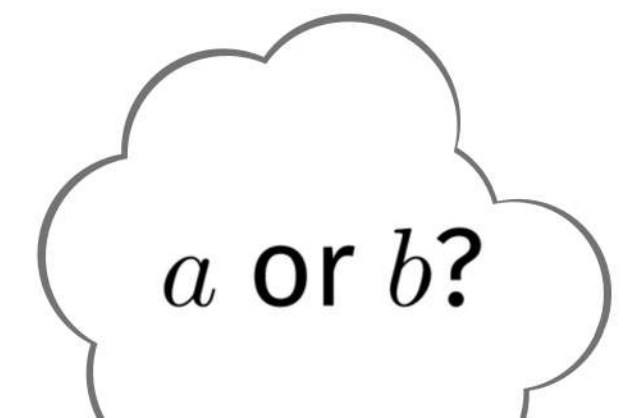
But agents don't know this.



...



...



MODEL

A group of n agents has to decide between two alternatives, a and b .

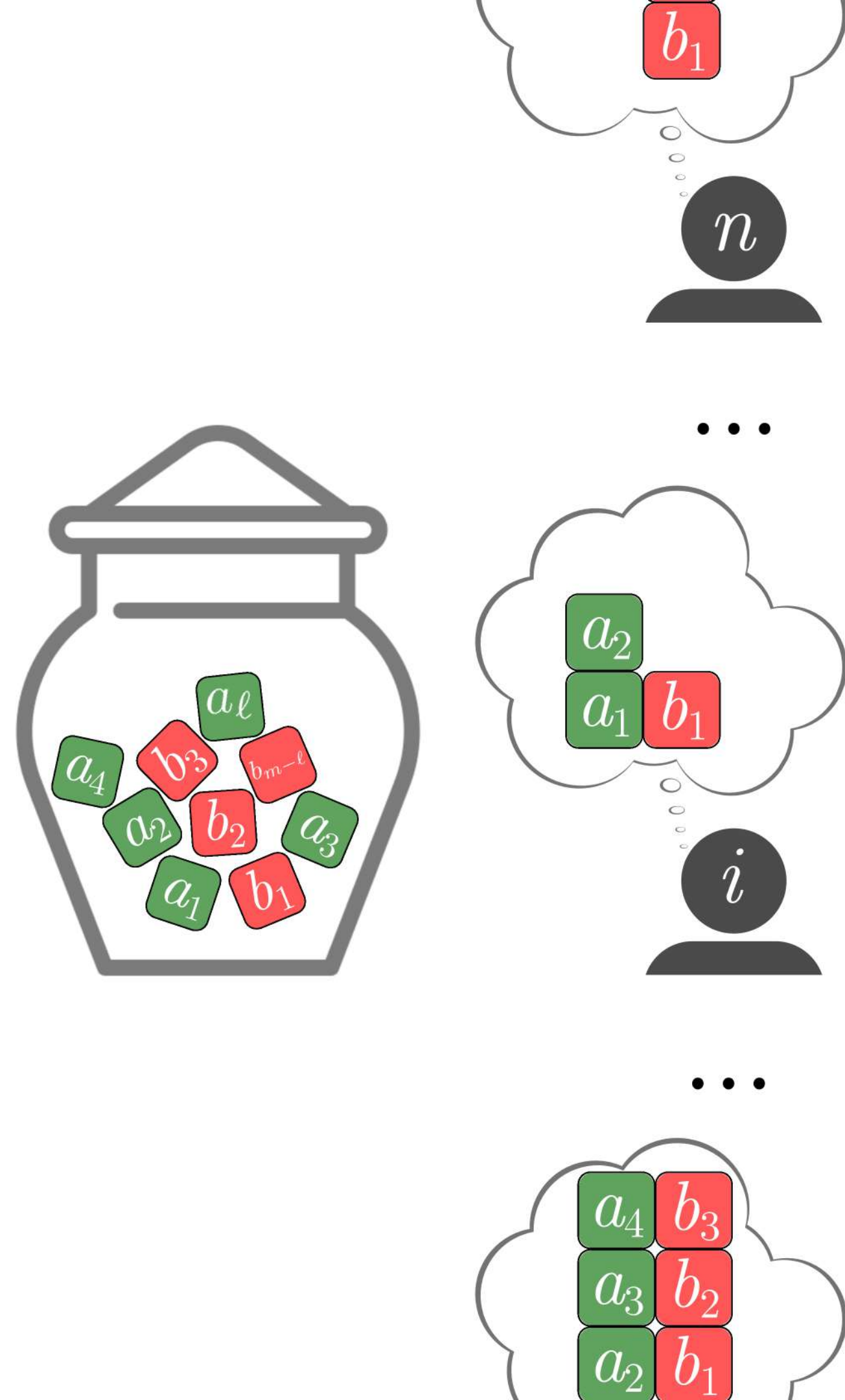
Each alternative is supported by objective evidence:

$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$

Alternative a is better than b ($\ell > m - \ell$).

But agents don't know this.

Each agent has private evidence for a and b



MODEL

A group of n agents has to decide between two alternatives, a and b .

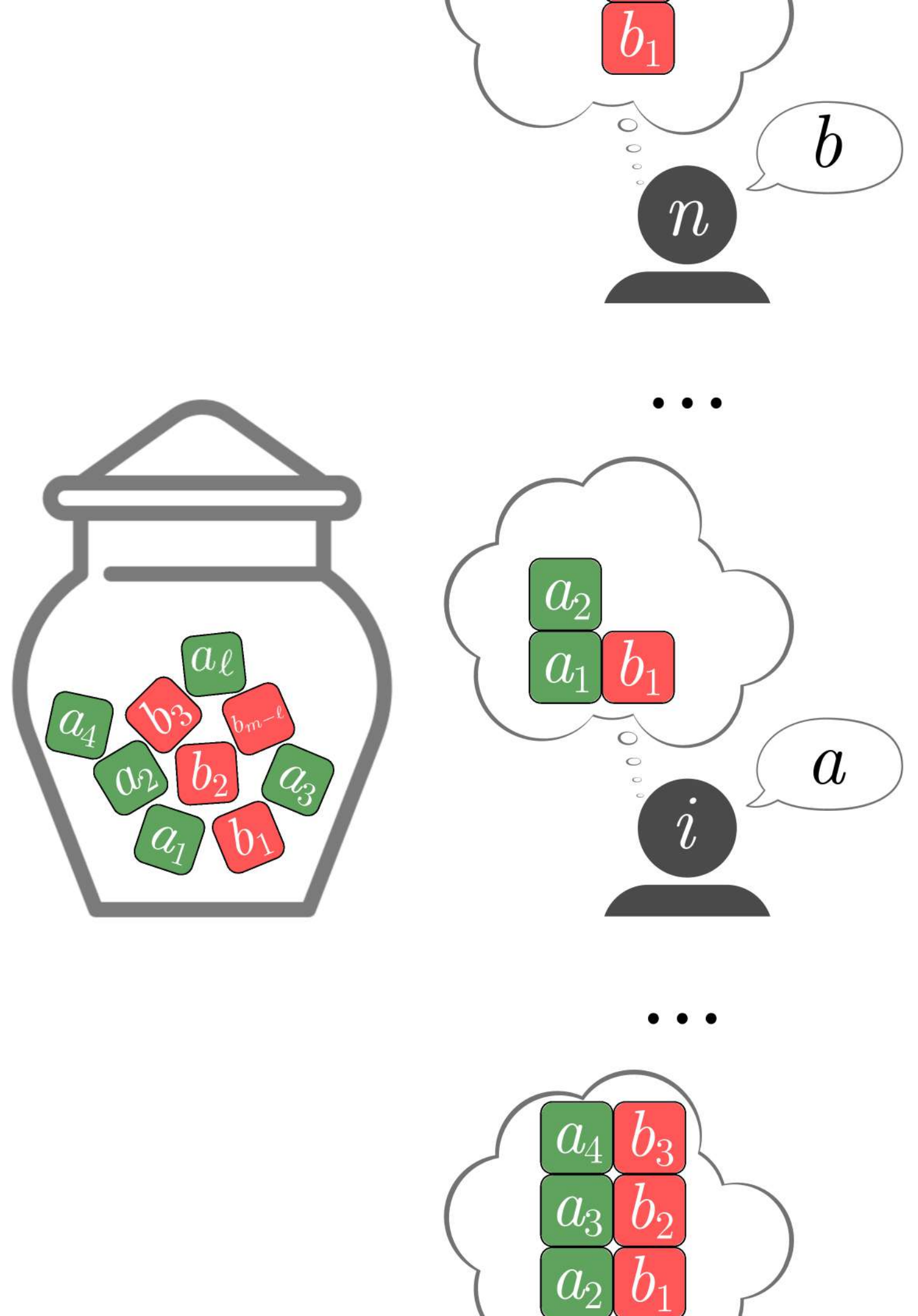
Each alternative is supported by objective evidence:

$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$

Alternative a is better than b ($\ell > m - \ell$).

But agents don't know this.

Each agent has private evidence for a and b , and votes for the alternative with more evidence.*



MODEL

A group of n agents has to decide between two alternatives, a and b .

Each alternative is supported by objective evidence:

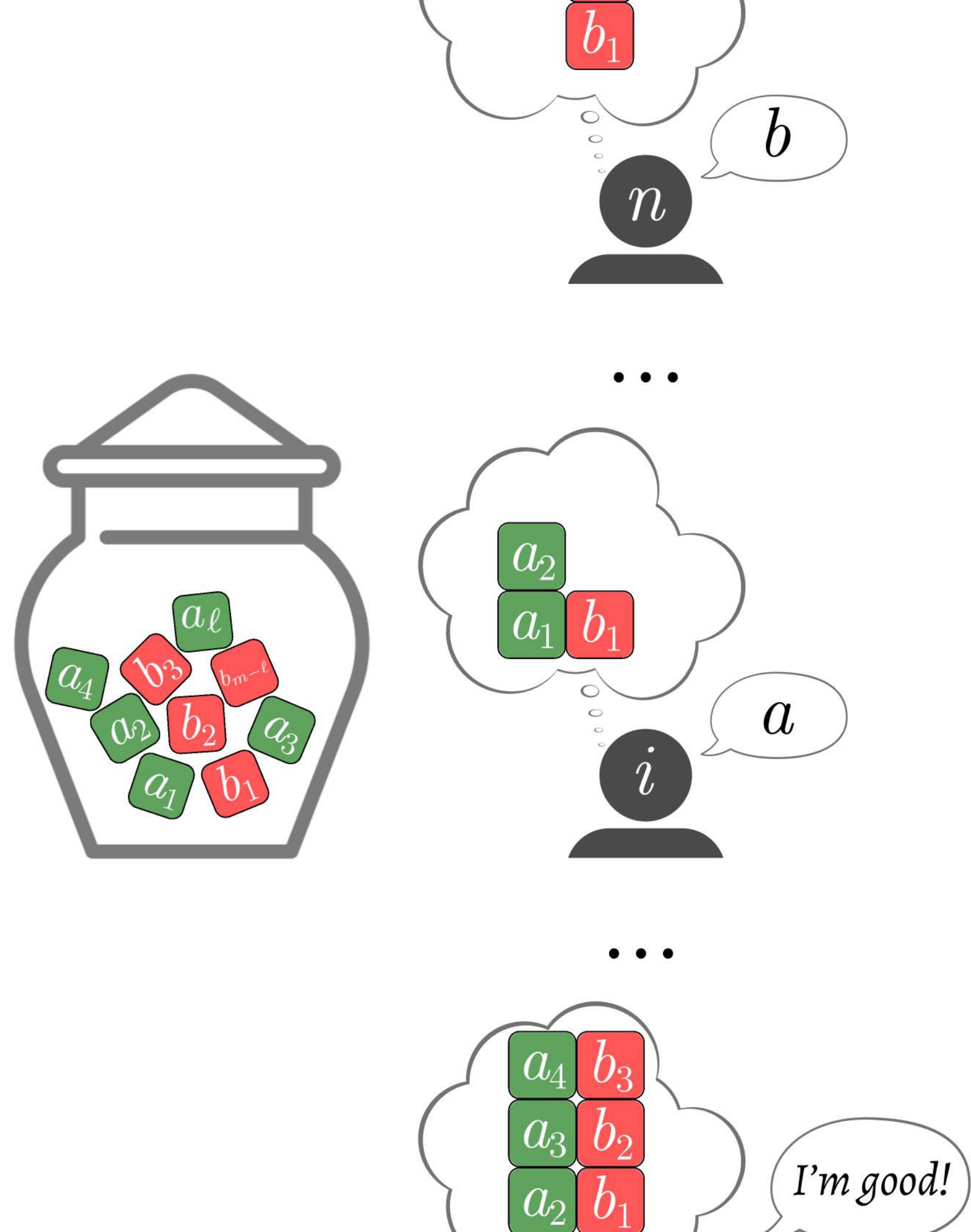
$$a \leftarrow \{a_1, \dots, a_\ell\}, \quad b \leftarrow \{b_1, \dots, b_{m-\ell}\}.$$

Alternative a is better than b ($\ell > m - \ell$).

But agents don't know this.

Each agent has private evidence for a and b , and votes for the alternative with more evidence.*

*In case of a tie they abstain.

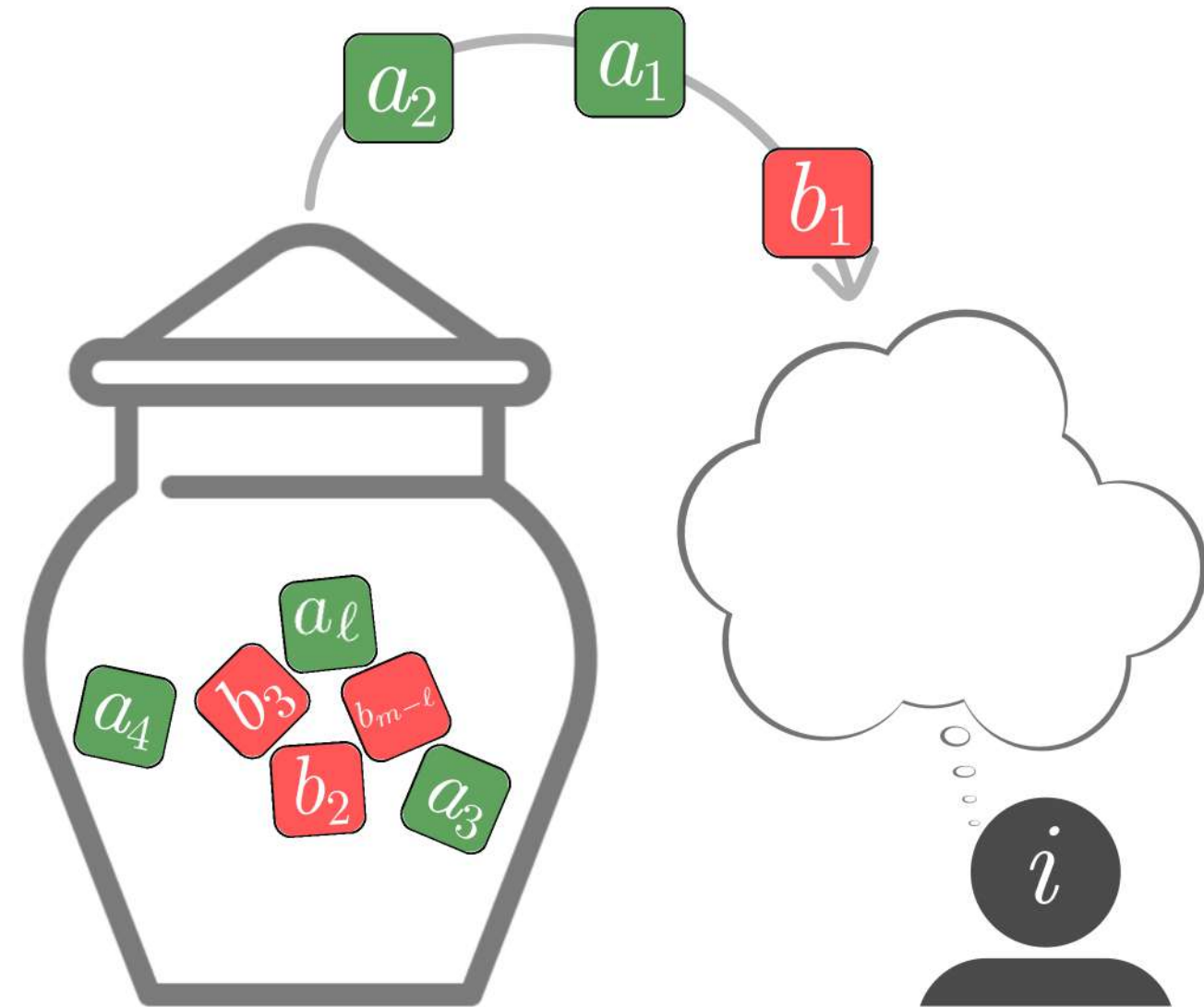


Deliberation occurs in rounds.

MODEL

ROUND o

Each agent samples k items of evidence from the pool of m items.*



*Independently, without replacement.

MODEL

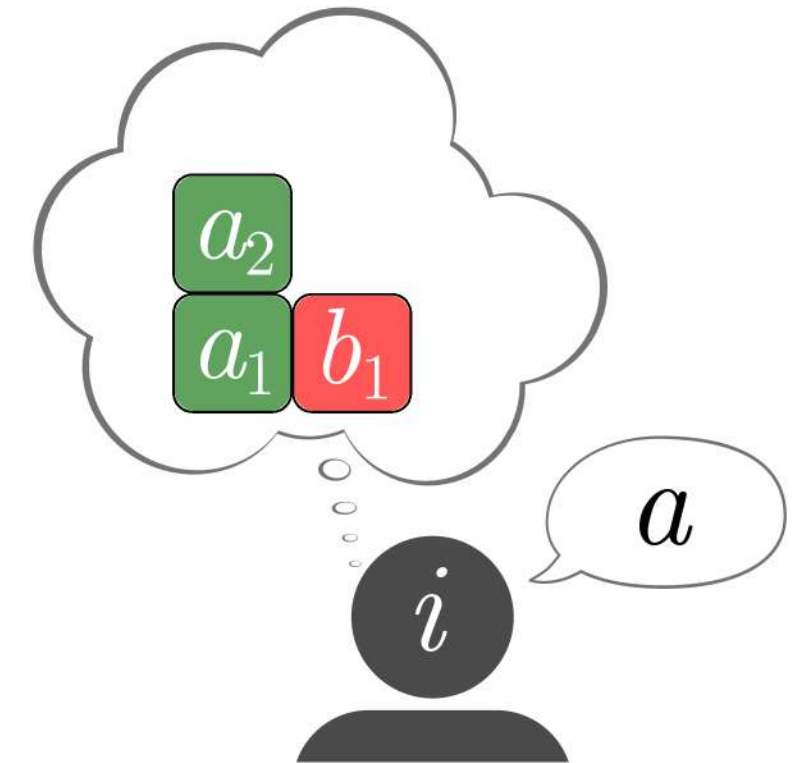
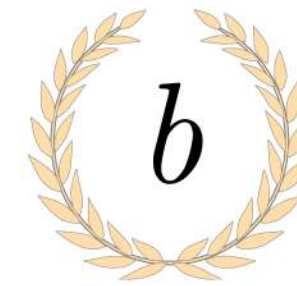
ROUND 0

Each agent samples k items of evidence from the pool of m items.*

ROUND $t \geq 1$

Agents compare their vote with the collective decision.

If they disagree



*Independently, without replacement.

MODEL

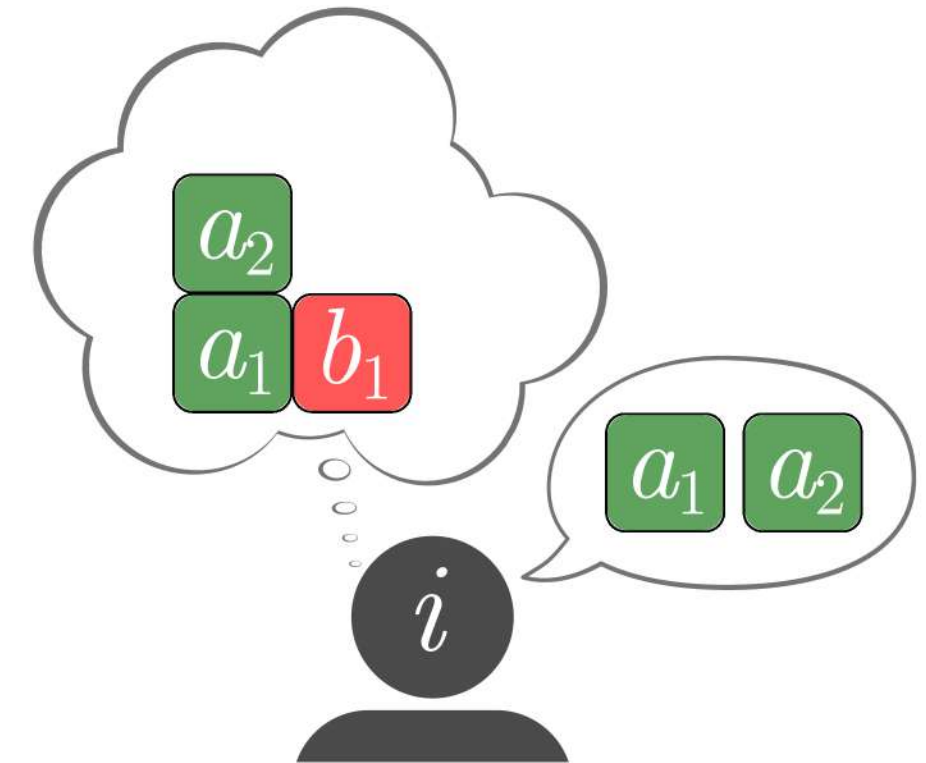
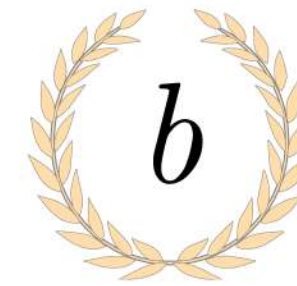
ROUND 0

Each agent samples k items of evidence from the pool of m items.*

ROUND $t \geq 1$

Agents compare their vote with the collective decision.

If they disagree, they disclose all evidence for their preferred alternative (that hasn't been disclosed yet).



*Independently, without replacement.

MODEL

ROUND 0

Each agent samples k items of evidence from the pool of m items.*

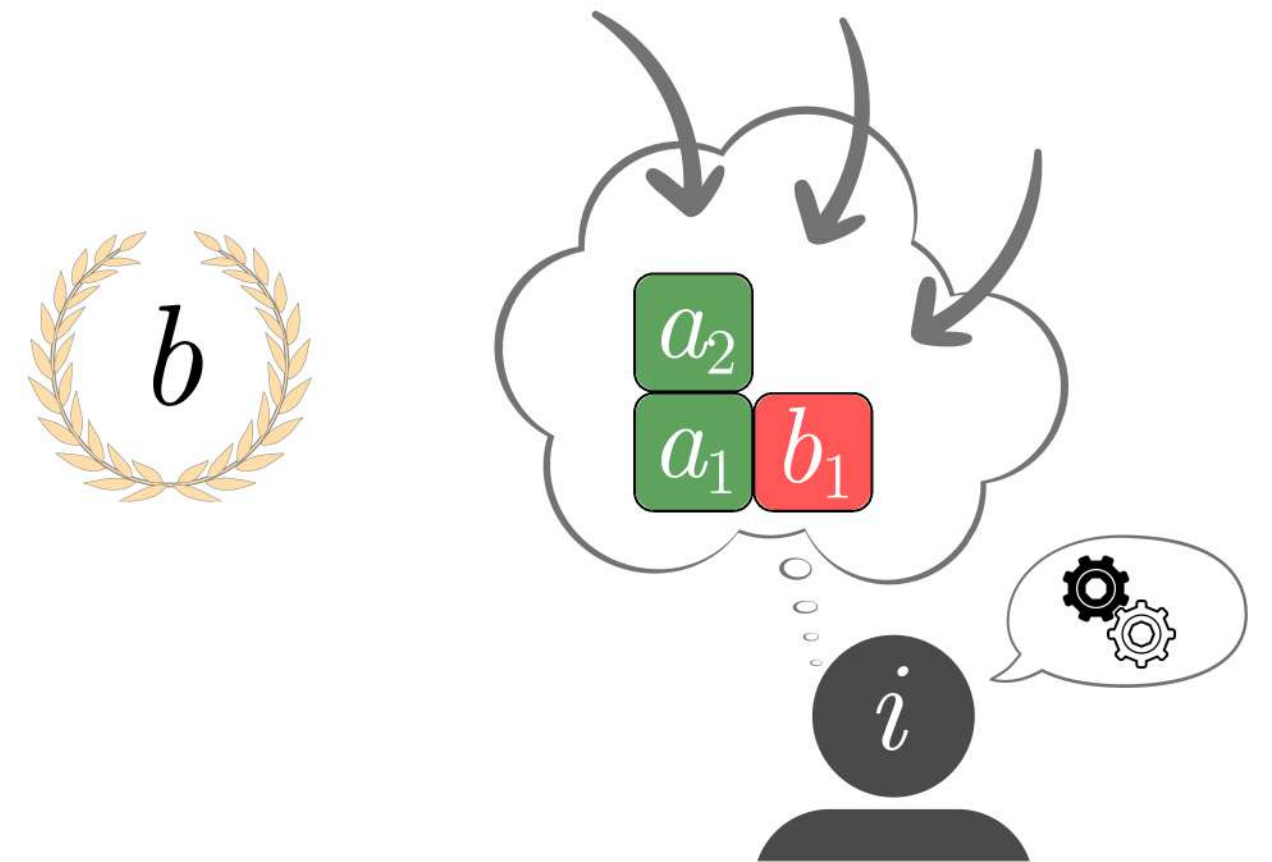
ROUND $t \geq 1$

Agents compare their vote with the collective decision.

If they disagree, they disclose all evidence for their preferred alternative (that hasn't been disclosed yet).

Agents then update private evidence and votes.

*Independently, without replacement.



MODEL

ROUND 0


Each agent samples k items of evidence from the pool of m items.*

ROUND $t \geq 1$

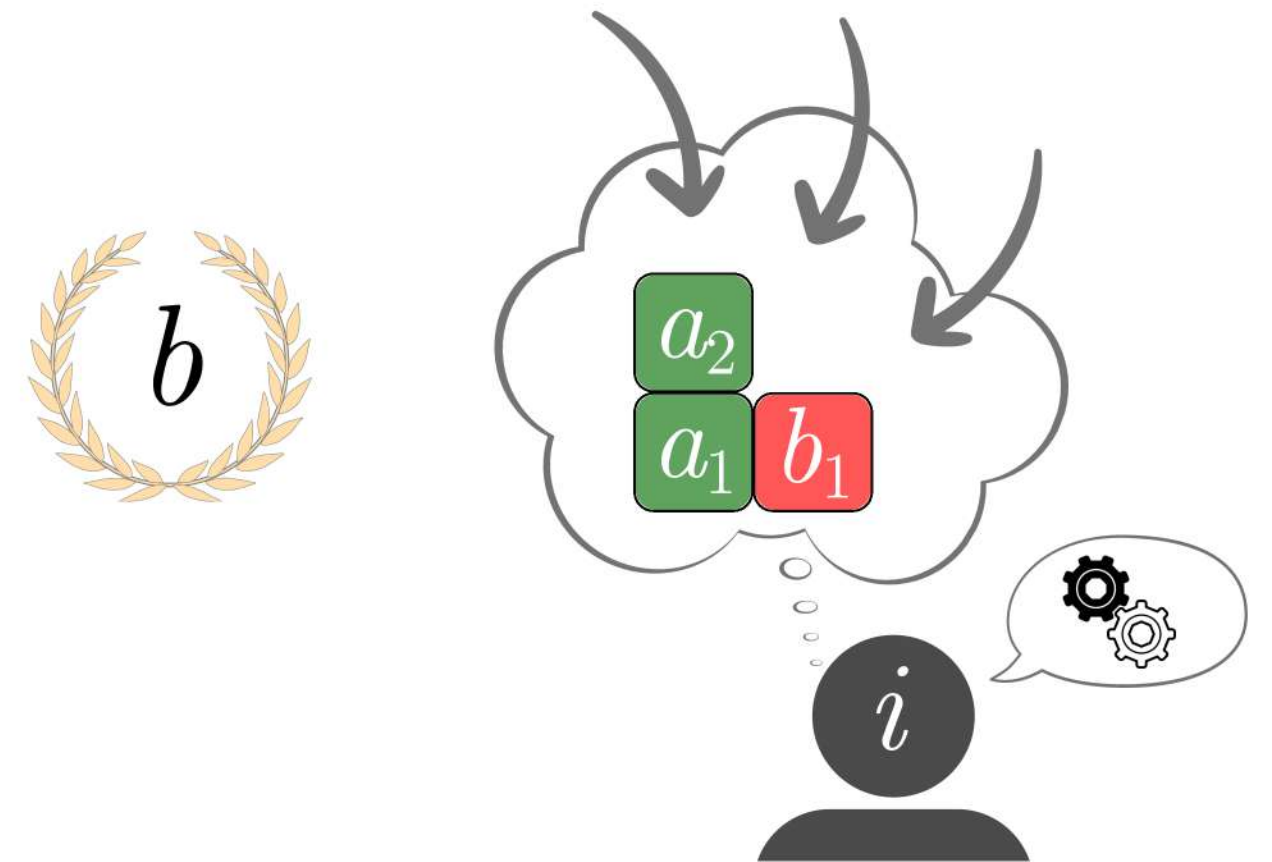
Agents compare their vote with the collective decision.

If they disagree, they disclose all evidence for their preferred alternative (that hasn't been disclosed yet).

Agents then update private evidence and votes.

This keeps going until no new evidence is disclosed.
The collective decision at termination is the outcome of deliberation.** 

*Independently, without replacement.



MODEL

ROUND 0


Each agent samples k items of evidence from the pool of m items.*

ROUND $t \geq 1$

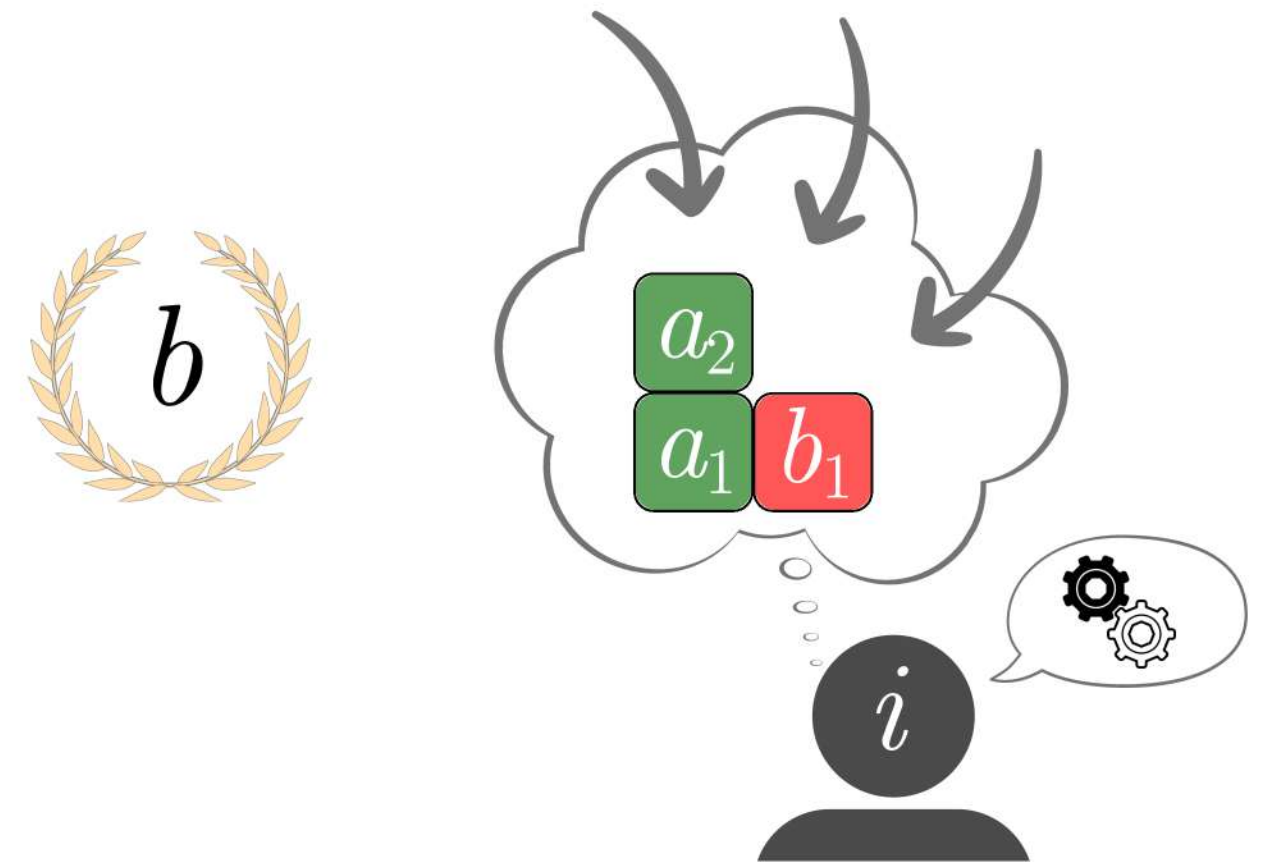
Agents compare their vote with the collective decision.

If they disagree, they disclose all evidence for their preferred alternative (that hasn't been disclosed yet).

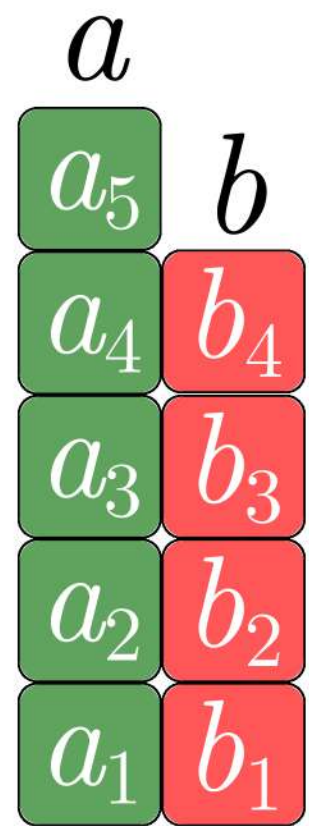
Agents then update private evidence and votes.

This keeps going until no new evidence is disclosed.
The collective decision at termination is the outcome of deliberation.** 

*Independently, without replacement.
**Deliberation is guaranteed to terminate.

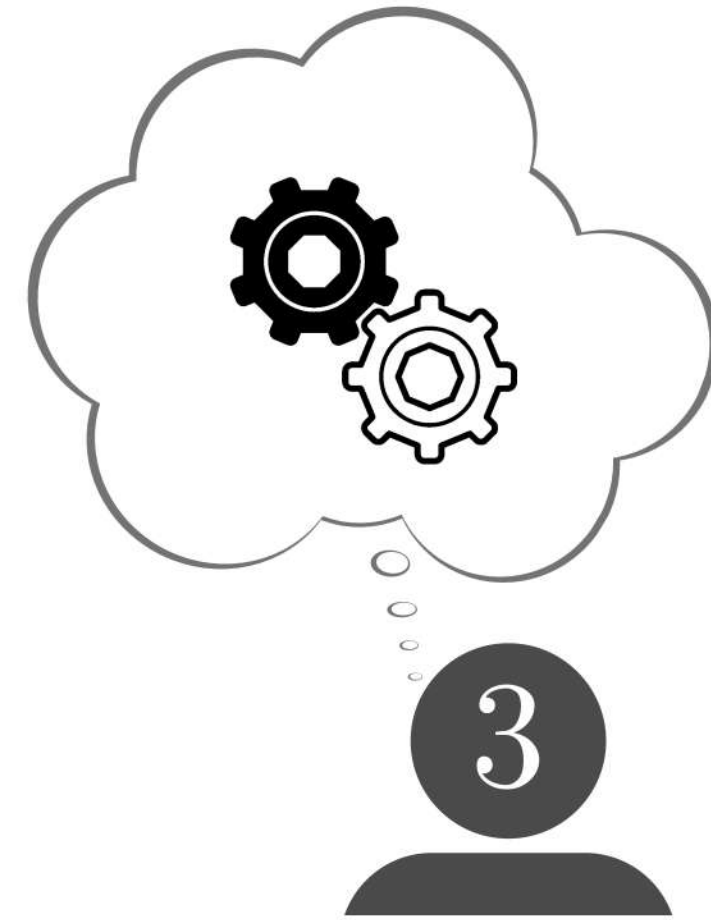
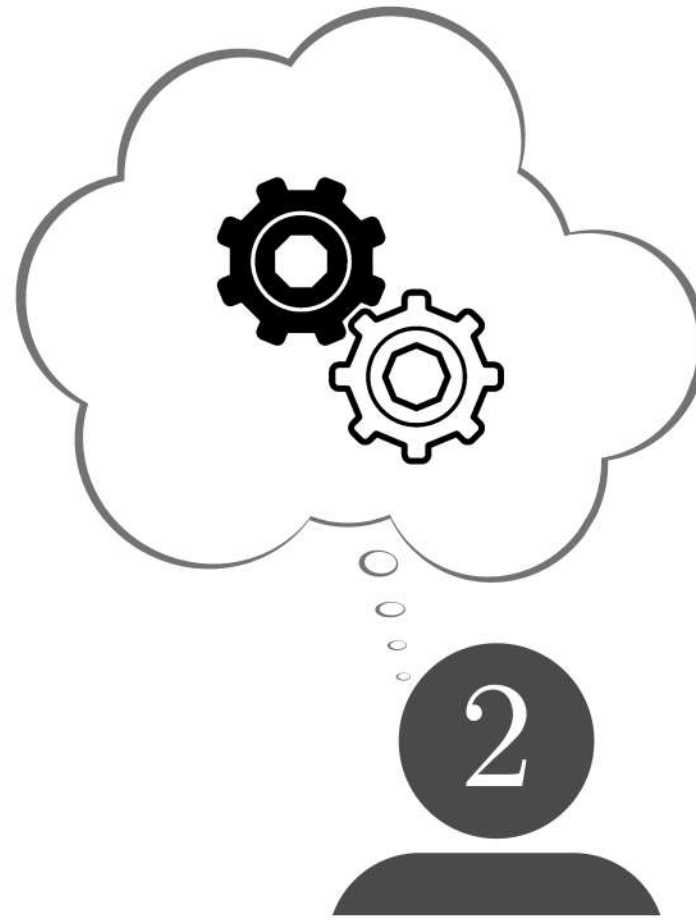
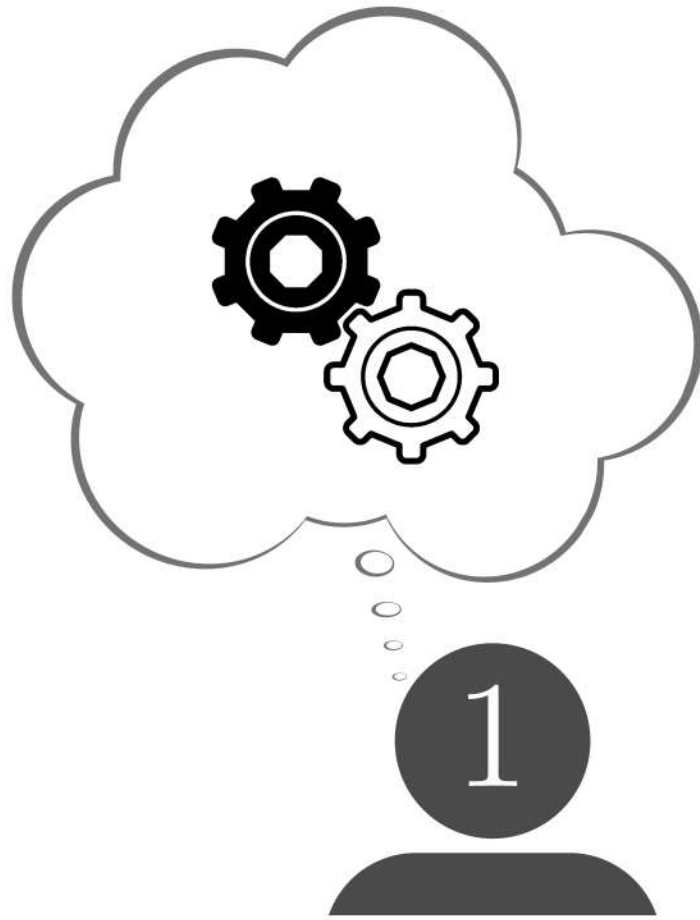
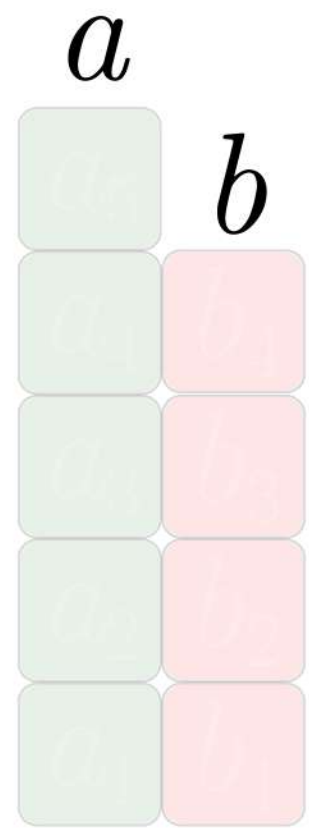


EXAMPLE



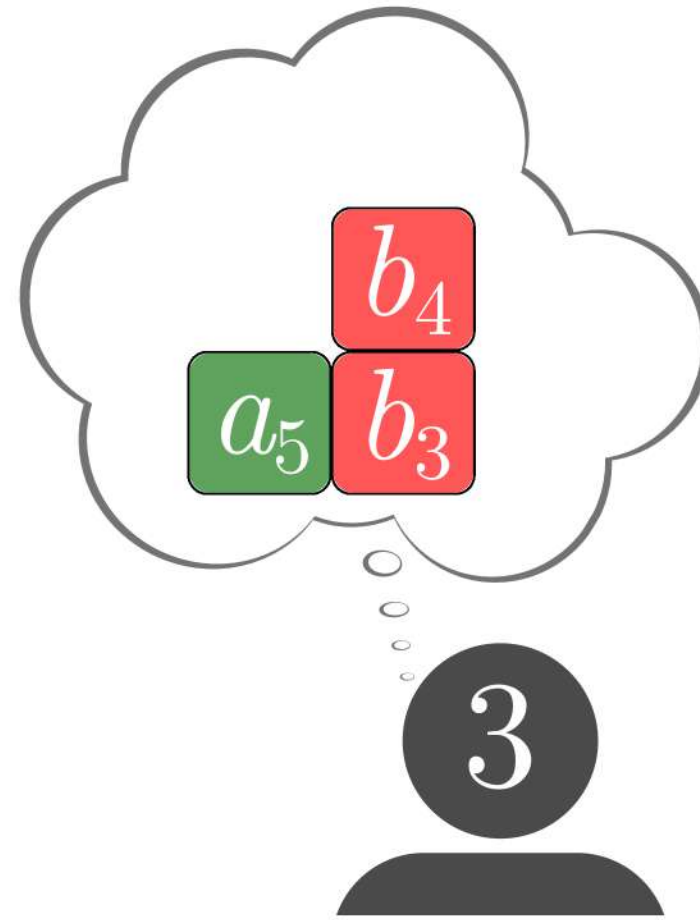
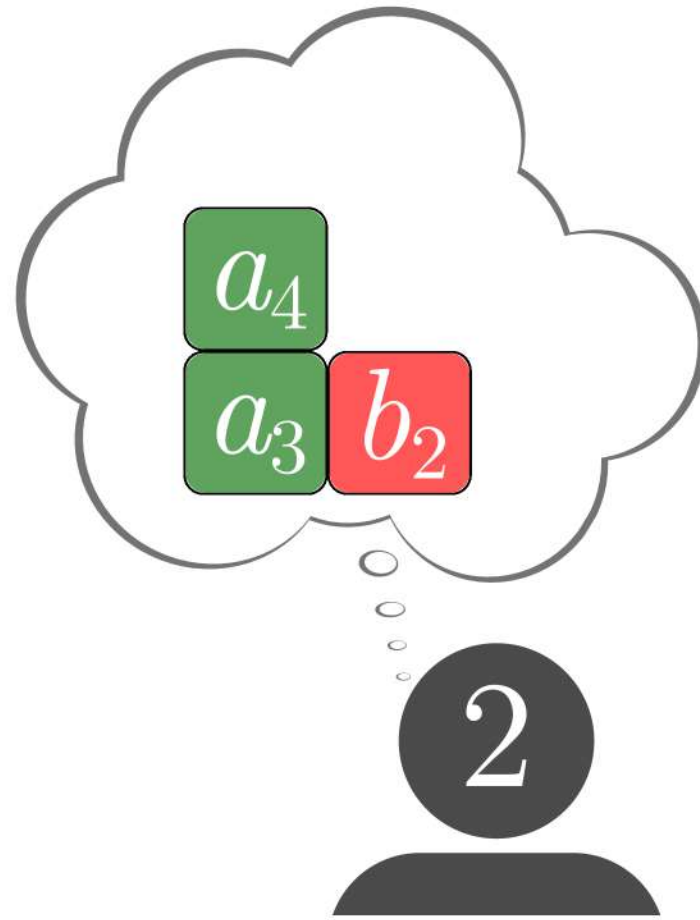
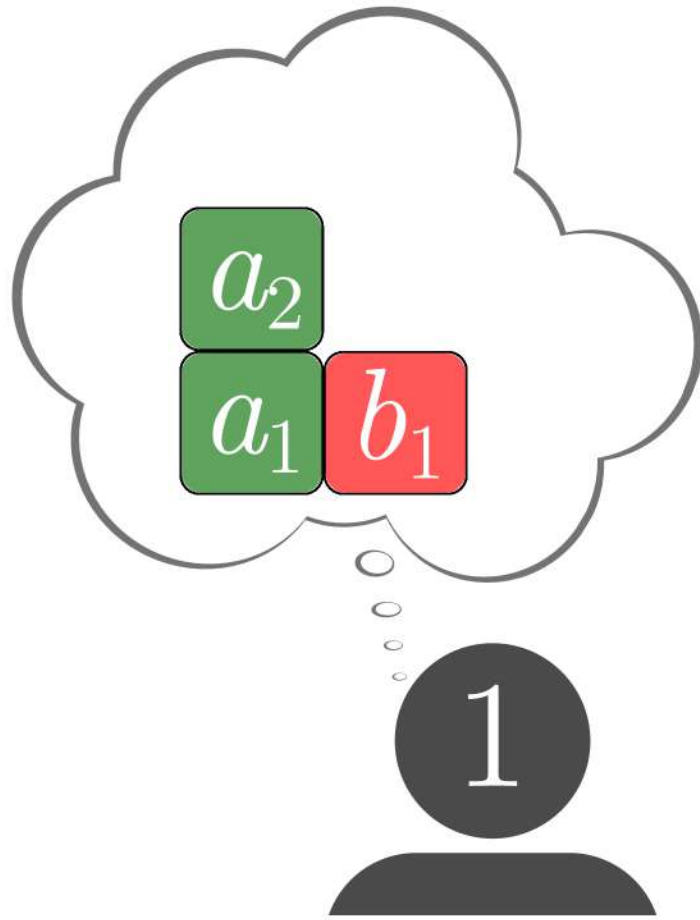
EXAMPLE

$t = 0$



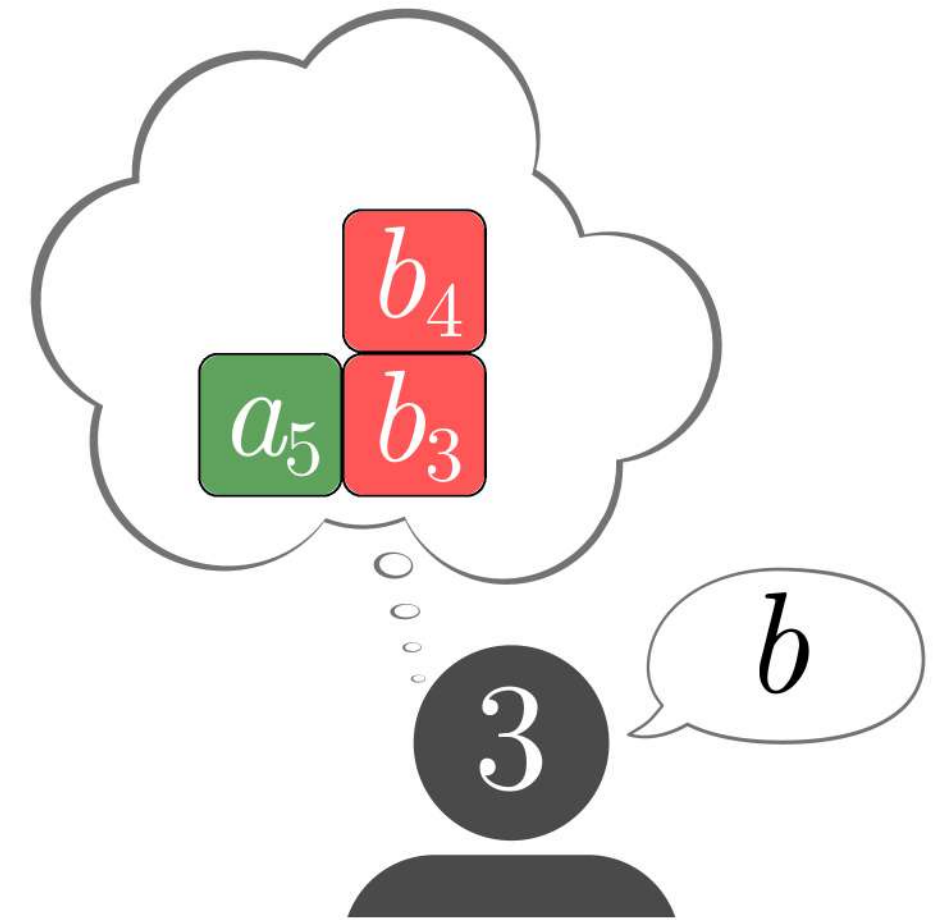
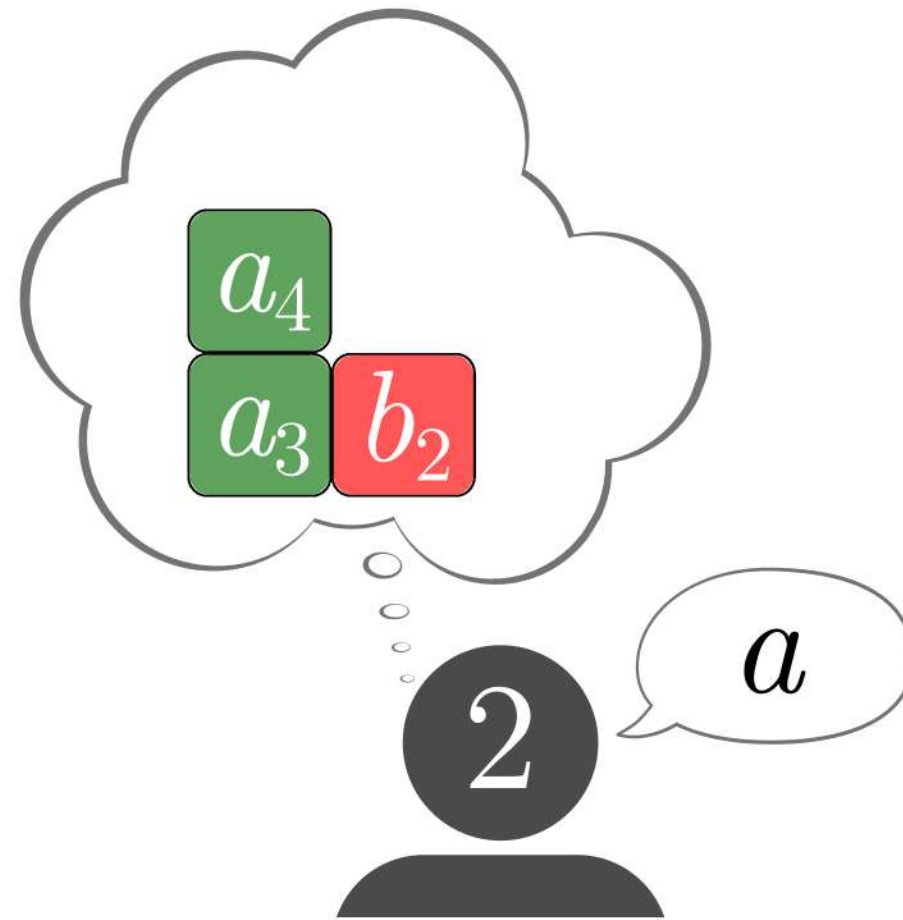
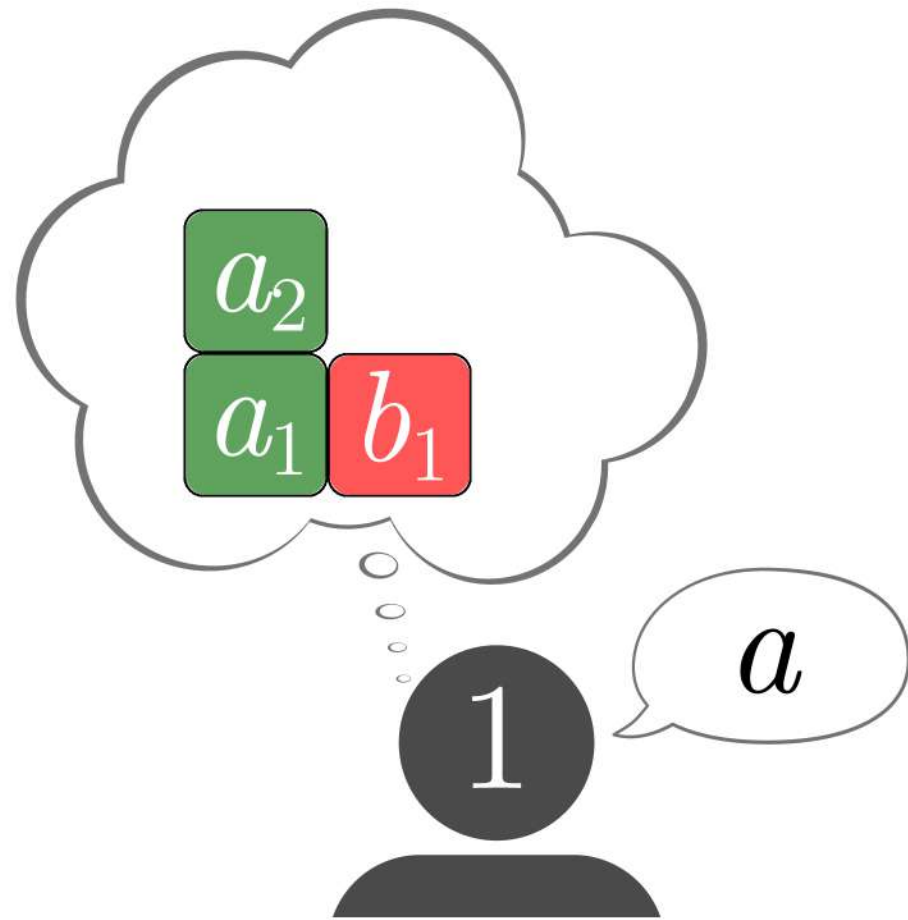
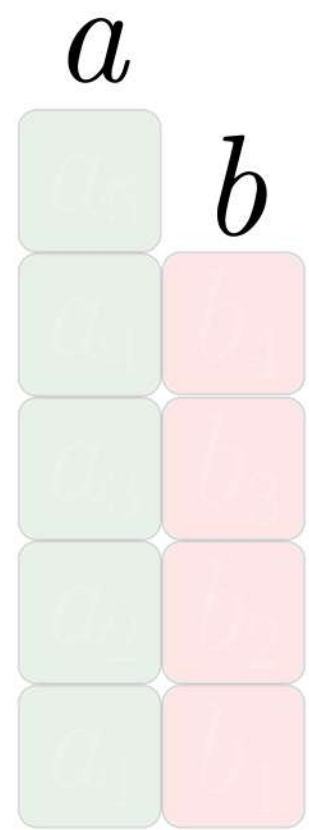
EXAMPLE

$t = 0$



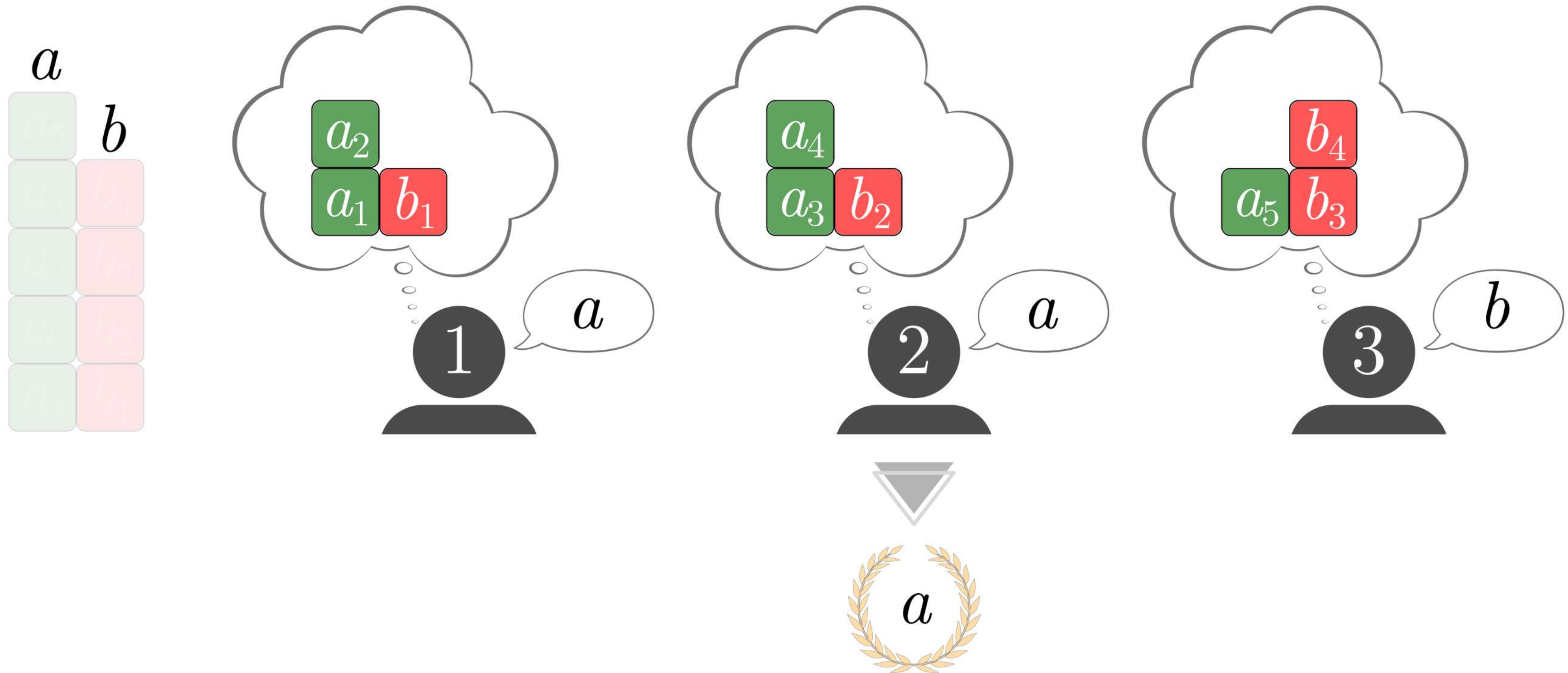
EXAMPLE

$t = 0$



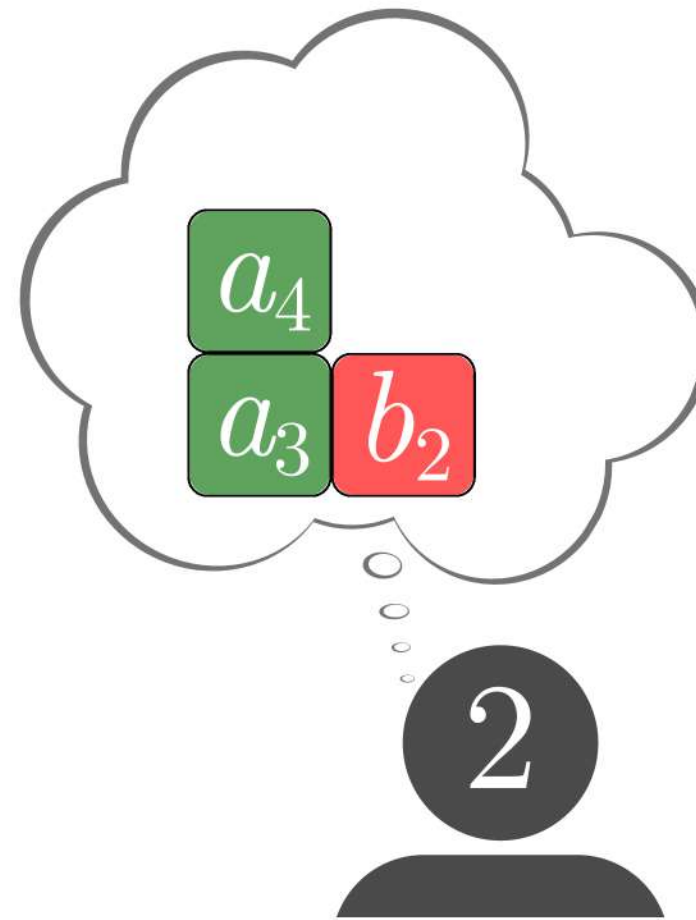
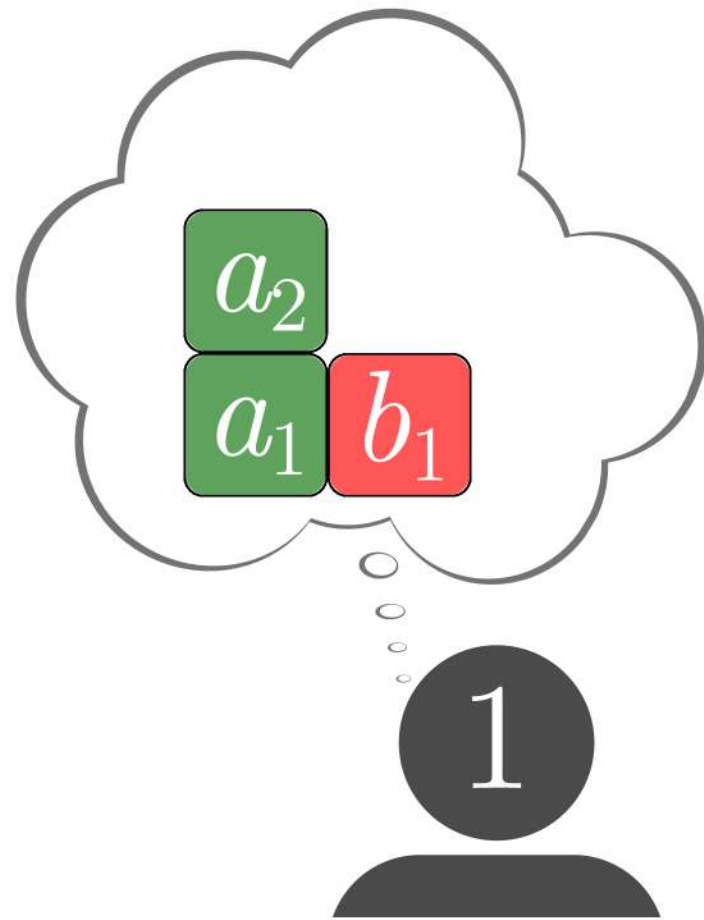
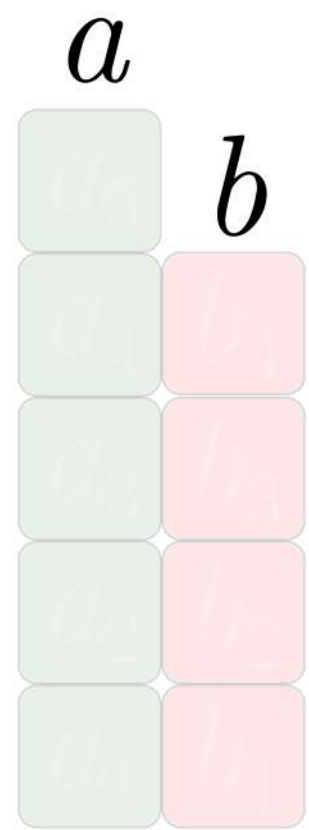
EXAMPLE

$t = 0$



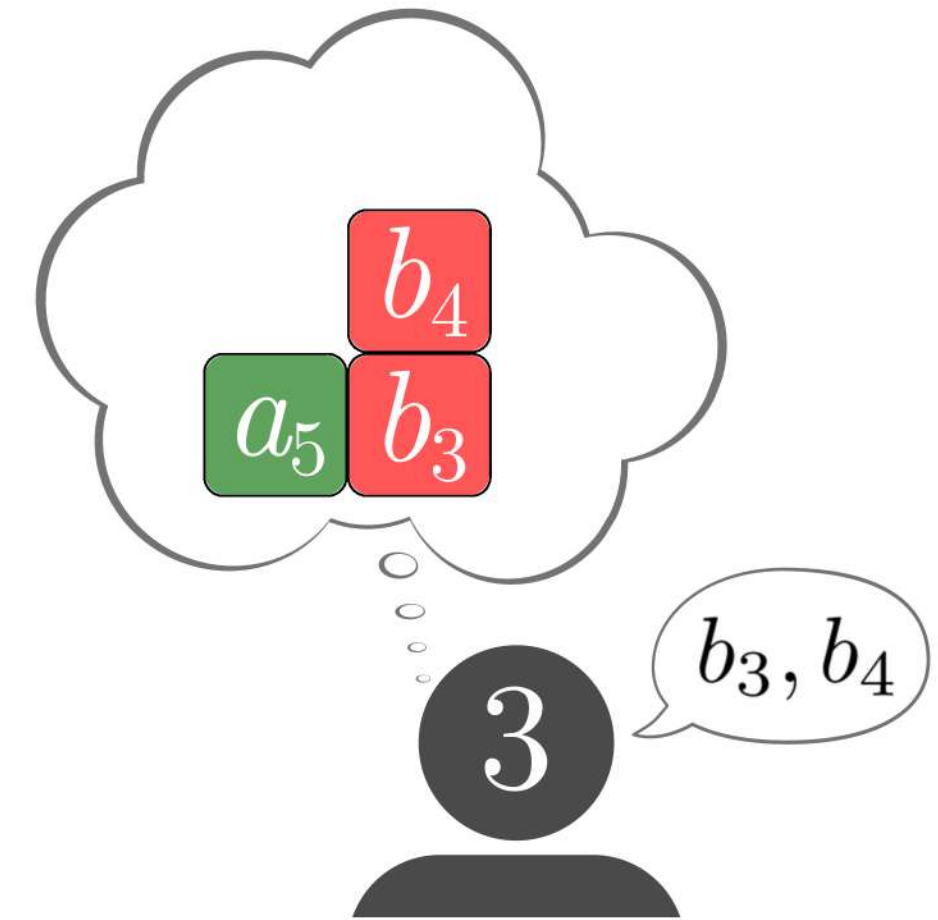
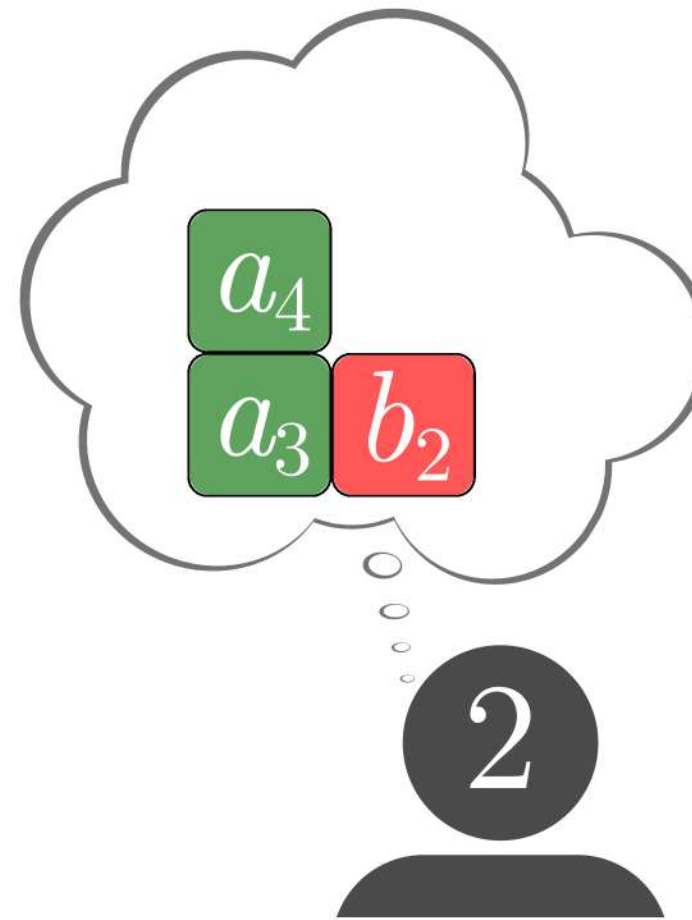
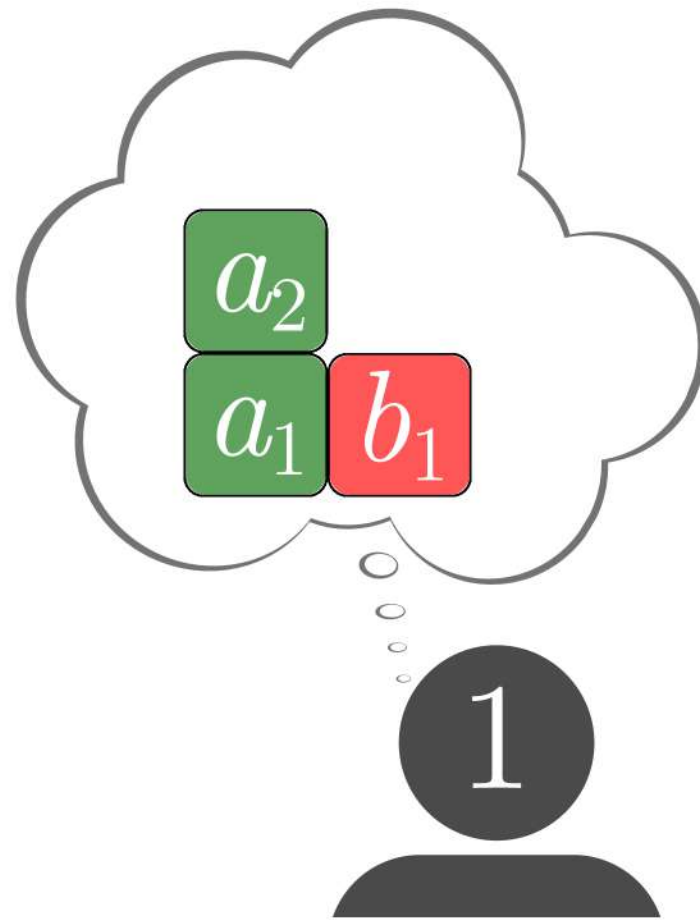
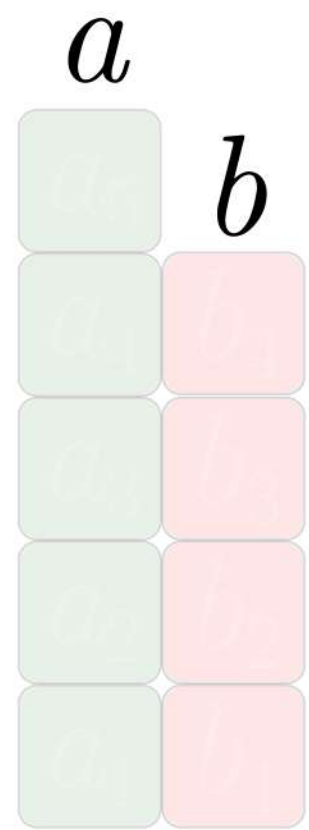
EXAMPLE

$t = 1$



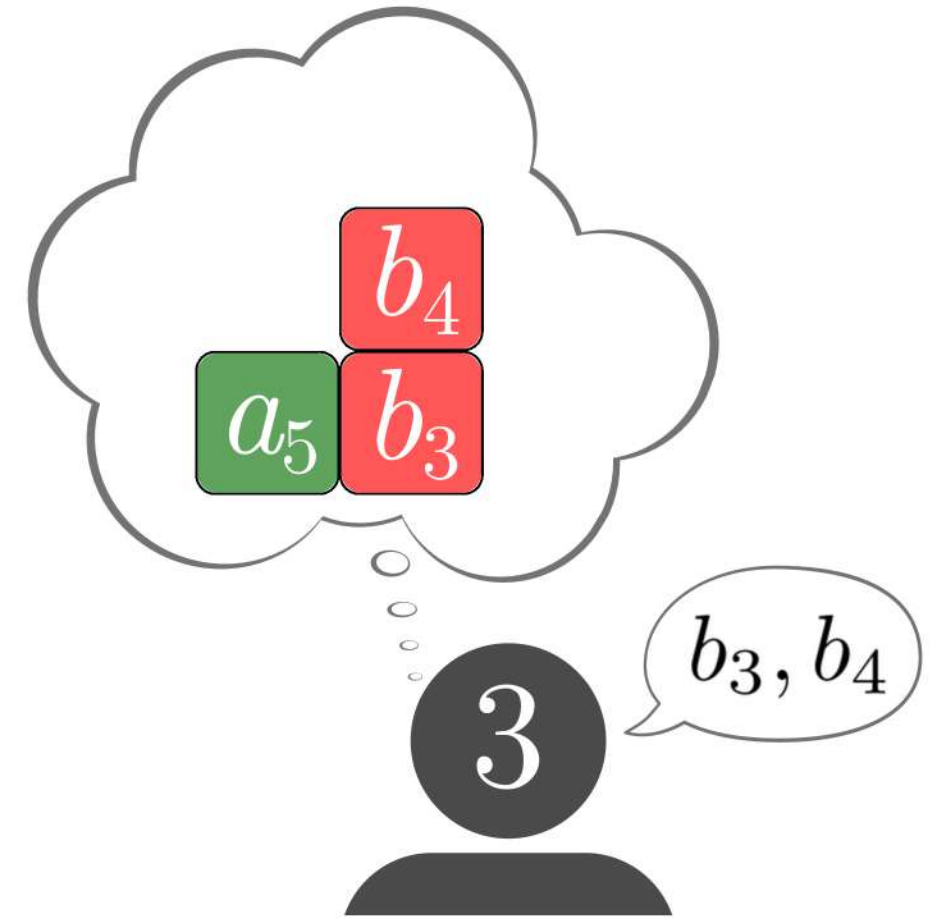
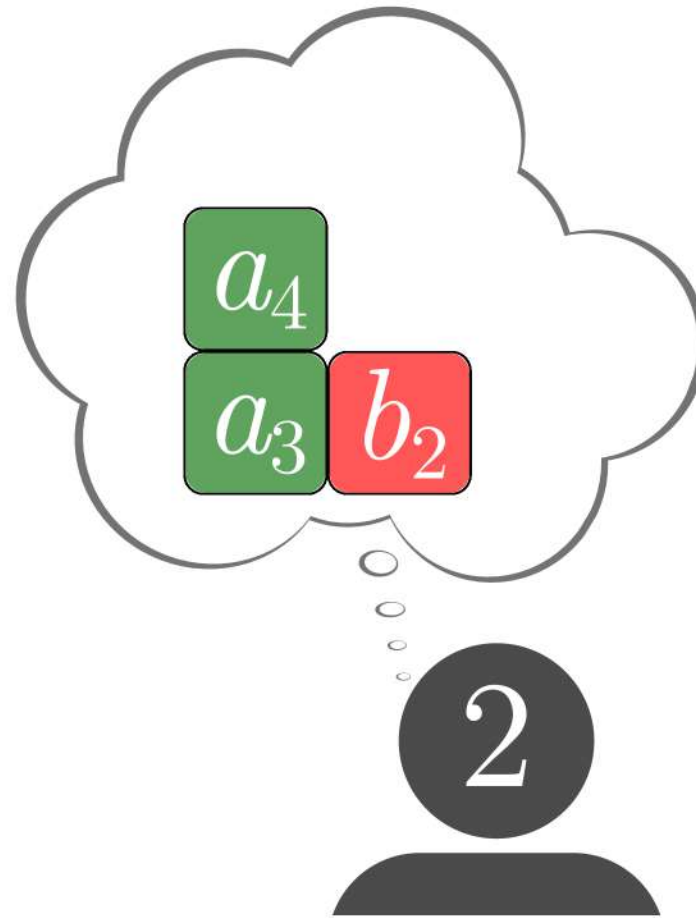
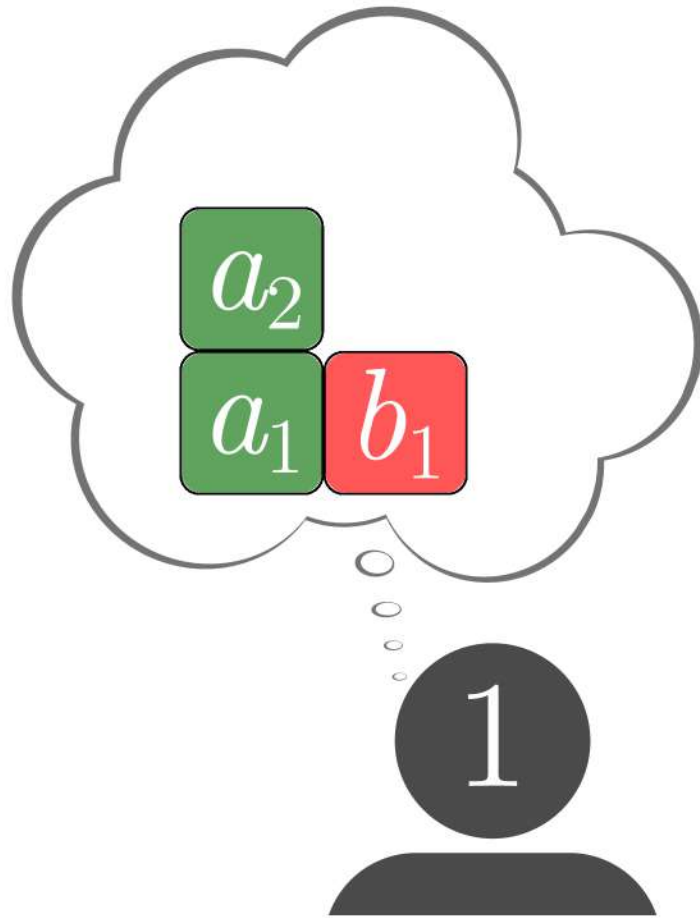
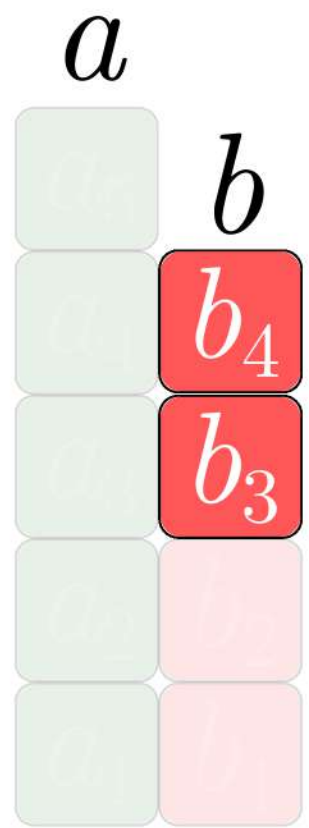
EXAMPLE

$t = 1$



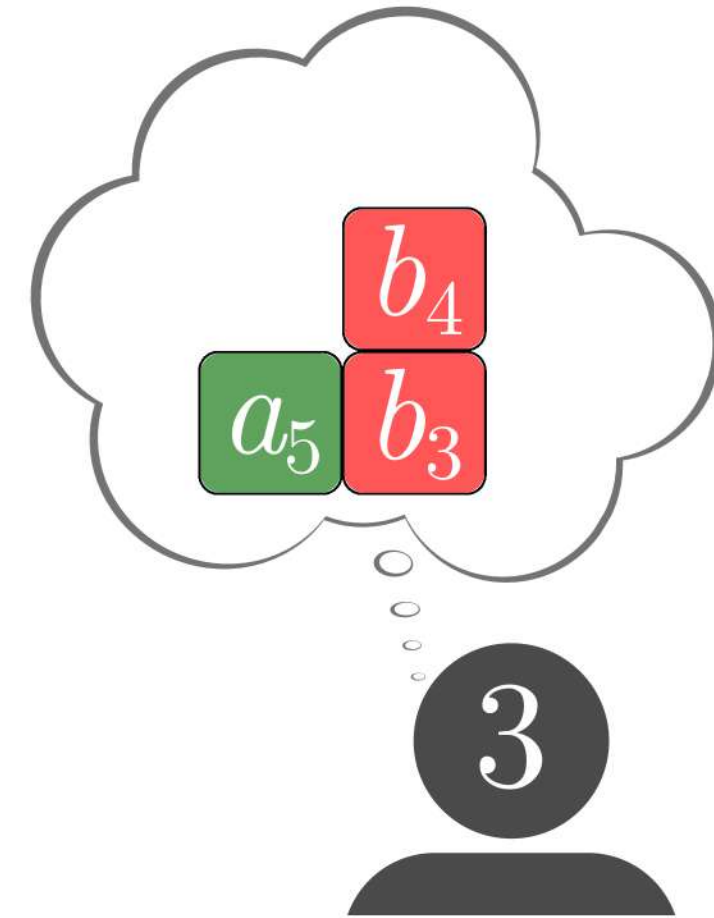
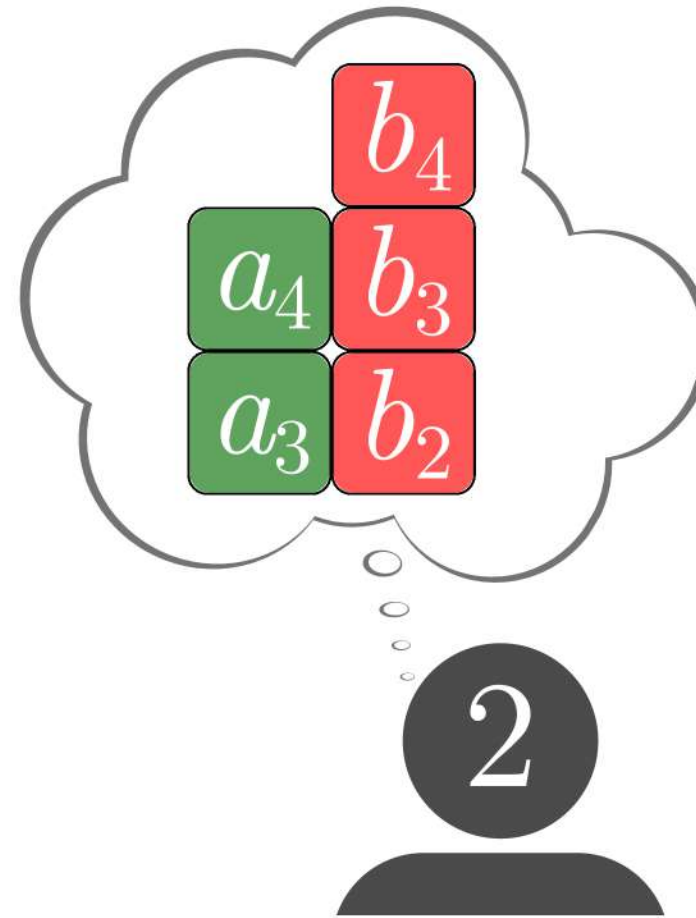
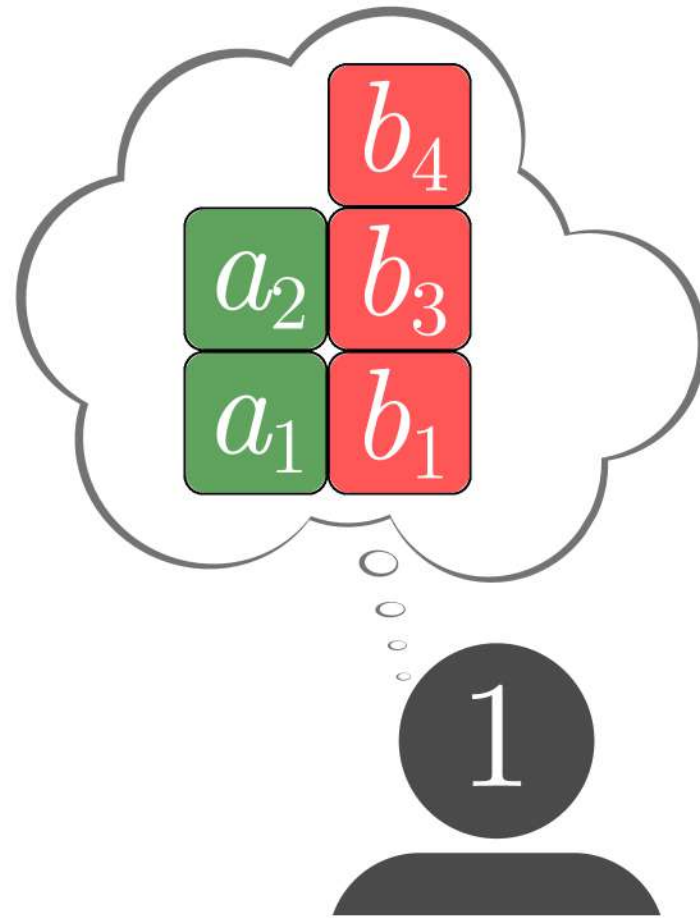
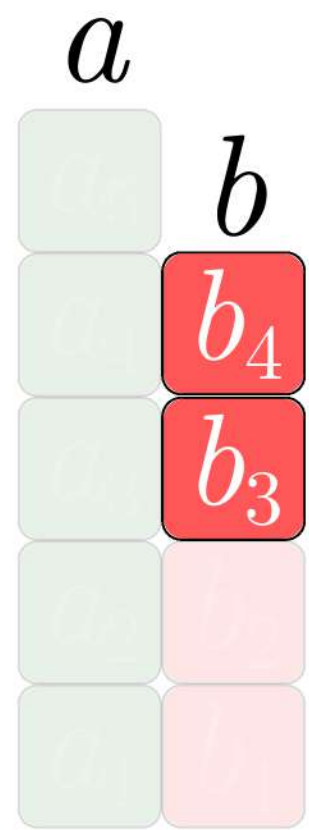
EXAMPLE

$t = 1$



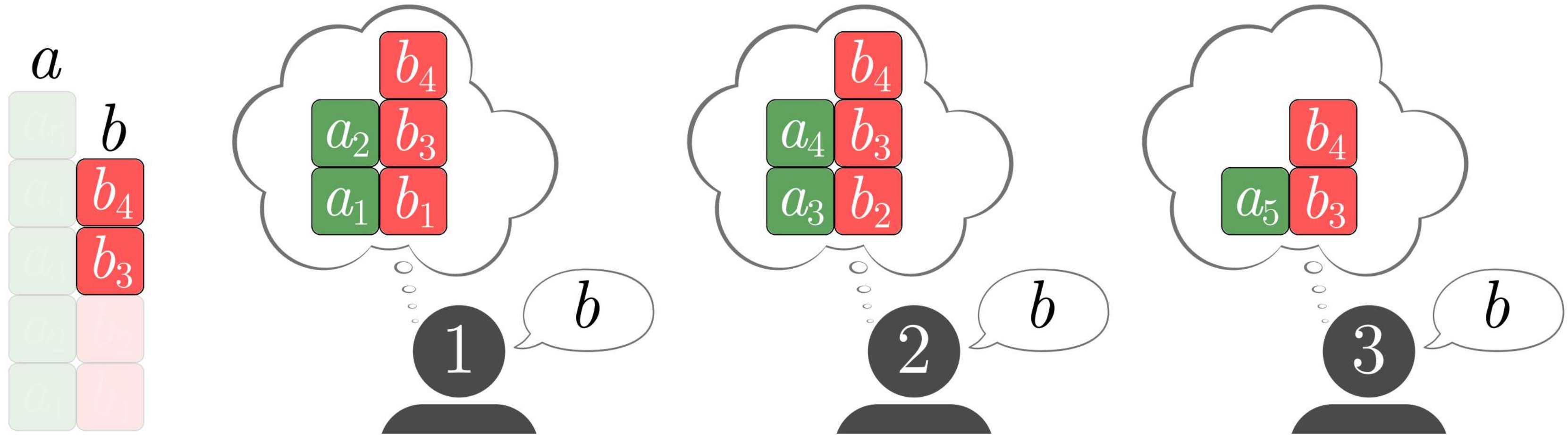
EXAMPLE

$t = 1$



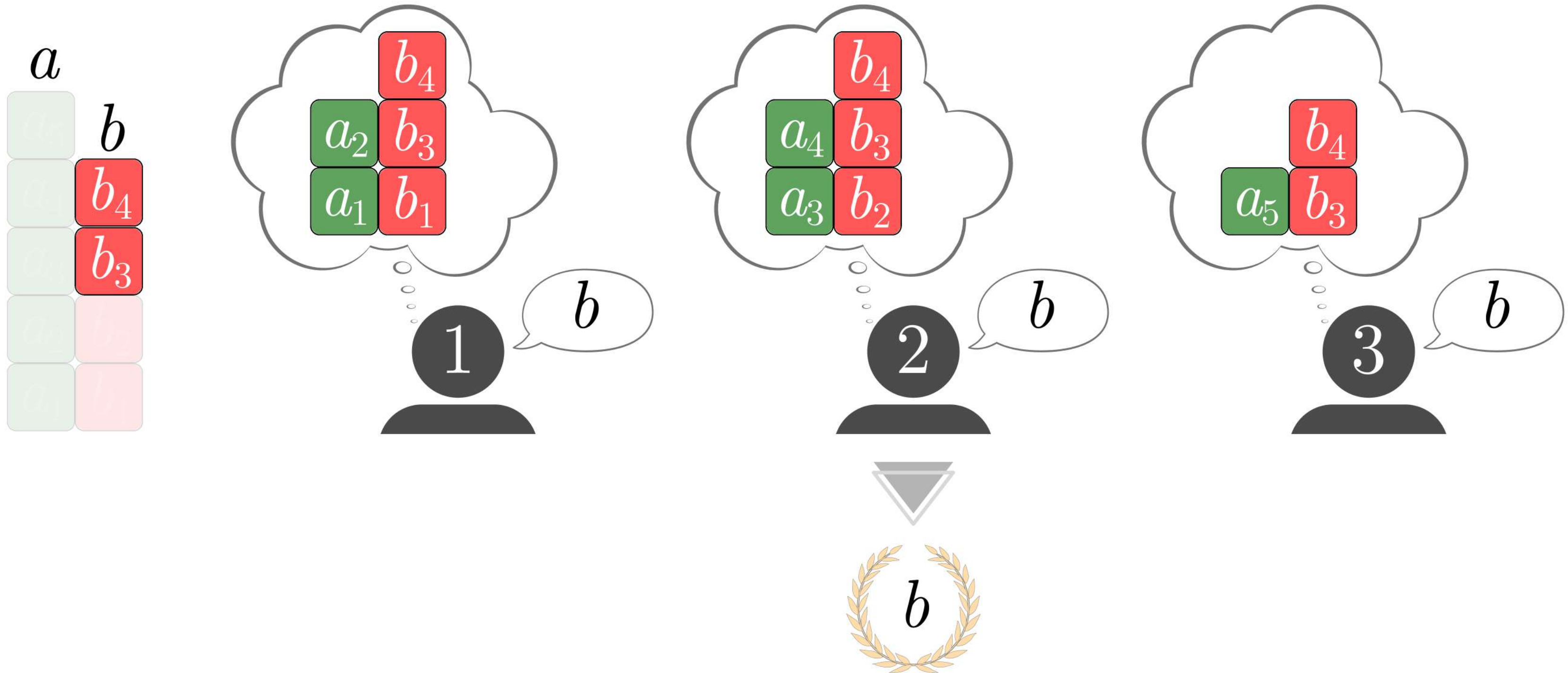
EXAMPLE

$t = 1$



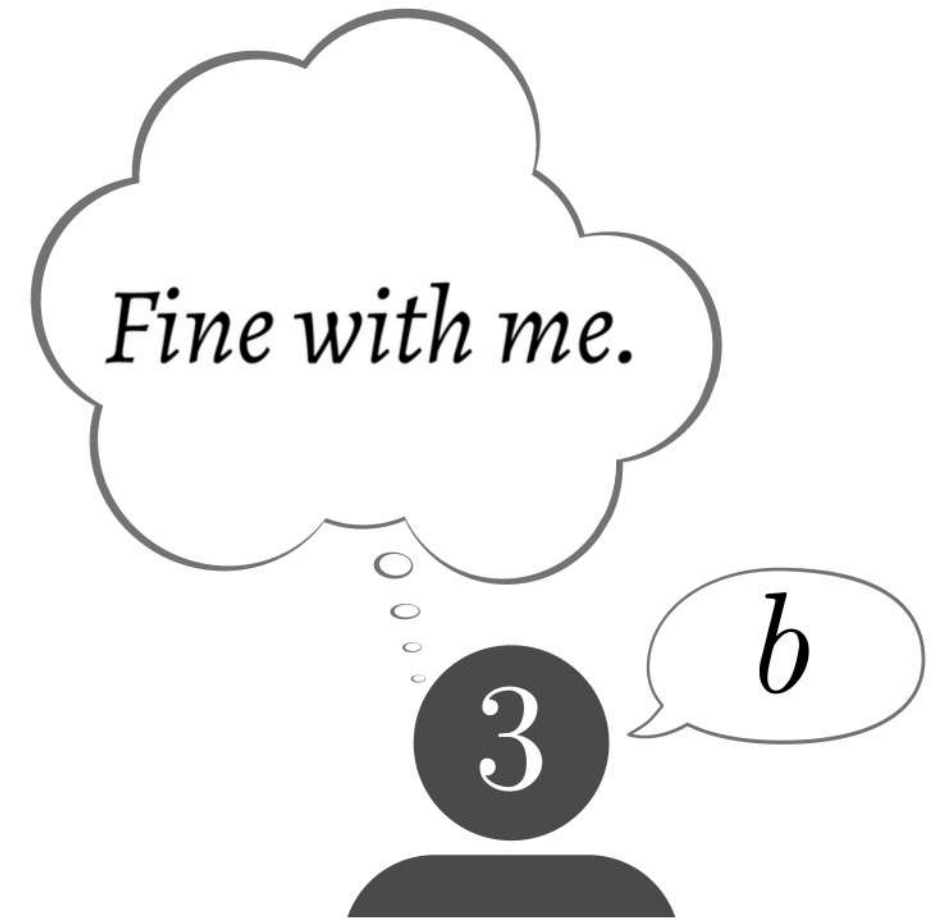
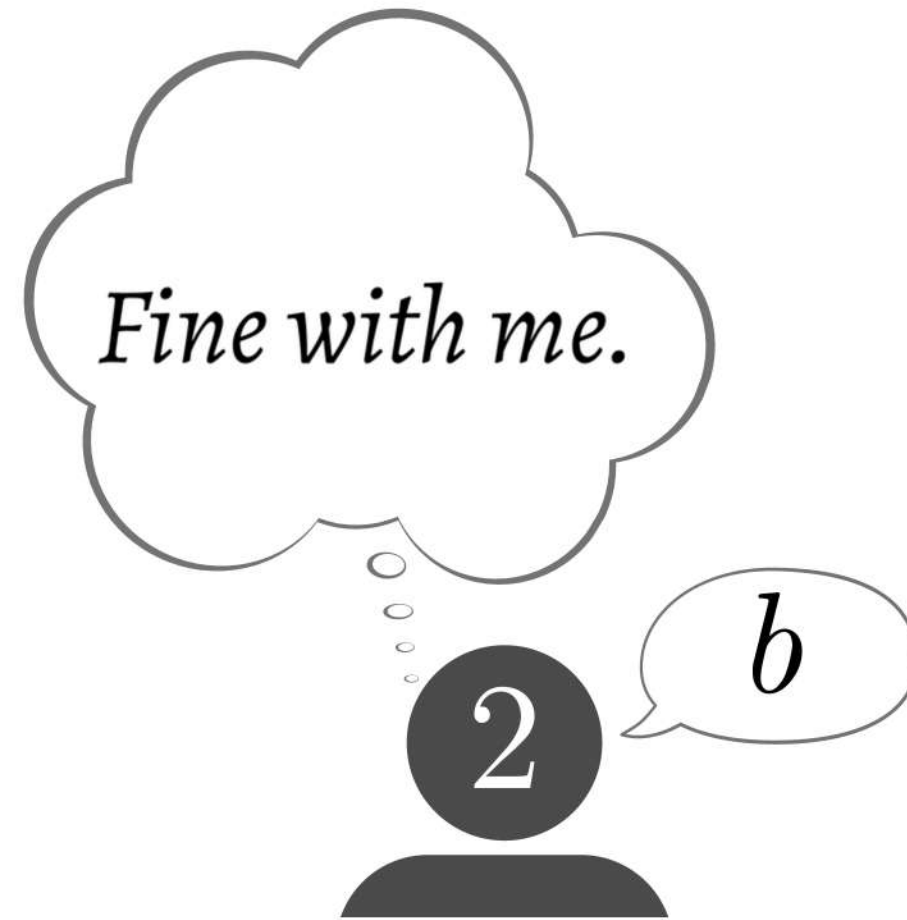
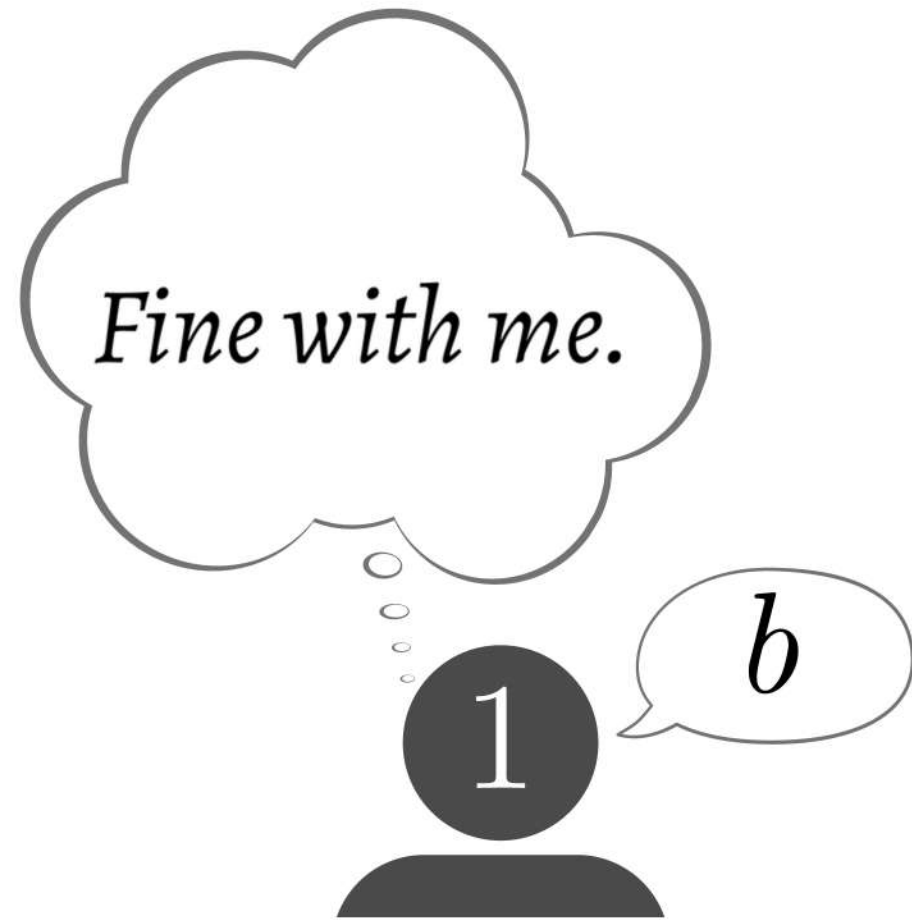
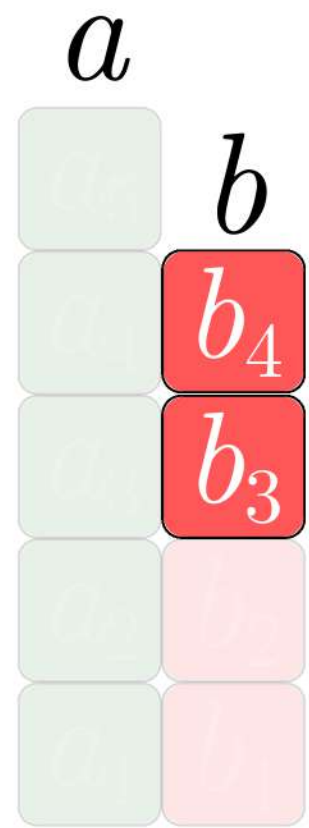
EXAMPLE

$t = 1$



EXAMPLE

$t = 2$



In this example deliberation leads to the wrong answer...

In this example deliberation leads to the wrong answer... but how *likely* is that to happen?

In this example deliberation leads to the wrong answer... but how *likely* is that to happen?

Let's look at snapshots of the deliberation process.

RESULT

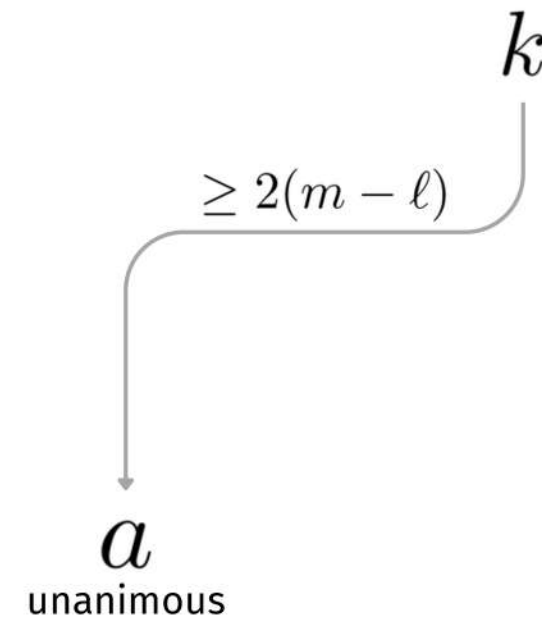
Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it.



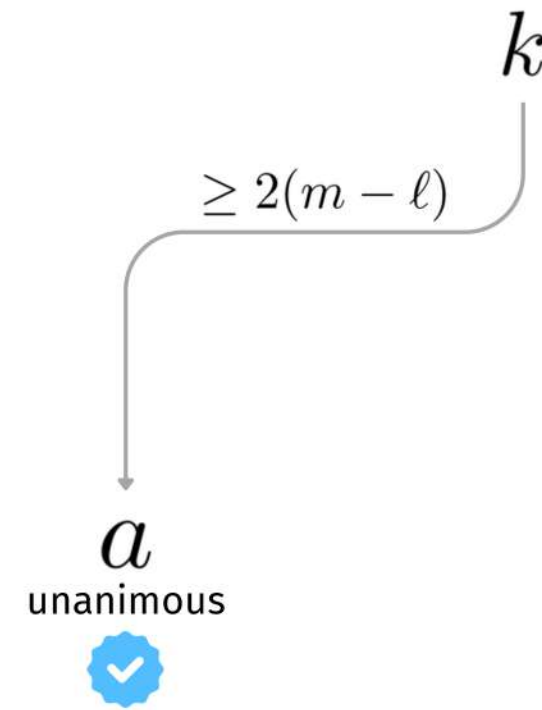
ROUND 0

RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.



ROUND 0

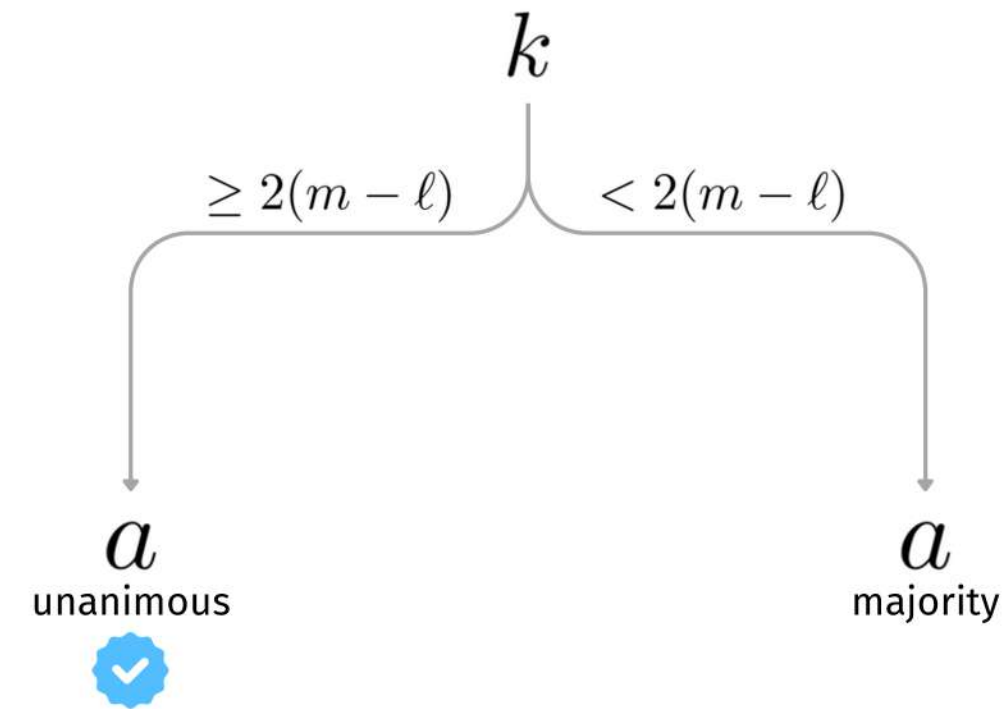
RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.

If $k < 2(m - \ell)$, we get a majority* for a : every agent has a probability $p > 1/2$ to vote for a .



ROUND 0

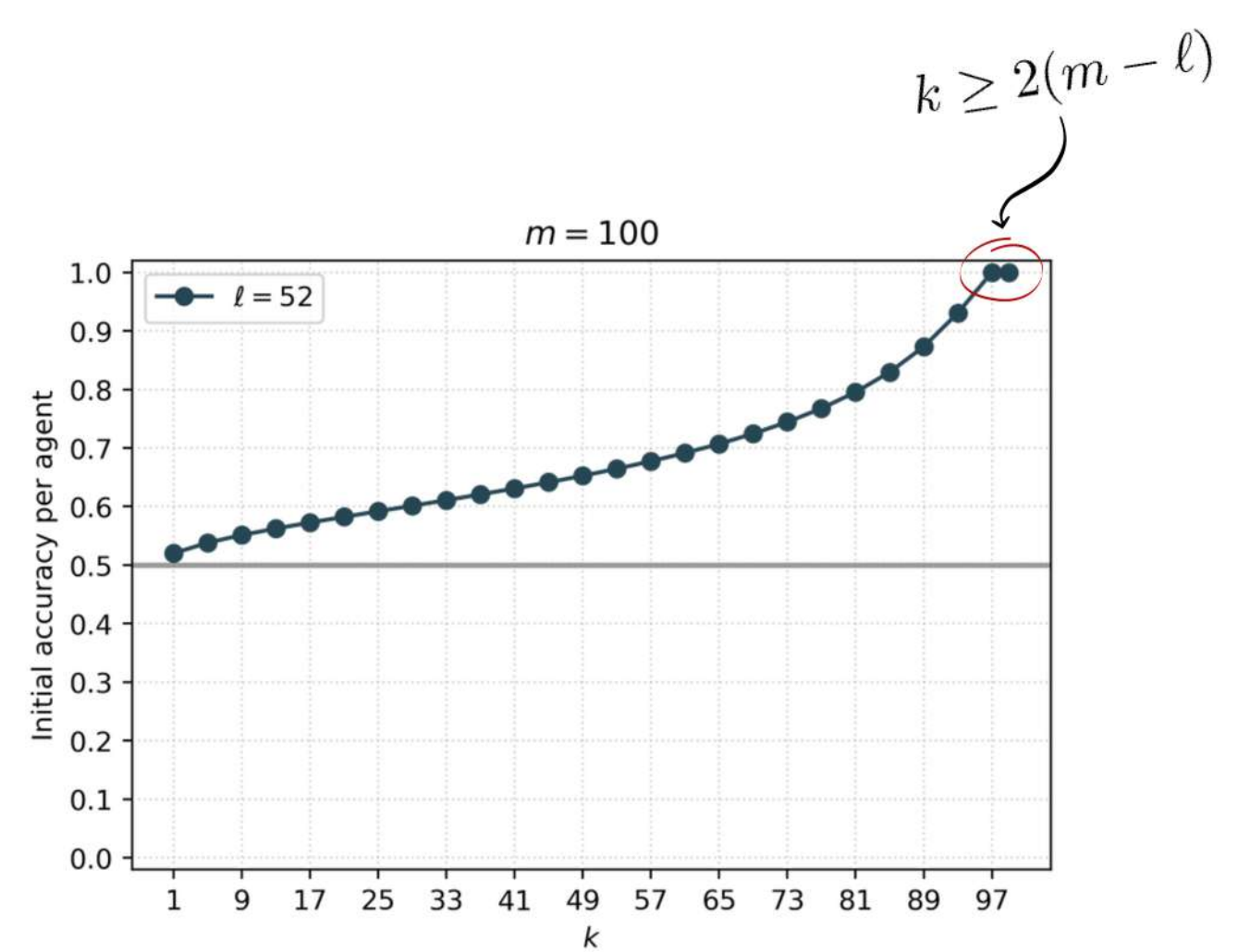
RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.

If $k < 2(m - \ell)$, we get a majority* for a : every agent has a probability $p > 1/2$ to vote for a .



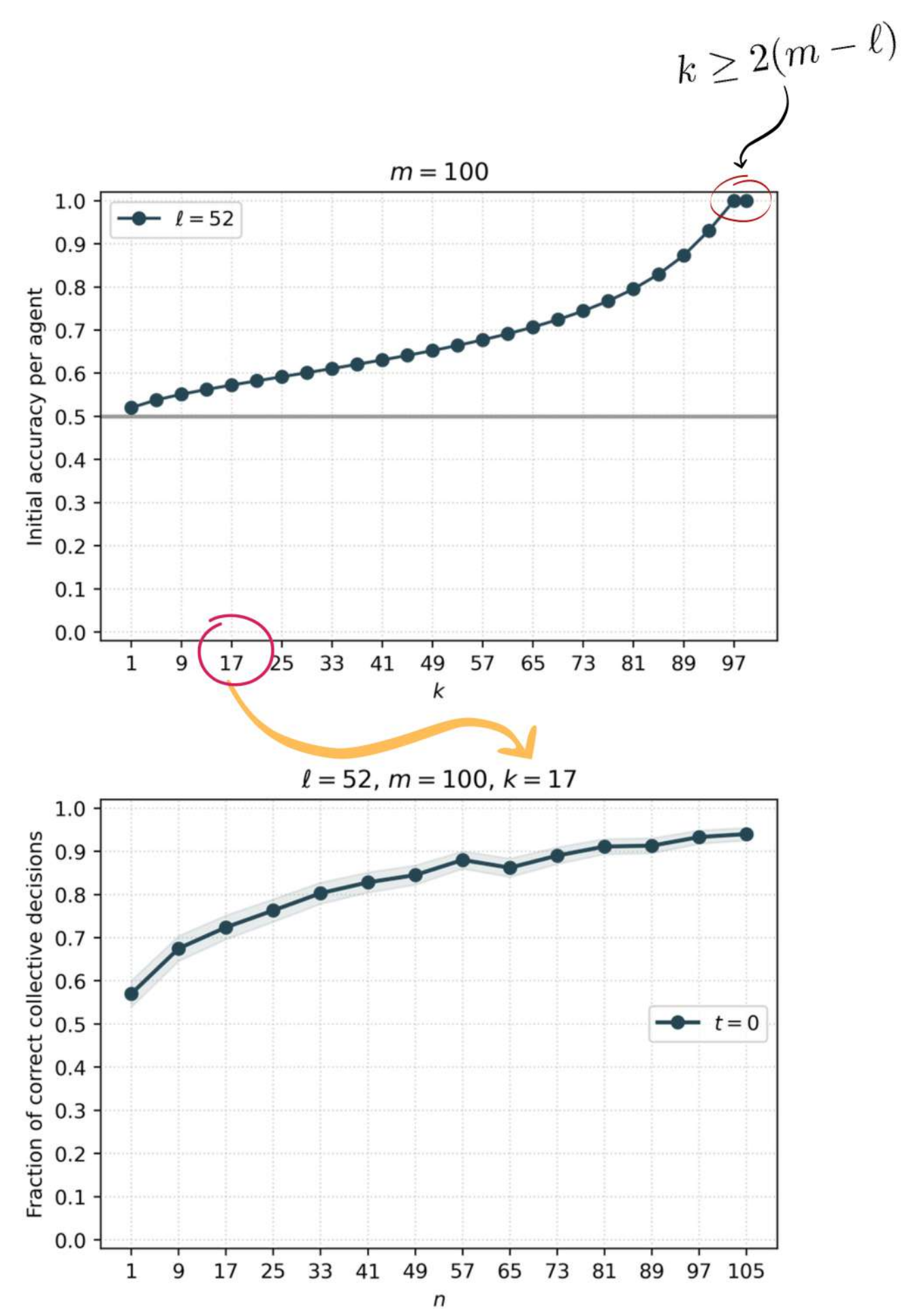
RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.

If $k < 2(m - \ell)$, we get a majority* for a : every agent has a probability $p > 1/2$ to vote for a .



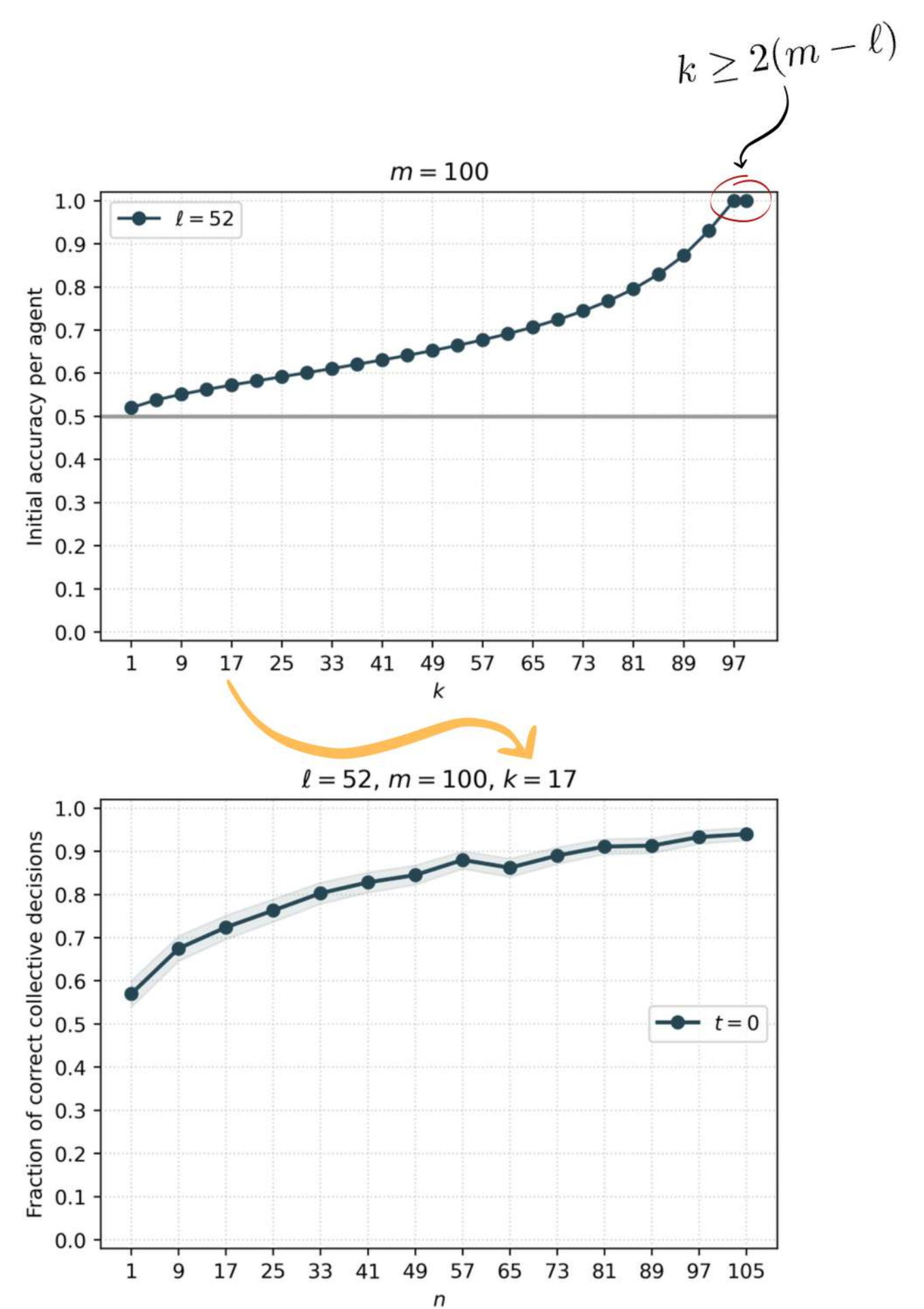
RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.

If $k < 2(m - \ell)$, we get a majority* for a : every agent has a probability $p > 1/2$ to vote for a . This is just the Condorcet Jury Theorem!



RESULT

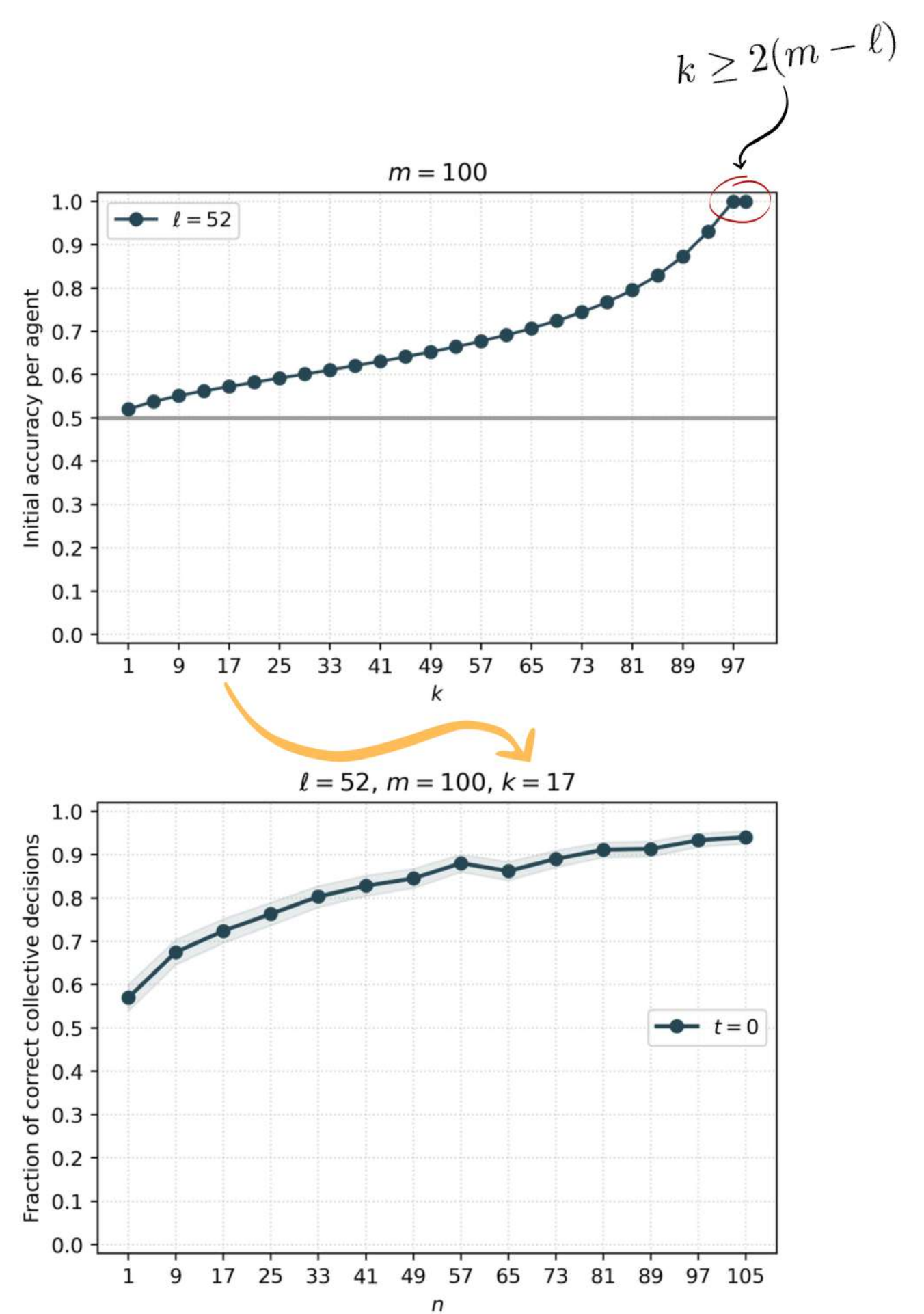
Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 0

If $k \geq 2(m - \ell)$, no one votes for b : not enough total evidence for b to make anyone vote for it. Deliberation terminates here.

If $k < 2(m - \ell)$, we get a majority* for a : every agent has a probability $p > 1/2$ to vote for a . This is just the Condorcet Jury Theorem!

*But not unanimity! $p < 1$, so some agents vote for b .

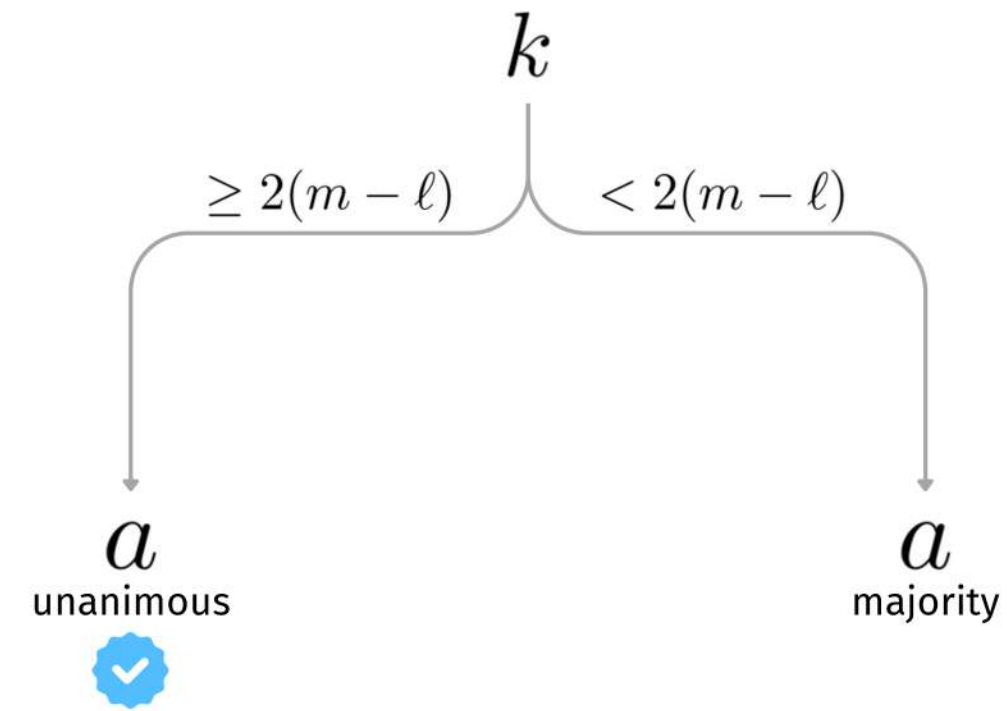


RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b .



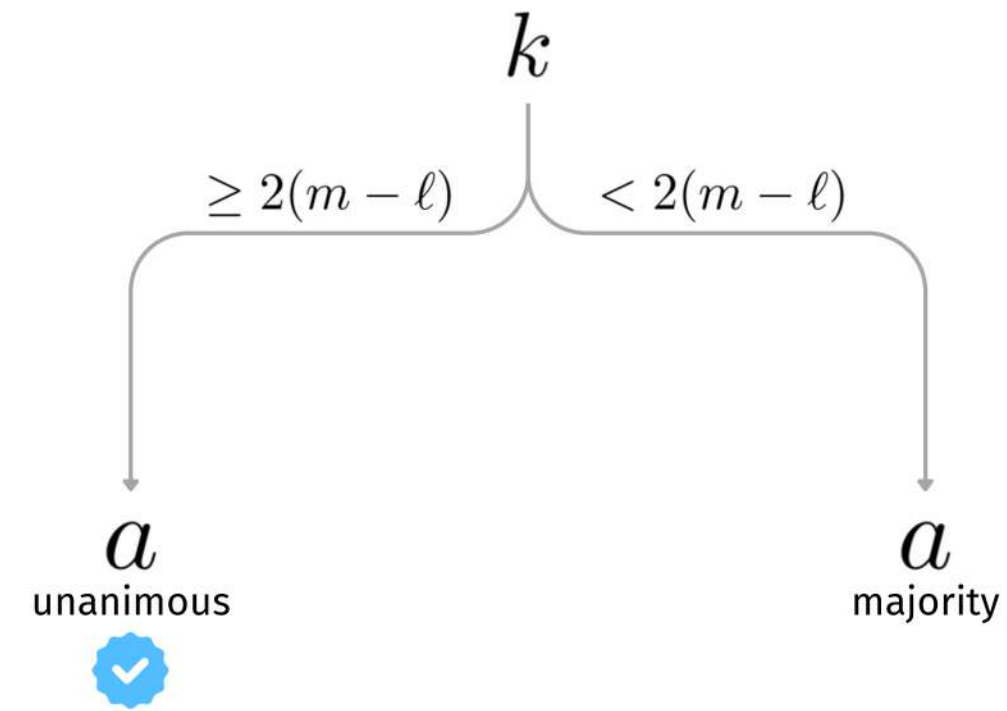
ROUND 0

RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b .
They disclose all their evidence for b .



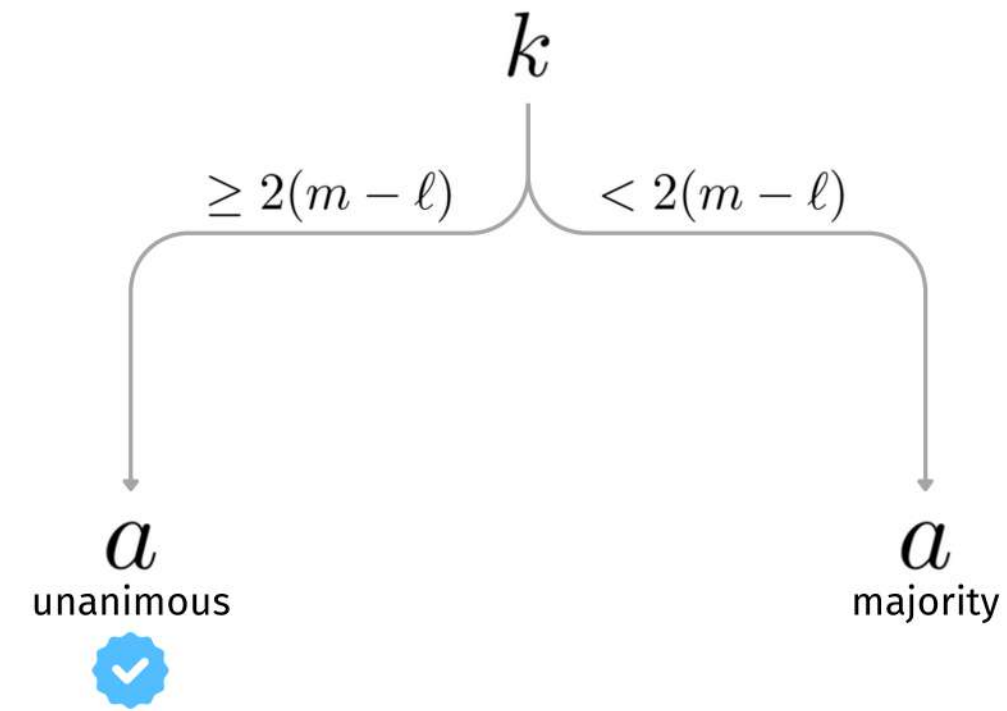
ROUND 0

RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b . They disclose all their evidence for b . As $n \rightarrow \infty$, this is *all* (!) existing evidence for b .



ROUND 0

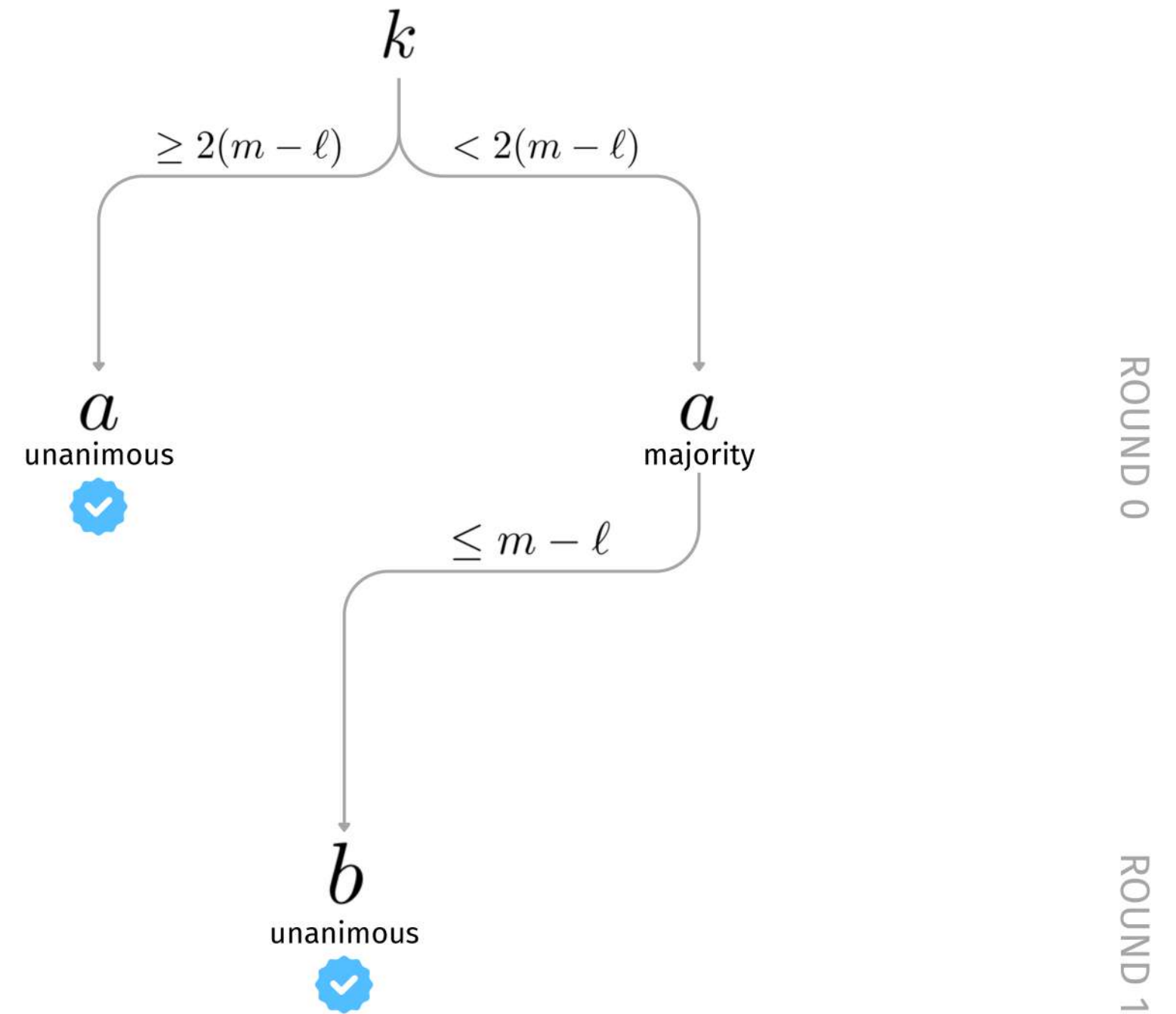
RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b . They disclose all their evidence for b . As $n \rightarrow \infty$, this is *all* (!) existing evidence for b .

If $k \leq m - \ell$, then no a -supporter is left. Deliberation terminates here.



RESULT

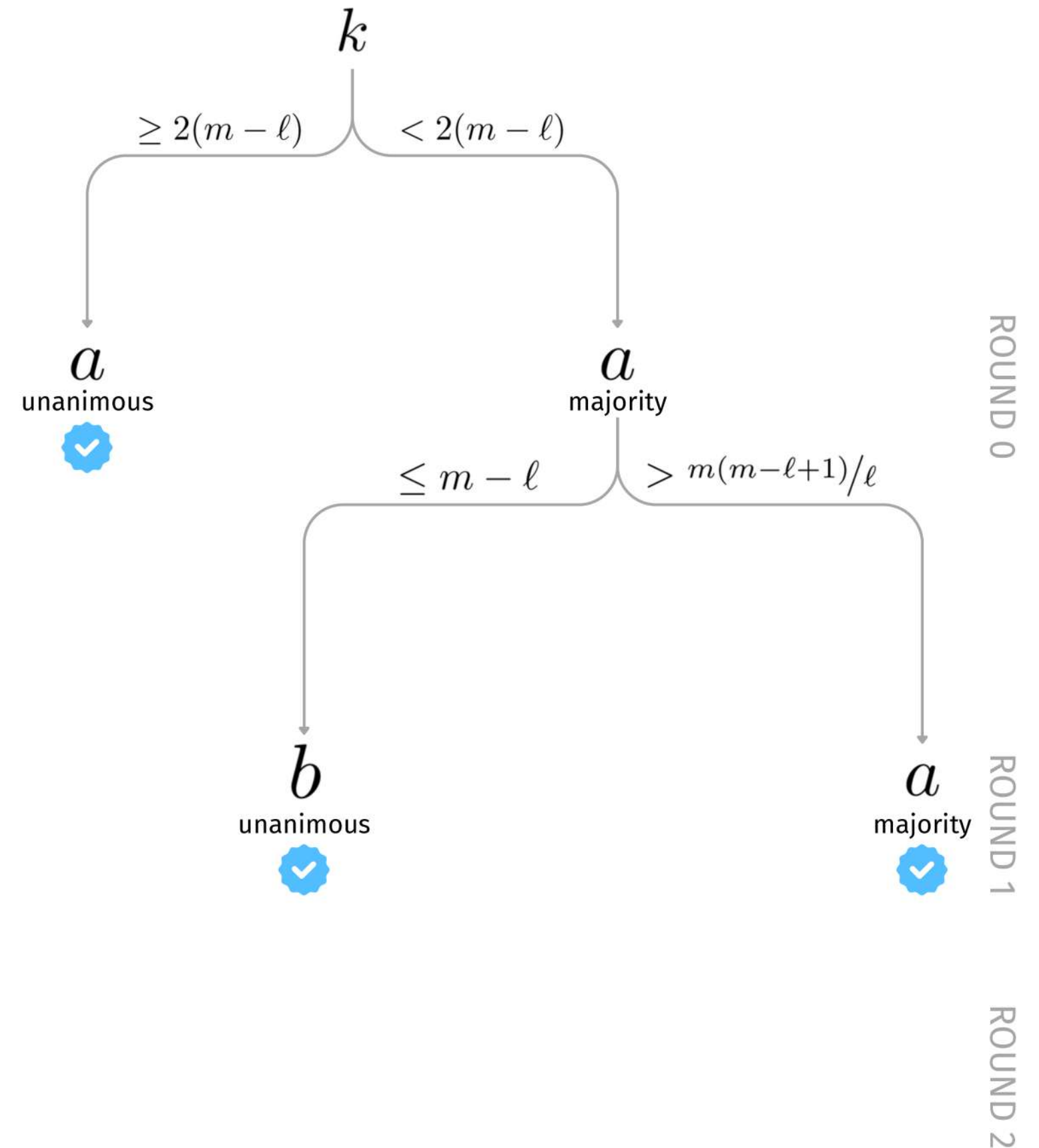
Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b . They disclose all their evidence for b . As $n \rightarrow \infty$, this is *all* (!) existing evidence for b .

If $k \leq m - \ell$, then no a -supporter is left. Deliberation terminates here.

If $m(m - \ell + 1)/\ell < k < 2(m - \ell)$, the majority for a survives. There's no more evidence for b to be learned, so deliberation terminates here.



RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

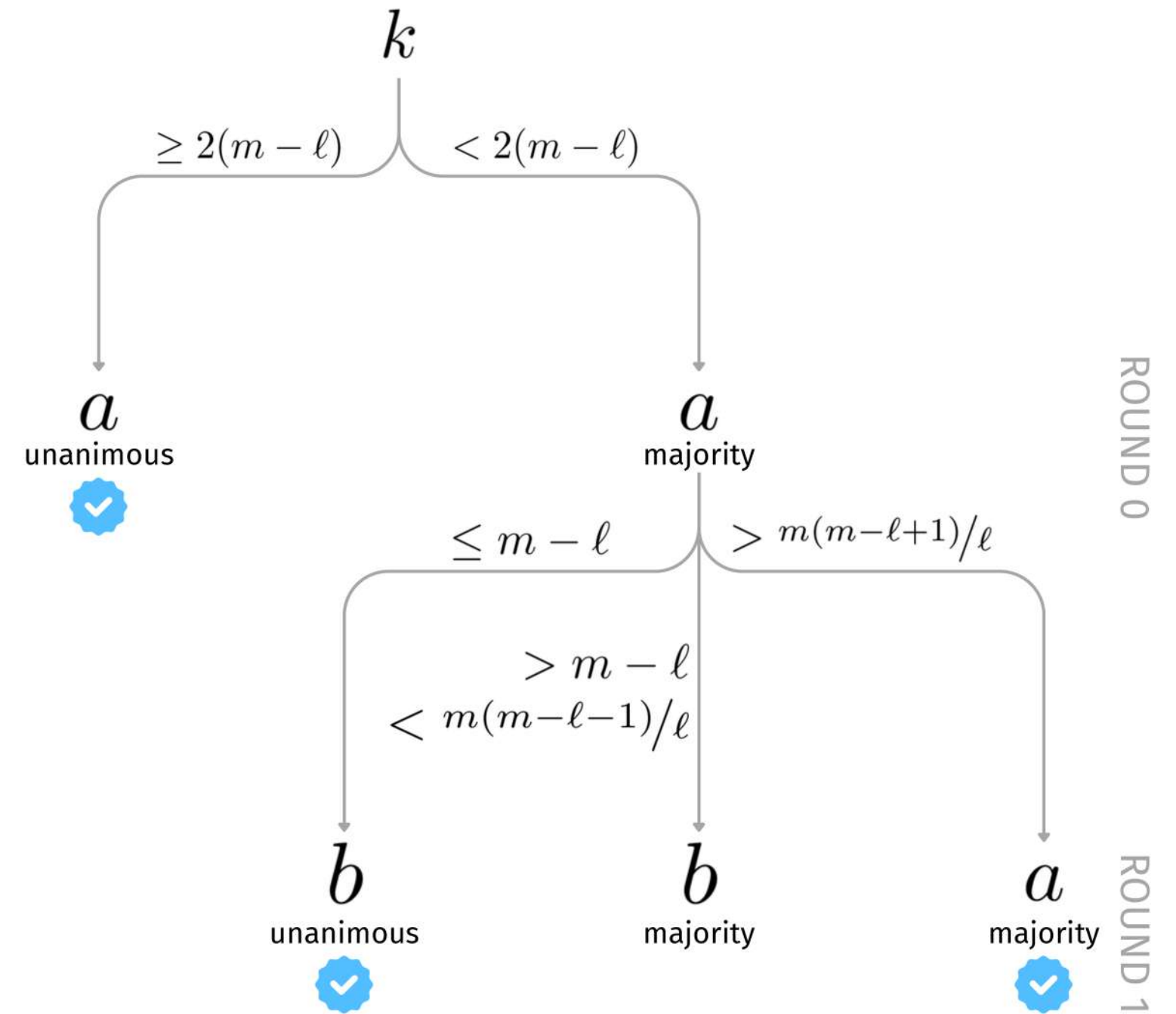
ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b . They disclose all their evidence for b . As $n \rightarrow \infty$, this is *all* (!) existing evidence for b .

If $k \leq m - \ell$, then no a -supporter is left. Deliberation terminates here.

If $m(m - \ell + 1)/\ell < k < 2(m - \ell)$, the majority for a survives. There's no more evidence for b to be learned, so deliberation terminates here.

If $m - \ell < k < m(m - \ell - 1)/\ell$, we get a majority for b .



RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

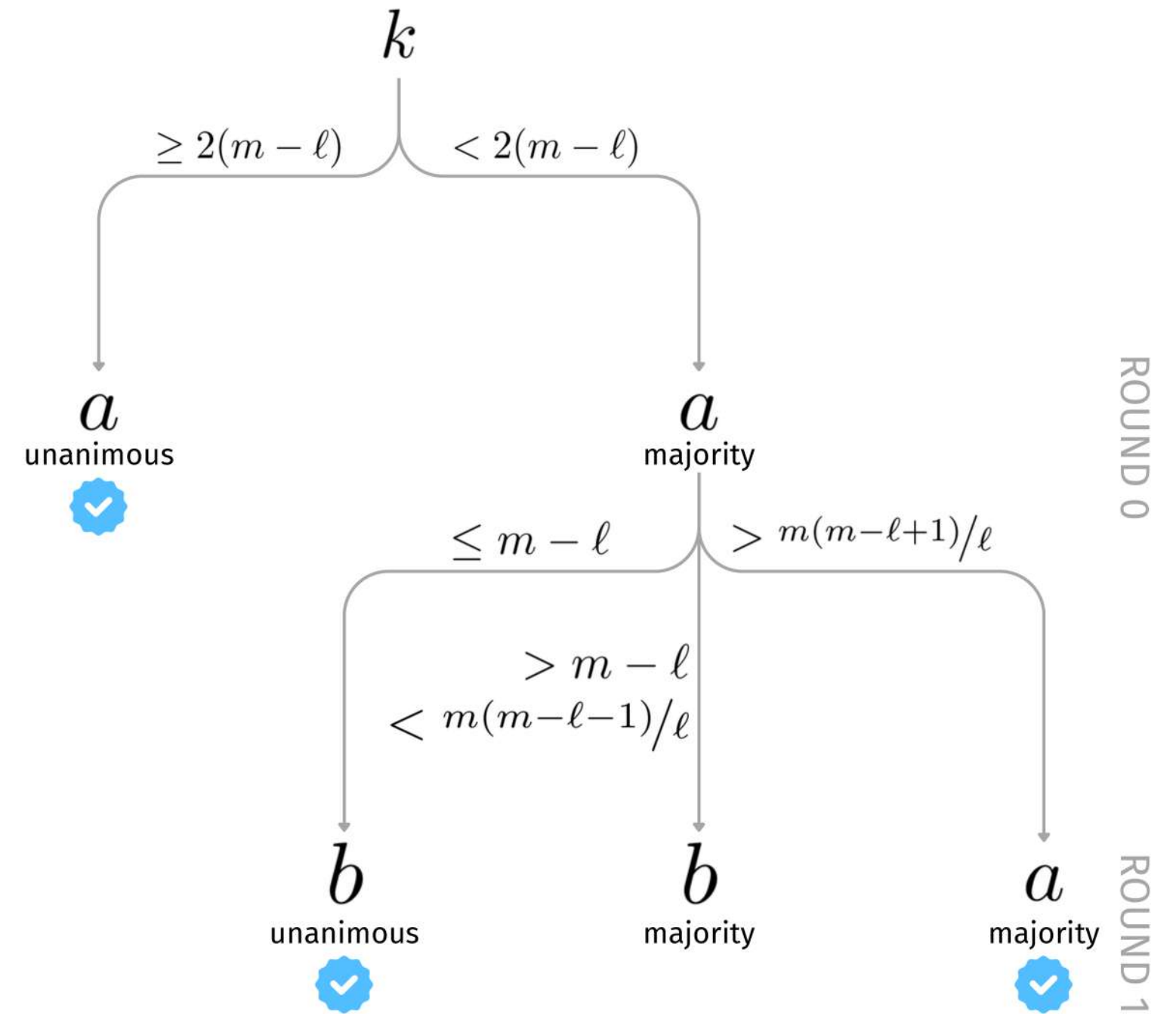
ROUND 1

There's some ($\approx (1 - p)n$) agents voting for b . They disclose all their evidence for b . As $n \rightarrow \infty$, this is *all* (!) existing evidence for b .

If $k \leq m - \ell$, then no a -supporter is left. Deliberation terminates here.

If $m(m - \ell + 1)/\ell < k < 2(m - \ell)$, the majority for a survives. There's no more evidence for b to be learned, so deliberation terminates here.

If $m - \ell < k < m(m - \ell - 1)/\ell$, we get a majority for b . But not unanimity!...

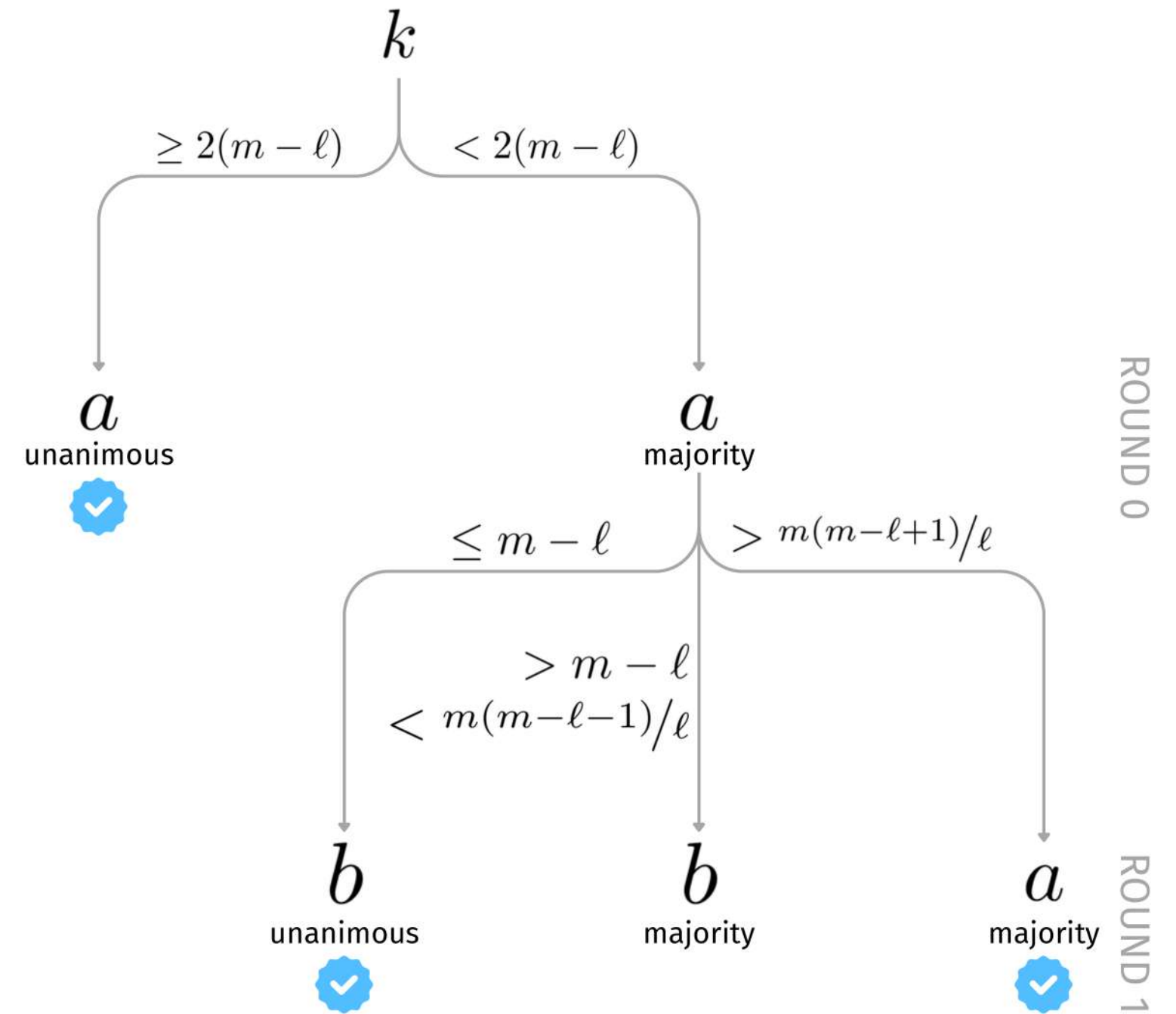


RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 2

If $m - \ell < k < m(m - \ell - 1)/\ell$, there is still a fraction of a -supporters left.

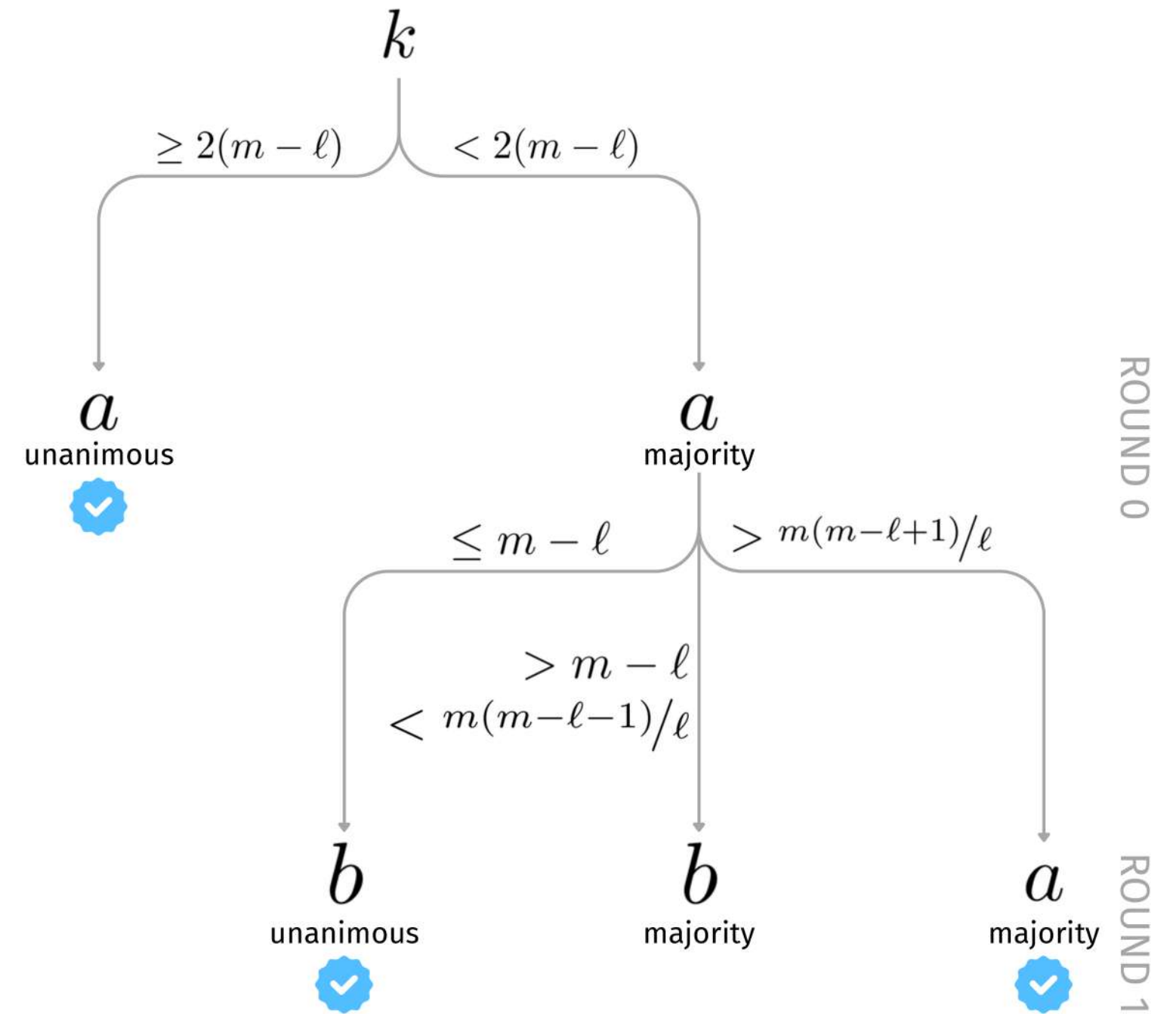


RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

ROUND 2

If $m - \ell < k < m(m - \ell - 1)/\ell$, there is still a fraction of a -supporters left. As $n \rightarrow \infty$, they hold all evidence for a .

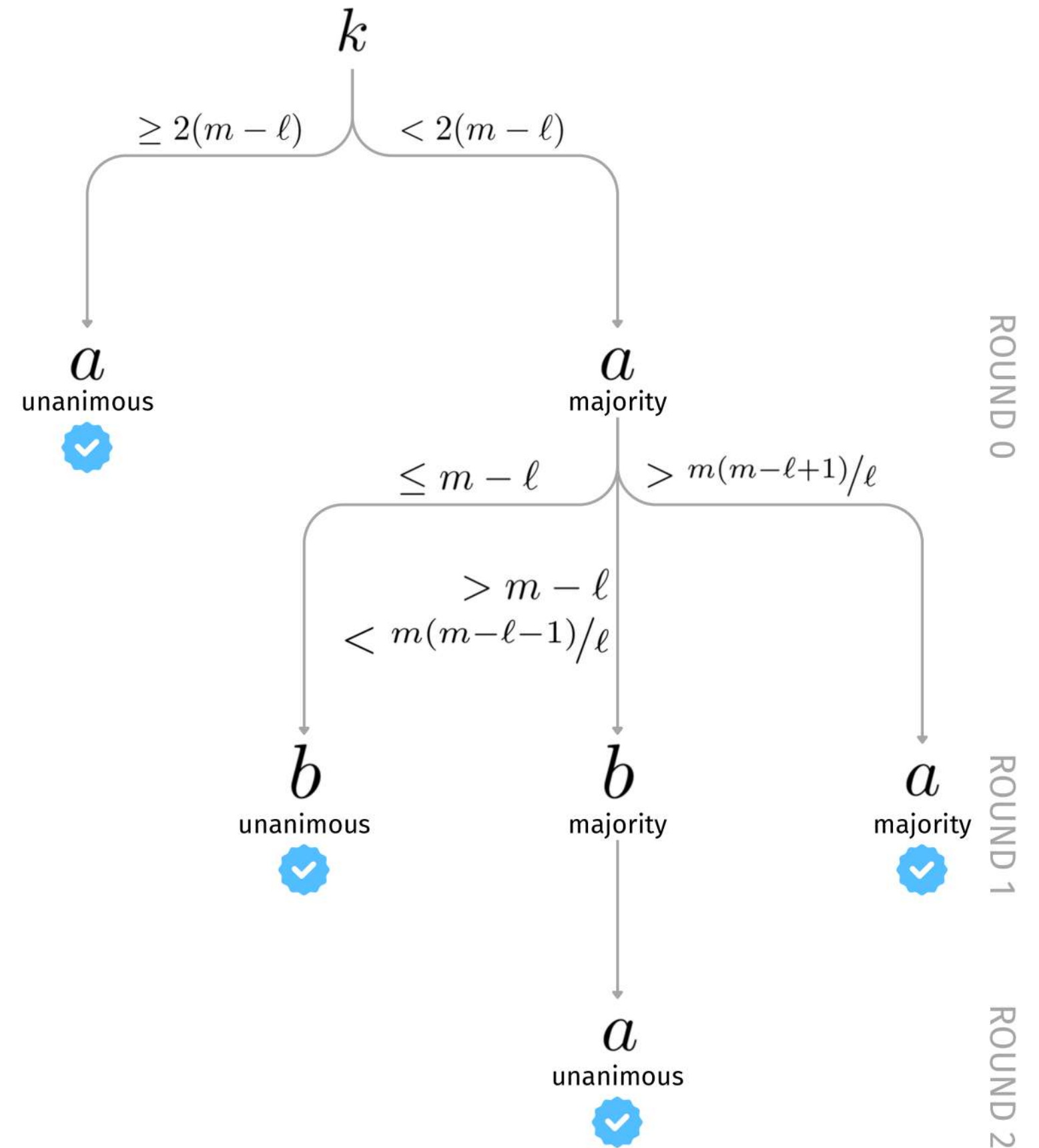


RESULT

Take odd k and $\ell/m > 1/2$. As $n \rightarrow \infty$, we have:

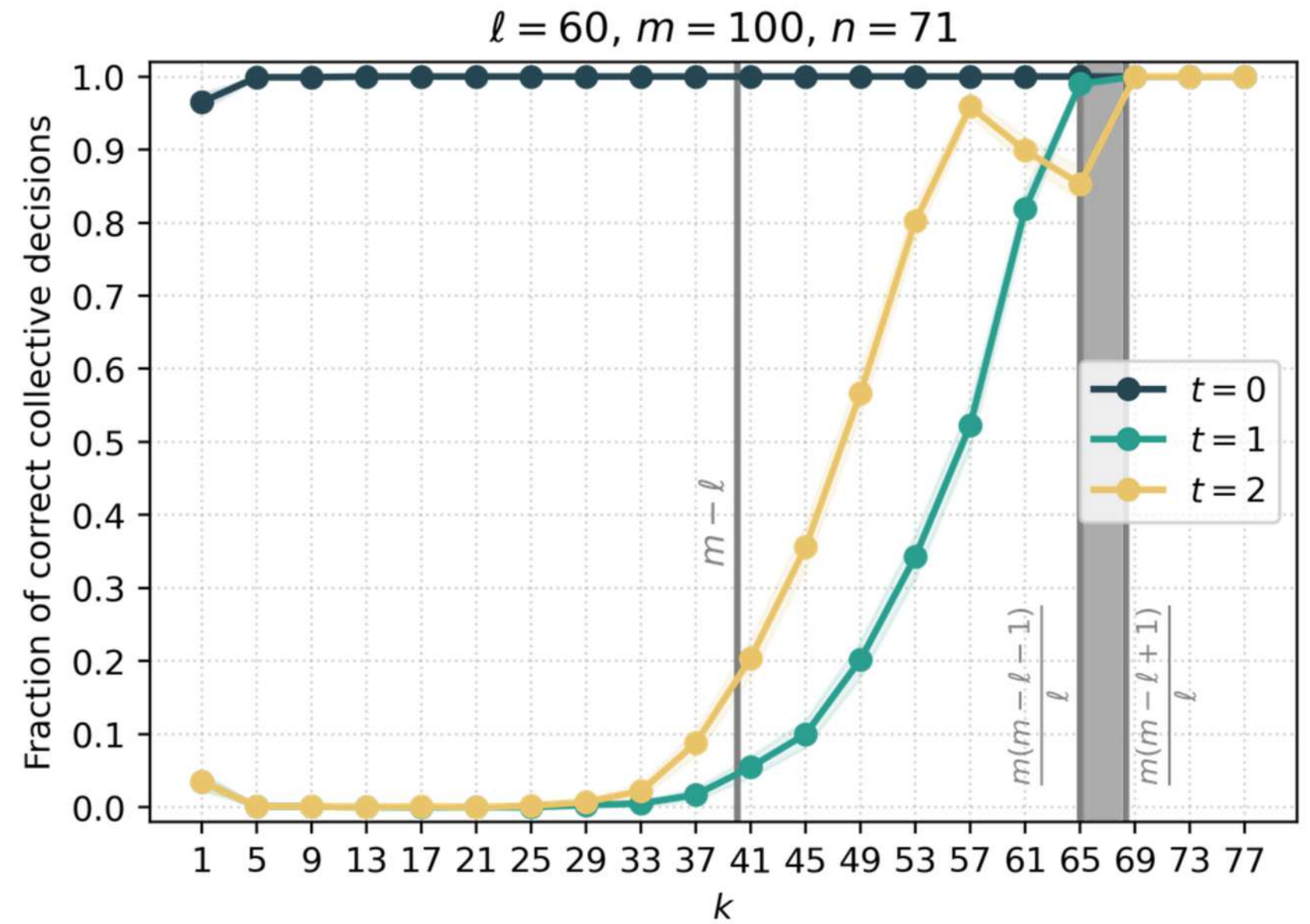
ROUND 2

If $m - \ell < k < m(m - \ell - 1)/\ell$, there is still a fraction of a -supporters left. As $n \rightarrow \infty$, they hold all evidence for a . Once they disclose, it's game over.



RESULT IN A NUTSHELL

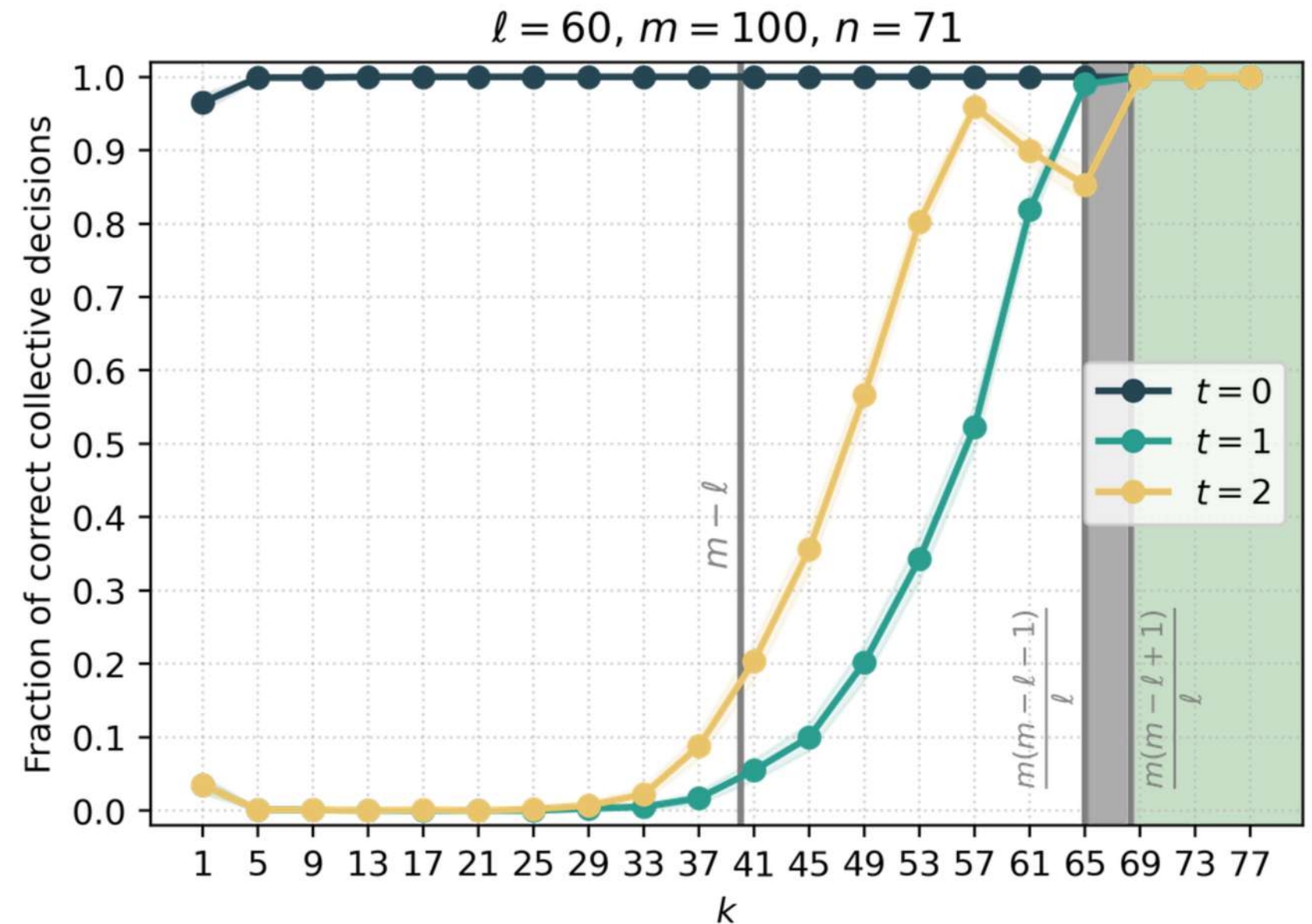
At round 0 the group does well.



RESULT IN A NUTSHELL

At round 0 the group does well.

At round 1 it does well for $k > m(m-\ell+1)/\ell$. For smaller k we get the flip to b .

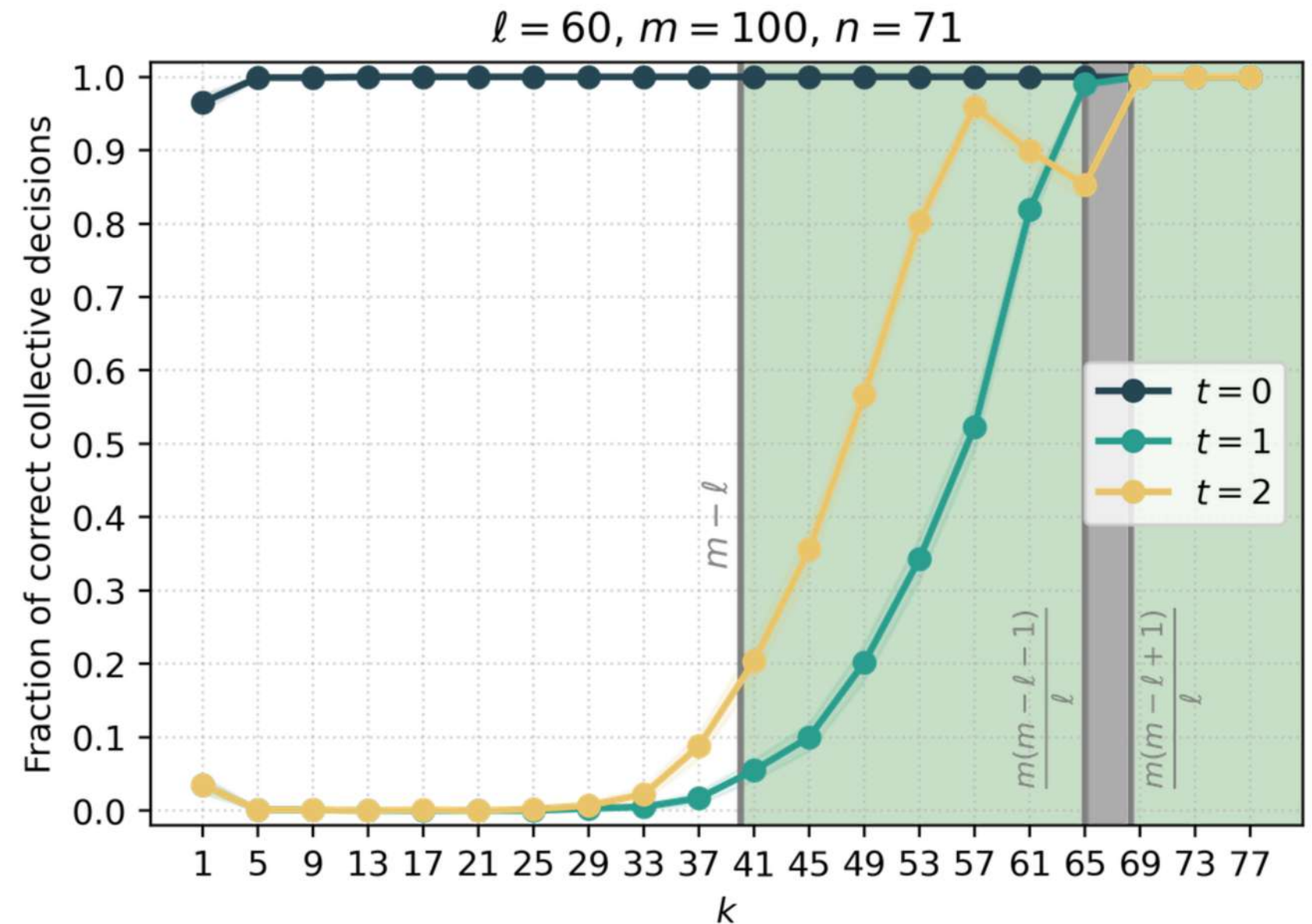


RESULT IN A NUTSHELL

At round 0 the group does well.

At round 1 it does well for $k > m(m-\ell+1)/\ell$. For smaller k we get the flip to b .

At round 2 it does well for $k > m - \ell$: a takes over again.



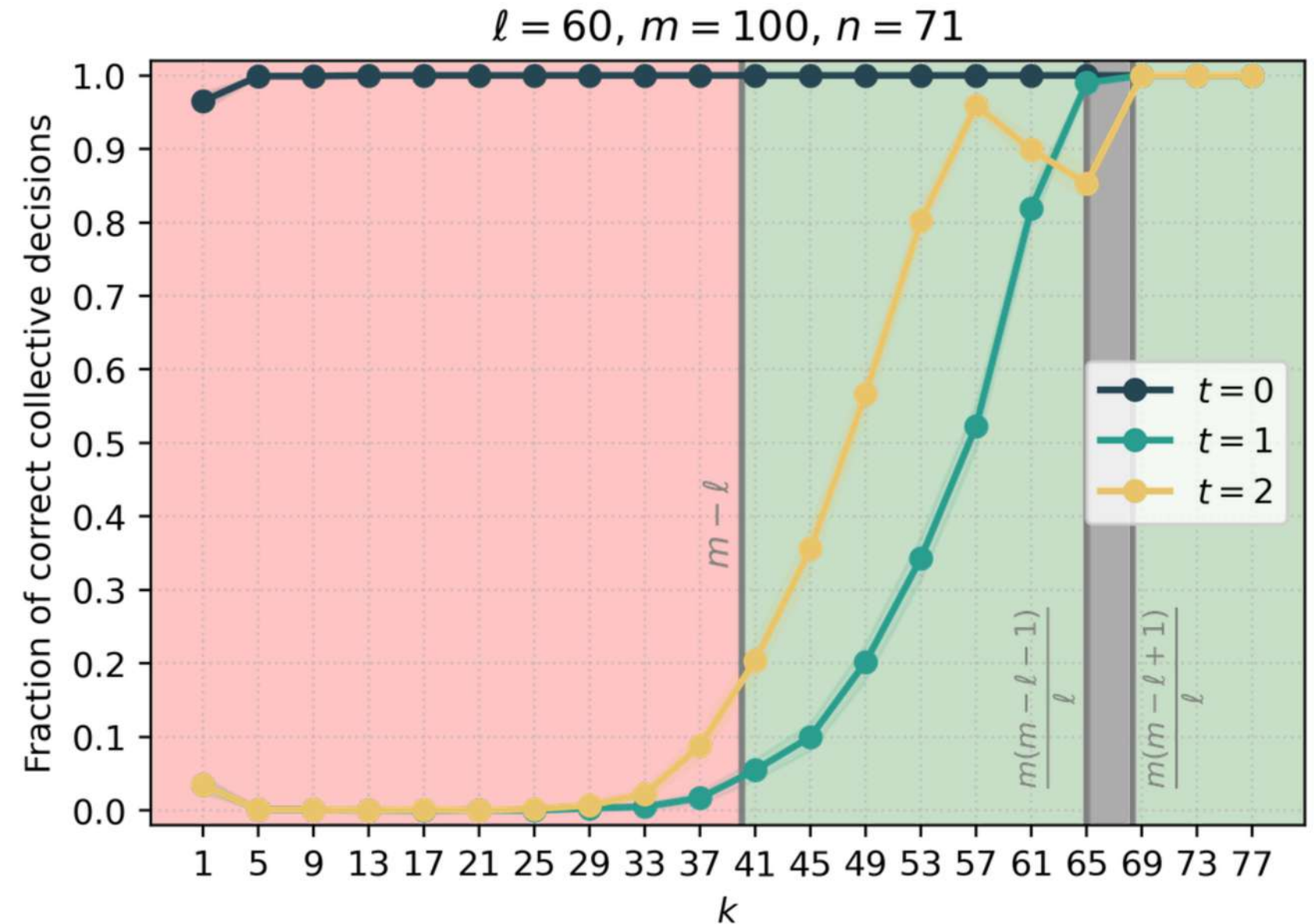
RESULT IN A NUTSHELL

At round 0 the group does well.

At round 1 it does well for $k > m(m-\ell+1)/\ell$. For smaller k we get the flip to b .

At round 2 it does well for $k > m - \ell$: a takes over again.

But for smaller k we get b .



RESULT IN A NUTSHELL

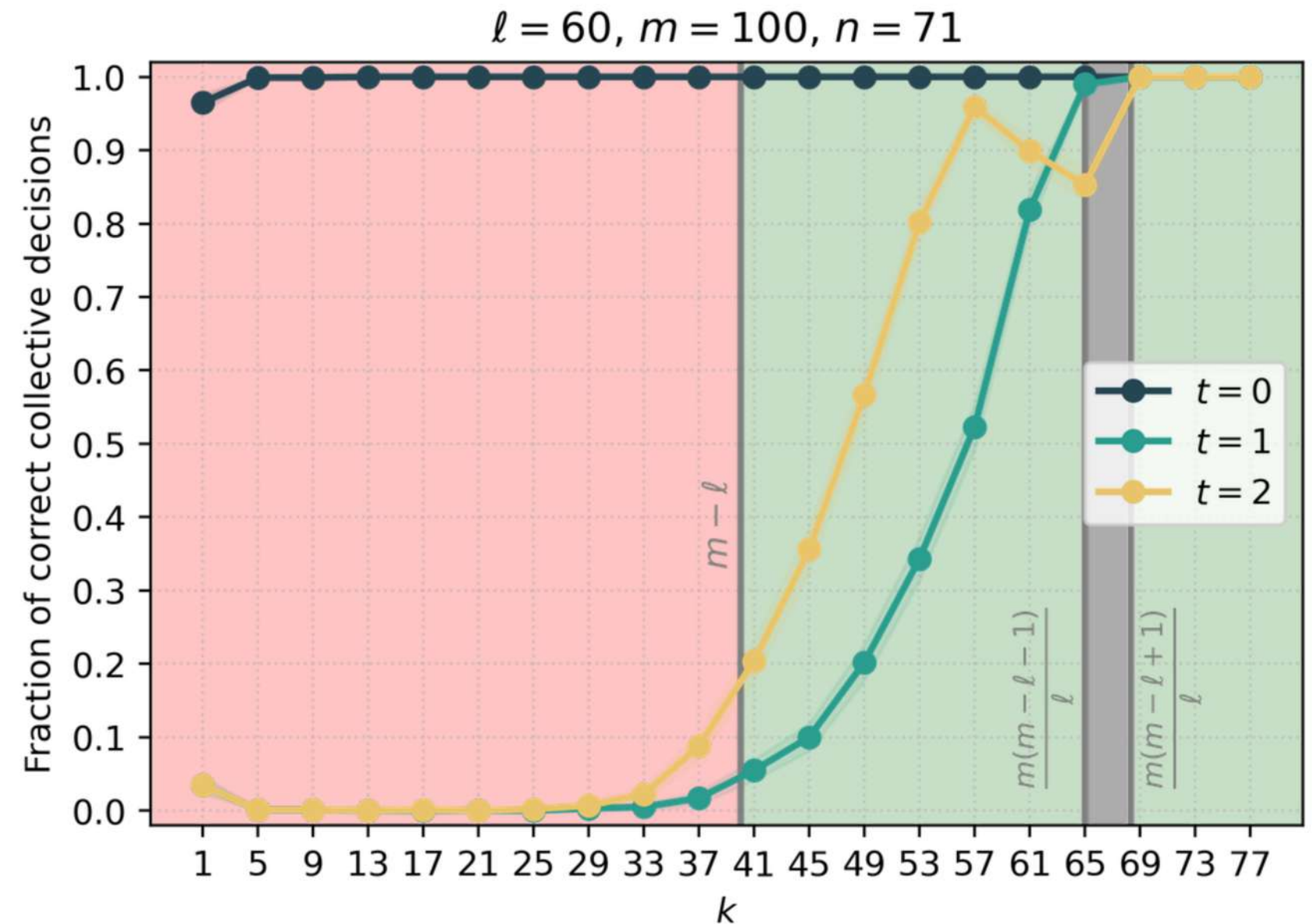
At round 0 the group does well.

At round 1 it does well for $k > m(m-\ell+1)/\ell$. For smaller k we get the flip to b .

At round 2 it does well for $k > m - \ell$: a takes over again.

But for smaller k we get b .

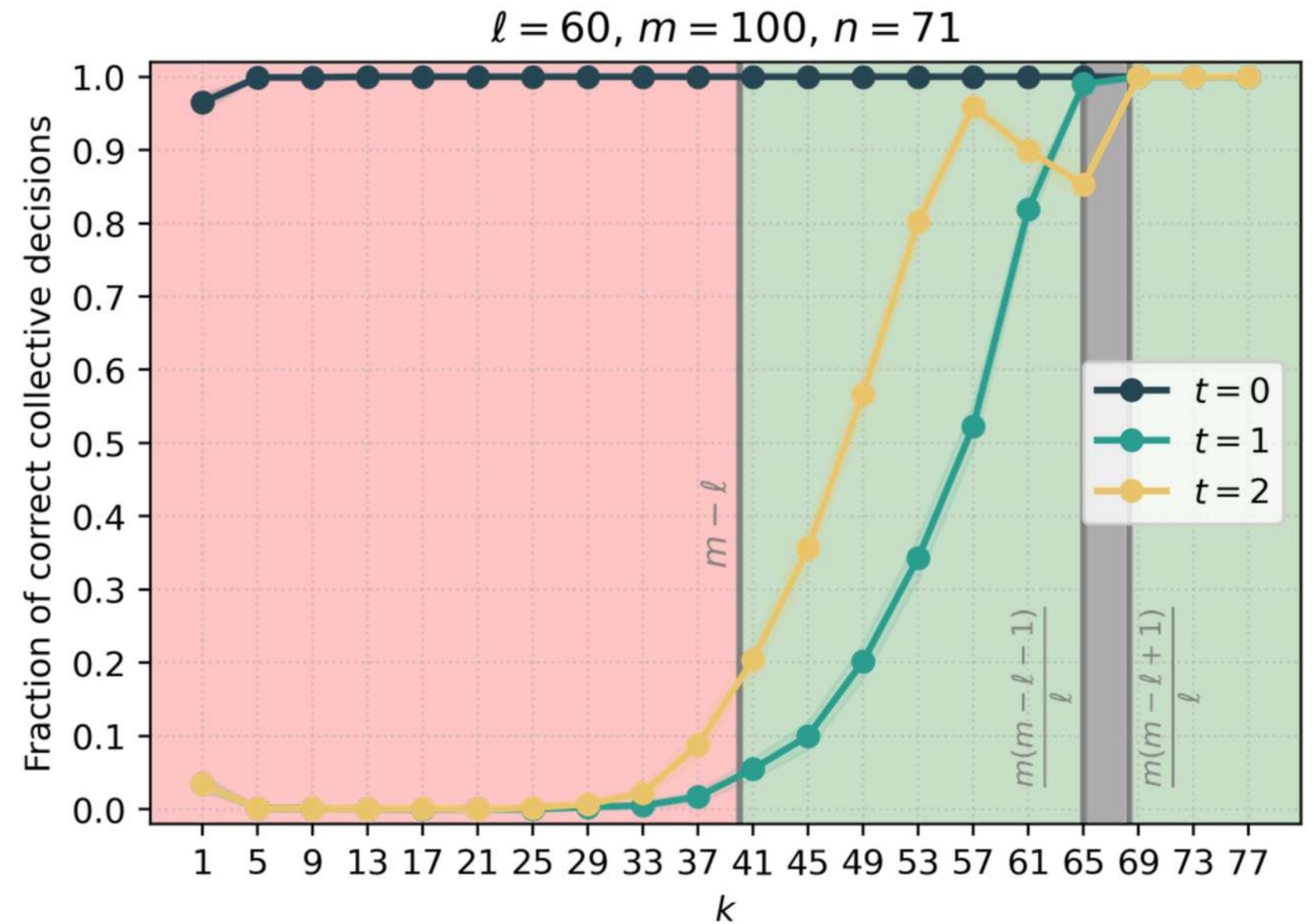
If k is the amount of effort agents are willing to put in, then smaller k is perhaps more realistic.



MORALS

TRADEOFF

To prevent bad evidence from flooding the zone, agents really need to do their homework.

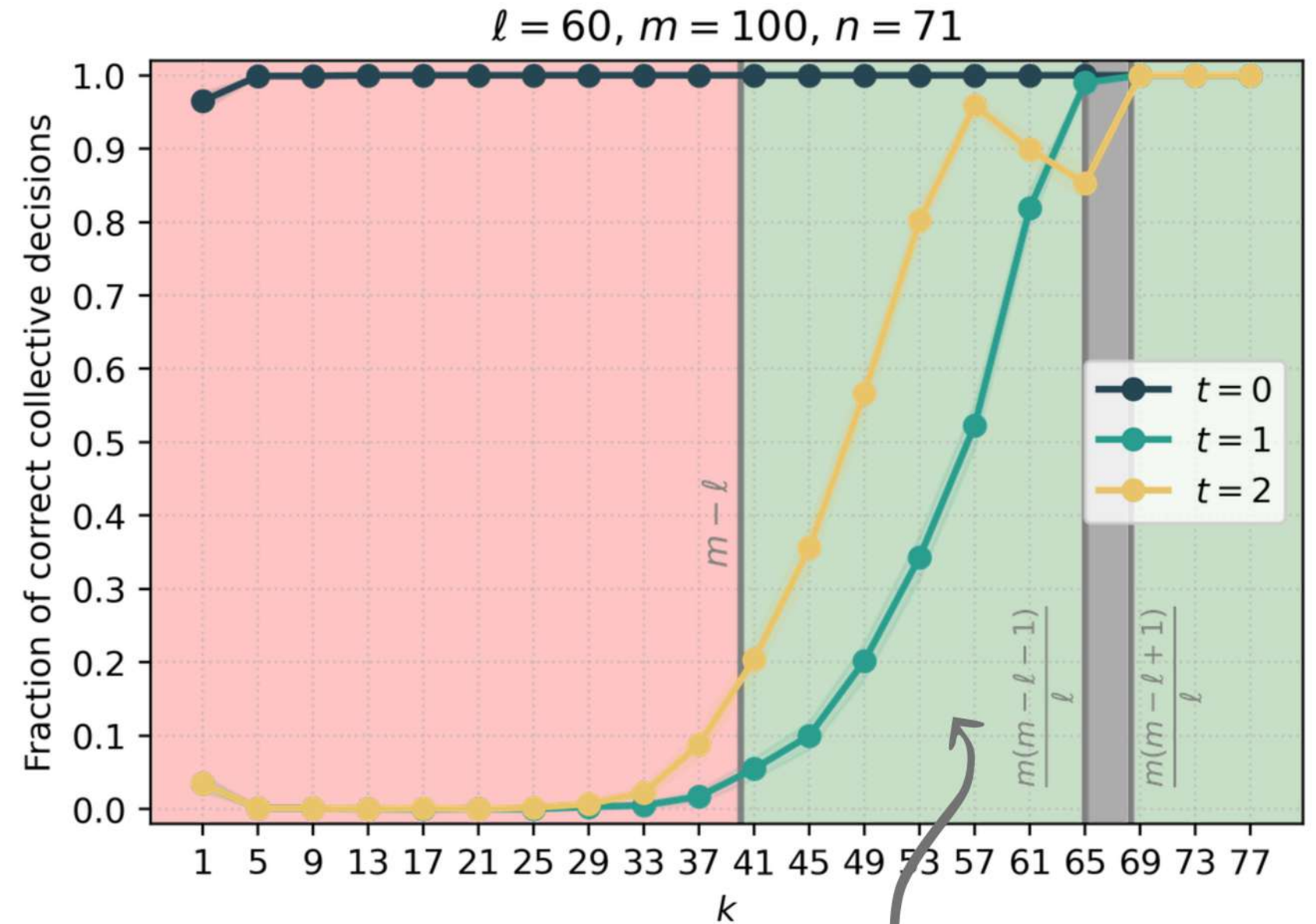


MORALS

especially if deliberation ends before round 2

TRADEOFF

To prevent bad evidence from flooding the zone, agents really need to do their homework.



only at round 2

MORALS

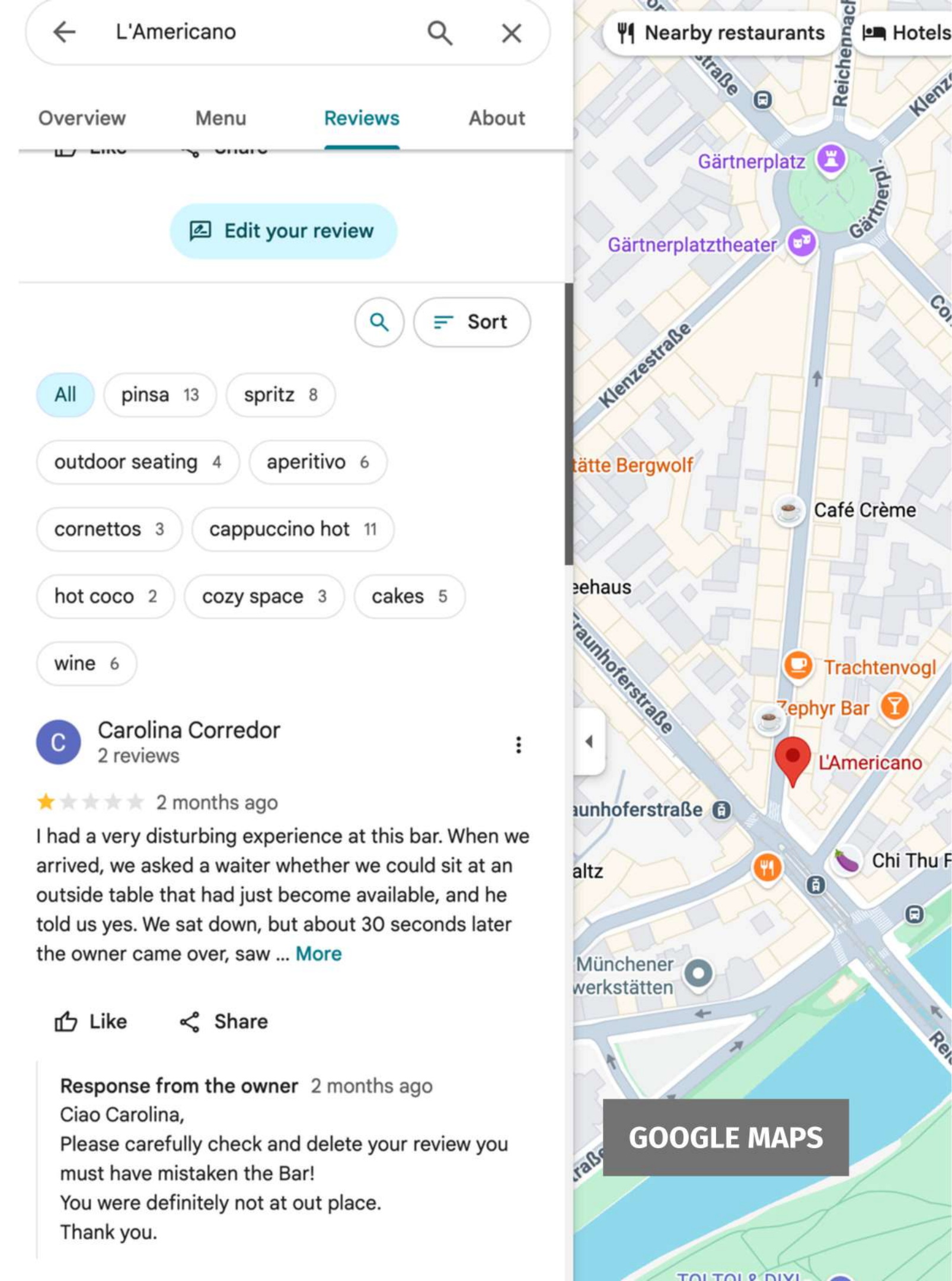
especially if deliberation ends before round 2

TRADEOFF

To prevent bad evidence from flooding the zone, agents really need to do their homework.

AND A DESIGN RECOMMENDATION

Capping discussion early → bad.



MORALS

especially if deliberation ends before round 2

TRADEOFF

To prevent bad evidence from flooding the zone, agents really need to do their homework.

AND A DESIGN RECOMMENDATION

Capping discussion early → bad.

Allowing discussion to follow its course → better.

Sharing is Caring: A Probabilistic Model of Deliberation as Selective Evidence Disclosure



Adrian Haret, Rosa Bauer

23 Feb 2026 (modified: 24 Feb 2026) LOFT 2026 Conference Submission
Conference, Reviewers, Authors Revisions CC BY-SA 4.0

Keywords: deliberation, voting, wisdom of crowds

TL;DR: We study wisdom of crowds in a formal model of deliberation where agents have probabilistic access to evidence and selectively disclose it to their peers in order to induce opinion change.

Abstract:

We analyze a formal model of deliberation in which agents have probabilistic access to evidence and selectively disclose it to their peers in order to induce opinion change. We present theoretical and simulation results that tackle the question: does deliberation improve the collective accuracy of the group? We also study the impact of diversity on the quality of deliberation outcomes.

Submission Number: 35

Filter by reply type... Filter by author... Search keywords... Sort: Newest First

Everyone Program Chairs Submission35 Reviewers Submission35 Authors 4 / 4 replies shown

Add: **Withdrawal**

Paper Decision
Decision by Program Chairs 19 Apr 2026, 14:09 (modified: 19 Apr 2026, 14:46)
Program Chairs, Reviewers, Authors Revisions
Decision: Accept
Comment:
Reviewers found the model

Very interesting extension of previous deliberating models
Official Review by Reviewer MPmt 13 Apr 2026, 04:01 (modified: 19 Apr 2026, 14:58)
Program Chairs, Reviewers, Authors Revisions
Review:
This paper presents a formal model of deliberation as selective evidence disclosure, extending

Can we do any better?

SALIENCE

Make some items of evidence more *salient* than others.

SALIENCE

Make some items of evidence more *salient* than others.

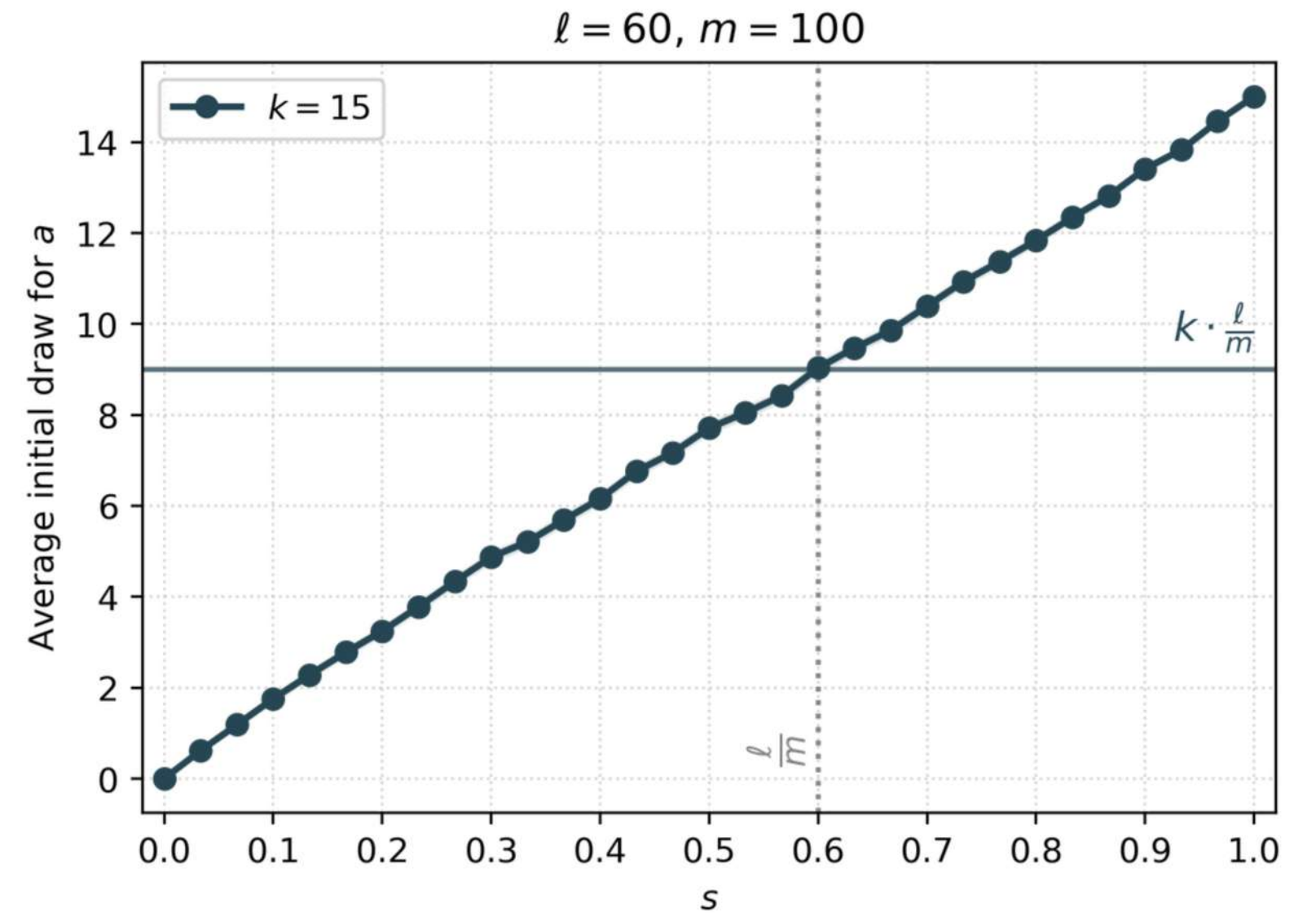
Salience determined by s : higher s means a -evidence more salient.

SALIENCE

Make some items of evidence more *salient* than others.

Salience determined by s : higher s means a -evidence more salient.

When drawing evidence, evidence is sampled in proportion to salience.



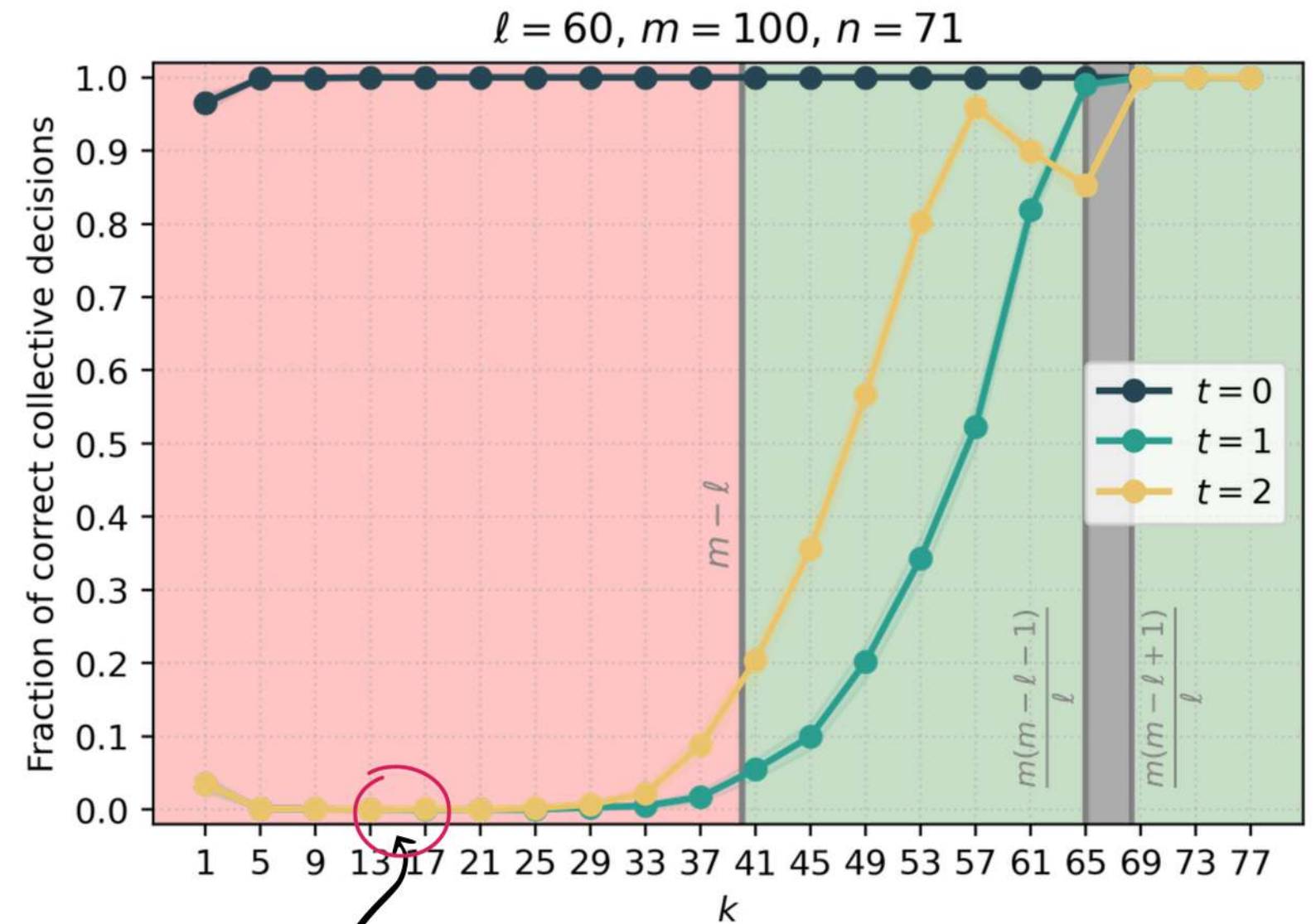
SALIENCE

Make some items of evidence more *salient* than others.

Salience determined by s : higher s means a -evidence more salient.

When drawing evidence, evidence is sampled in proportion to salience.

Simulation results



$k = 15$

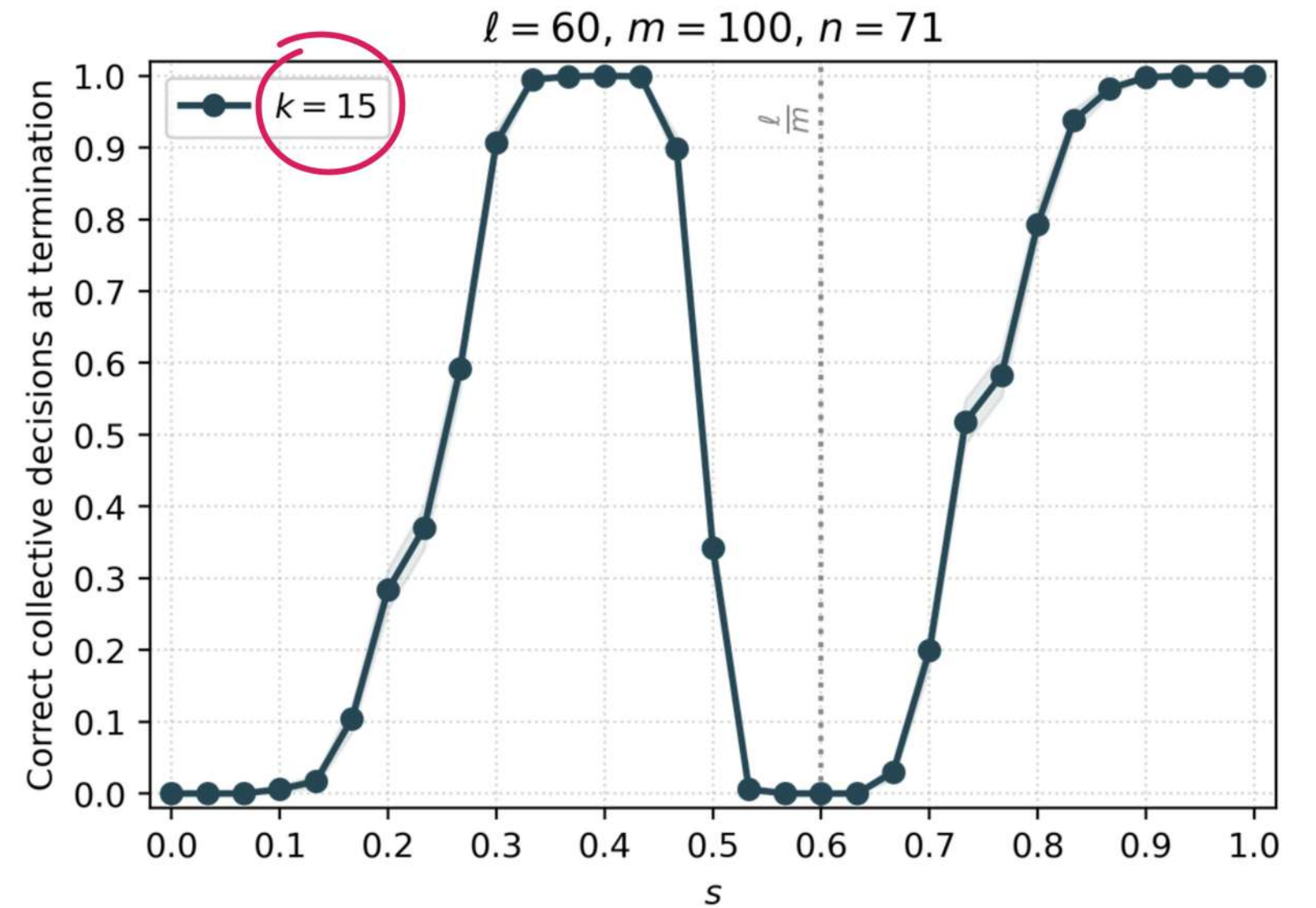
SALIENCE

Make some items of evidence more *salient* than others.

Salience determined by s : higher s means a -evidence more salient.

When drawing evidence, evidence is sampled in proportion to salience.

Simulation results show double hump.



SALIENCE

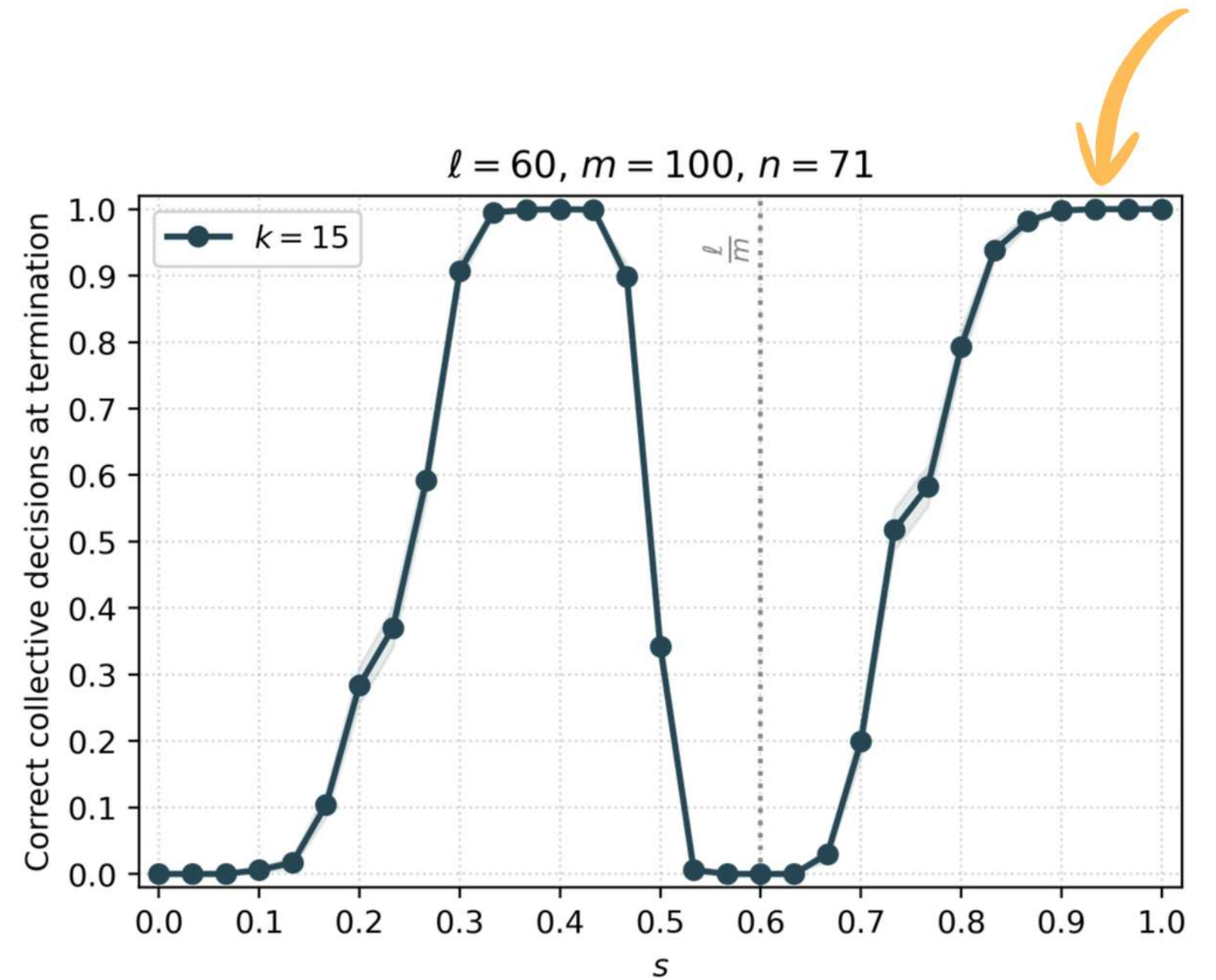
Make some items of evidence more *salient* than others.

Salience determined by s : higher s means a -evidence more salient.

When drawing evidence, evidence is sampled in proportion to salience.

Simulation results show double hump.

In the second ($w > 0.7$): a -evidence more salient, a wins easily.



SALIENCE

Make some items of evidence more *salient* than others.

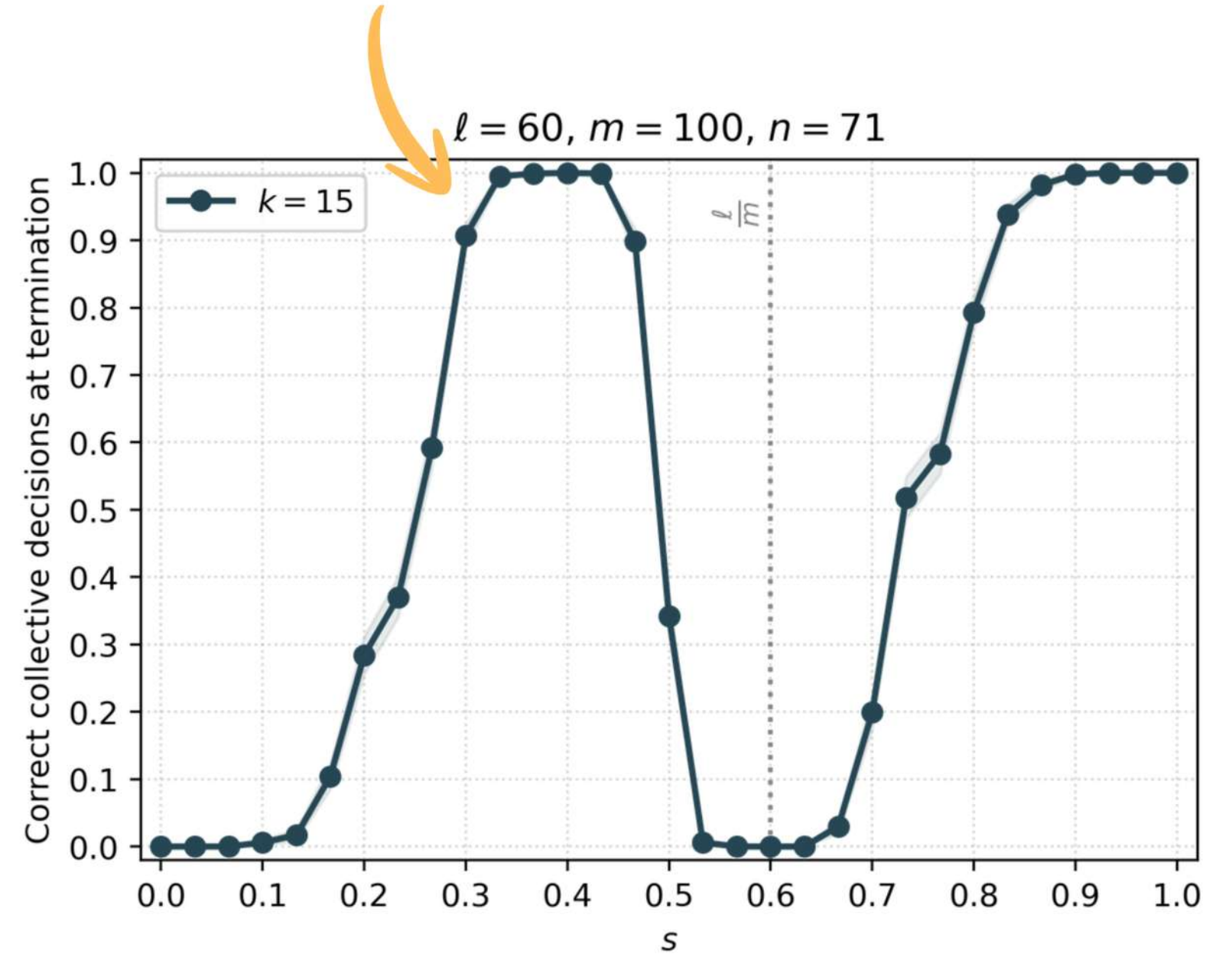
Salience determined by s : higher s means a -evidence more salient.

When drawing evidence, evidence is sampled in proportion to salience.

Simulation results show double hump.

In the second ($w > 0.7$): a -evidence more salient, a wins easily.

In the first ($0.3 < w < 0.5$): b -evidence more salient, but a wins. Initial vote is for b , which makes a -voters dissenters.



EXTRA MORALS

SALIENCE

Making the optimal decision a (bit of an) underdog can go a long way.