

Games with noisy feedback about emotions

Pierpaolo Battigalli
Bocconi and IGER

Nicolò Generoso
Yale

August 22, 2022
EEA-ESEM 2022

Introduction

- Starting point: players' utilities depend on **emotions**.

Introduction

- Starting point: players' utilities depend on **emotions**.
- Key idea: **noisy signals** (e.g., facial cues) may betray emotions.

Introduction

- Starting point: players' utilities depend on **emotions**.
- Key idea: **noisy signals** (e.g., facial cues) may betray emotions.
 - Everyday experience: blushing = embarrassment, smiling = happiness, etc.
 - Experimental evidence: nonverbal communication matters when lying, deceiving, negotiating.

Introduction

- Starting point: players' utilities depend on **emotions**.
- Key idea: **noisy signals** (e.g., facial cues) may betray emotions.
 - Everyday experience: blushing = embarrassment, smiling = happiness, etc.
 - Experimental evidence: nonverbal communication matters when lying, deceiving, negotiating.
- How can we represent emotional signals?

Can such signals shape behavior when players reason strategically?

Leading example: “Buy me an ice-cream”

Leading example: “Buy me an ice-cream”

- Child is at home alone: he chooses *Homework* or *Videogames*.

Leading example: “Buy me an ice-cream”

- Child is at home alone: he chooses *Homework* or *Videogames*.
- Mom gets back home. Child: “Mom, can you please buy me an ice-cream?” Mom: “Did you do your homework?”
Child can choose to answer *Yes* or *No* (note: *Yes* = “I did my homework”). But Child may *blush* if he says *Yes* lying.

Leading example: “Buy me an ice-cream”

- Child is at home alone: he chooses *Homework* or *Videogames*.
- Mom gets back home. Child: “Mom, can you please buy me an ice-cream?” Mom: “Did you do your homework?”
Child can choose to answer *Yes* or *No* (note: *Yes* = “I did my homework”). But Child may *blush* if he says *Yes* lying.
- Mom decides whether to *Buy* the ice-cream or *Not*.

Leading example: “Buy me an ice-cream”

- Child is at home alone: he chooses *Homework* or *Videogames*.
- Mom gets back home. Child: “Mom, can you please buy me an ice-cream?” Mom: “Did you do your homework?”
Child can choose to answer *Yes* or *No* (note: *Yes* = “I did my homework”). But Child may *blush* if he says *Yes* lying.
- Mom decides whether to *Buy* the ice-cream or *Not*.

More general and relevant problem: *disclosure of information*.

Is lying worth it if lies can be spotted?

Roadmap

① Introduction

② Framework

③ Rationality and behavioral predictions

④ Conclusion

Roadmap

① Introduction

② Framework

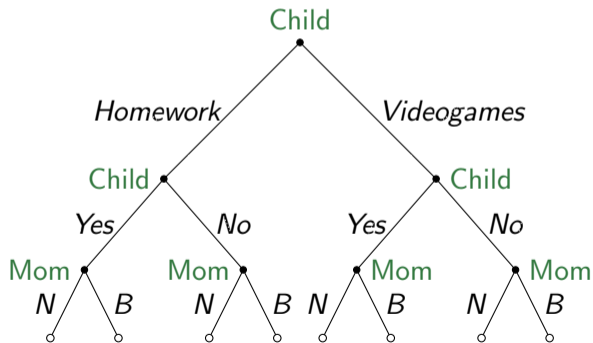
③ Rationality and behavioral predictions

④ Conclusion

The game tree: “Buy me an ice-cream”

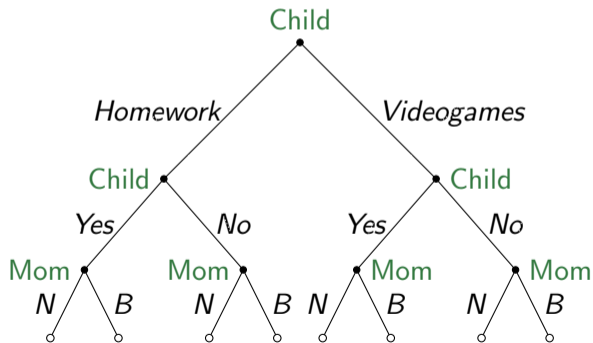
The game tree: “Buy me an ice-cream”

- The tree of action sequences:



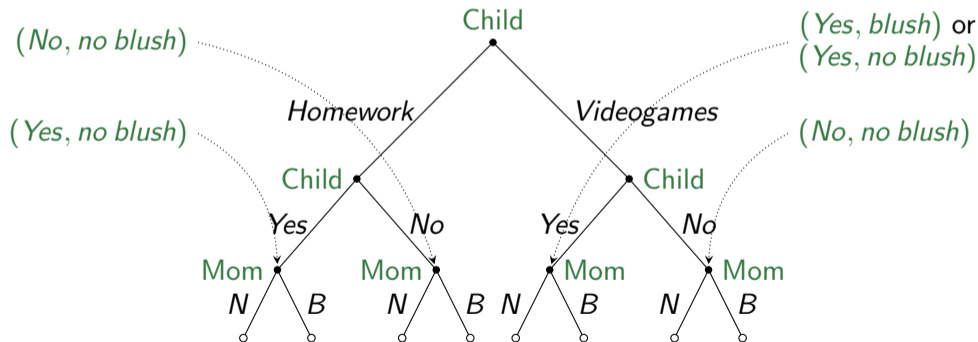
The game tree: “Buy me an ice-cream”

- The tree of action sequences.
- Mom does not observe if Child did *Homework* + Child may blush if he lies.



The game tree: “Buy me an ice-cream”

- The tree of action sequences.
- Mom does not observe if Child did *Homework* + Child may blush if he lies.
- Mom’s information looks like:



The game tree: More generally

The game tree: More generally

- We focus on **multistage games** → the play proceed through stages.

The game tree: More generally

- We focus on **multistage games** → the play proceed through stages.
- Players receive pieces of information (**signals/messages**) after each stage.

The game tree: More generally

- We focus on **multistage games** → the play proceed through stages.
- Players receive pieces of information (**signals/messages**) after each stage.
 - Signals provide information about **emotions** and about the **play**.
 - **Example:** signal (*Yes, blush*) informs Mom of Child's second-stage action and emotions.

The game tree: More generally

- We focus on **multistage games** → the play proceed through stages.
- Players receive pieces of information (**signals/messages**) after each stage.
 - Signals provide information about **emotions** and about the **play**.
 - **Example**: signal (*Yes, blush*) informs Mom of Child's second-stage action and emotions.
- A **history** is a sequence of action and signal/message profiles:

$$h = (\mathbf{a}^1, \mathbf{m}^1, \dots, \mathbf{a}^k, \mathbf{m}^k), \quad H = \text{set of histories.}$$

The game tree: More generally

- We focus on **multistage games** → the play proceed through stages.
- Players receive pieces of information (**signals/messages**) after each stage.
 - Signals provide information about **emotions** and about the **play**.
 - **Example**: signal (*Yes, blush*) informs Mom of Child's second-stage action and emotions.
- A **history** is a sequence of action and signal/message profiles:

$$h = (\mathbf{a}^1, \mathbf{m}^1, \dots, \mathbf{a}^k, \mathbf{m}^k), \quad H = \text{set of histories.}$$

- A **personal history** of player i is

$$h_i = (a_i^1, m_i^1, \dots, a_i^k, m_i^k), \quad H_i = \text{set of personal histories of } i.$$

- **Interpretation**: player i only has access to her actions and her messages.

Behavior and beliefs

Behavior and beliefs

- A **personal external state** of i is a map from **personal histories** to (feasible) **actions**.
 - **Interpretation**: a *description* of behavior (technically, a “strategy”).
 - S_i = set of personal external states of i ; $S = \times_i S_i$ = set of personal external state profiles.

Behavior and beliefs

- A **personal external state** of i is a map from **personal histories** to (feasible) **actions**.
 - *Interpretation*: a *description* of behavior (technically, a “strategy”).
 - S_i = set of personal external states of i ; $S = \times_i S_i$ = set of personal external state profiles.
- A **personal trait** of i is a parameter θ_i .
 - Θ_i = set of personal traits of i ; $\Theta = \times_i \Theta_i$ = set of personal trait profiles.

Behavior and beliefs

- A **personal external state** of i is a map from **personal histories** to (feasible) **actions**.
 - **Interpretation:** a *description* of behavior (technically, a “strategy”).
 - S_i = set of personal external states of i ; $S = \times_i S_i$ = set of personal external state profiles.
- A **personal trait** of i is a parameter θ_i .
 - Θ_i = set of personal traits of i ; $\Theta = \times_i \Theta_i$ = set of personal trait profiles.
- A(n epistemic) **type** of i is a map from **personal histories** to **infinite hierarchies of beliefs** over $S \times \Theta$.
 - **Note:** player i forms beliefs also over S_i (her plan) and Θ_i + no restrictions on beliefs.
 - T_i^∞ = set of types of i ; $T^\infty = \times_i T_i^\infty$ = set of type profiles.

Behavior and beliefs

- A **personal external state** of i is a map from **personal histories** to (feasible) **actions**.
 - **Interpretation:** a *description* of behavior (technically, a “strategy”).
 - S_i = set of personal external states of i ; $S = \times_i S_i$ = set of personal external state profiles.
- A **personal trait** of i is a parameter θ_i .
 - Θ_i = set of personal traits of i ; $\Theta = \times_i \Theta_i$ = set of personal trait profiles.
- A(n epistemic) **type** of i is a map from **personal histories** to **infinite hierarchies of beliefs** over $S \times \Theta$.
 - **Note:** player i forms beliefs also over S_i (her plan) and Θ_i + no restrictions on beliefs.
 - T_i^∞ = set of types of i ; $T^\infty = \times_i T_i^\infty$ = set of type profiles.
- A **state of the world** is $(s, \theta, t^\infty) \in S \times \Theta \times T^\infty$.
 - **Interpretation:** complete description of factors relevant for the strategic interaction.

Emotions

Emotions

- **Emotions** are determined by endogenous beliefs \rightarrow they depend on the type profile t^∞ .

Emotions

- **Emotions** are determined by endogenous beliefs \rightarrow they depend on the type profile t^∞ .
 - **Assumption:** only 1st order beliefs matter \rightarrow denote “truncated type” profiles as $t \in T$.
So $t_i : H_i \rightarrow \Delta(S \times \Theta)$.

Emotions

- **Emotions** are determined by endogenous beliefs \rightarrow they depend on the type profile t^∞ .
 - **Assumption:** only 1st order beliefs matter \rightarrow denote “truncated type” profiles as $t \in T$.
So $t_i : H_i \rightarrow \Delta(S \times \Theta)$.
 - **Example:** if i has **image concerns**, she cares about j 's beliefs about her trait θ_i .

Emotions

- **Emotions** are determined by endogenous beliefs \rightarrow they depend on the type profile t^∞ .
 - **Assumption:** only 1st order beliefs matter \rightarrow denote “truncated type” profiles as $t \in T$.
So $t_i : H_i \rightarrow \Delta(S \times \Theta)$.
 - **Example:** if i has **image concerns**, she cares about j 's beliefs about her trait θ_i .
- **Emotional feedback** is generated randomly according to the map

$$f : S \times \Theta \times T \times H \rightarrow \Delta(M).$$

Emotions

- **Emotions** are determined by endogenous beliefs \rightarrow they depend on the type profile t^∞ .
 - **Assumption:** only 1st order beliefs matter \rightarrow denote “truncated type” profiles as $t \in T$.
So $t_i : H_i \rightarrow \Delta(S \times \Theta)$.
 - **Example:** if i has **image concerns**, she cares about j 's beliefs about her trait θ_i .

- **Emotional feedback** is generated randomly according to the map

$$f : S \times \Theta \times T \times H \rightarrow \Delta(M).$$

- Player i 's **utility** is given by

$$u_i : S \times \Theta \times T \rightarrow \mathbb{R}.$$

- **Note:** t_i^∞ induces a belief over $S \times \Theta \times T_{-i} \rightarrow$ used to compute the expectation of u_i .

Emotions: “Buy me an ice-cream”

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.
 - ★ **Confidence** = he keeps a poker face if he thinks he can get away with the lie even if caught.

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.
 - ★ **Confidence** = he keeps a poker face if he thinks he can get away with the lie even if caught.
 - Thus, **probability of *no blush*** = belief that Mom will *Buy* him the ice-cream even if *blush*.

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.
 - ★ **Confidence** = he keeps a poker face if he thinks he can get away with the lie even if caught.
 - Thus, **probability of *no blush*** = belief that Mom will *Buy* him the ice-cream even if *blush*.
- **Utility:** Child dislikes being seen as a liar (**image concern**).

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.
 - ★ **Confidence** = he keeps a poker face if he thinks he can get away with the lie even if caught.
 - Thus, **probability of no blush** = belief that Mom will *Buy* him the ice-cream even if *blush*.
- **Utility:** Child dislikes being seen as a liar (**image concern**).
 - Child's utility looks like:

$$u_C(s, \theta, t) = \underbrace{\mathbf{1}(\text{ice-cream} | s)}_{\text{likes ice-cream}} + \underbrace{\theta_C}_{\text{videogames appreciation}} \cdot \underbrace{\mathbf{1}(\text{videogames} | s)}_{\text{likes videogames}} - \underbrace{t_M(\text{lie})}_{\text{image concern}}$$

where $\mathbf{1}(\cdot)$ is an indicator function and $t_M(\text{lie})$ is Mom's end-game belief.

Emotions: “Buy me an ice-cream”

- **Emotional feedback:** Child may blush if he plays *Videogames*, and then says he did *Homework*.
 - ★ **Confidence** = he keeps a poker face if he thinks he can get away with the lie even if caught.
 - Thus, **probability of no blush** = belief that Mom will *Buy* him the ice-cream even if *blush*.
- **Utility:** Child dislikes being seen as a liar (**image concern**).
 - Child's utility looks like:

$$u_C(s, \theta, t) = \underbrace{\mathbf{1}(\text{ice-cream} \mid s)}_{\text{likes ice-cream}} + \underbrace{\theta_C}_{\text{videogames appreciation}} \cdot \underbrace{\mathbf{1}(\text{videogames} \mid s)}_{\text{likes videogames}} - \underbrace{t_M(\text{lie})}_{\text{image concern}}$$

where $\mathbf{1}(\cdot)$ is an indicator function and $t_M(\text{lie})$ is Mom's end-game belief.

- **Mom's utility:** she wants to *Buy* the ice-cream iff Child did his *Homework*.

Roadmap

① Introduction

② Framework

③ Rationality and behavioral predictions

④ Conclusion

Behavioral predictions

Behavioral predictions

- We did not impose any requirements on states of the world (s, θ, t^∞) .
 - For example, beliefs may be incoherent or updated “incorrectly”.

Behavioral predictions

- We did not impose any requirements on states of the world (s, θ, t^∞) .
 - For example, beliefs may be incoherent or updated “incorrectly”.
- Our approach: consider sets of states of the world that satisfy some properties → derive behavioral predictions.

Behavioral predictions

- We did not impose any requirements on states of the world (s, θ, t^∞) .
 - For example, beliefs may be incoherent or updated “incorrectly”.
- Our approach: consider sets of states of the world that satisfy some properties → derive behavioral predictions.

For example:

- Consider states (s, θ, t^∞) where i acts “optimally” → joint restrictions on s_i , θ_i , and t_i^∞ .

A sketch of our approach

A sketch of our approach

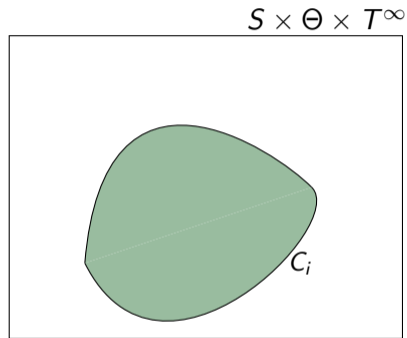
- Start with a (large!) set of states of the world.

$$S \times \Theta \times T^\infty$$



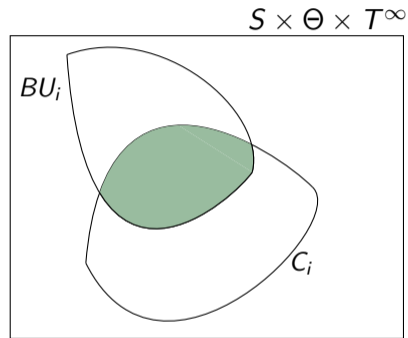
A sketch of our approach

- Start with a (large!) set of states of the world.
- Impose **coherence** (beliefs of different orders agree).



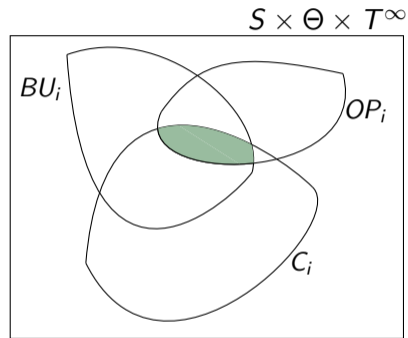
A sketch of our approach

- Start with a (large!) set of states of the world.
- Impose **coherence** (beliefs of different orders agree).
- Impose “**correct**” belief updating.



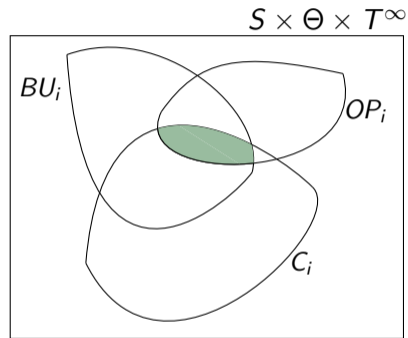
A sketch of our approach

- Start with a (large!) set of states of the world.
- Impose **coherence** (beliefs of different orders agree).
- Impose “correct” belief updating.
- Impose **optimal planning**.



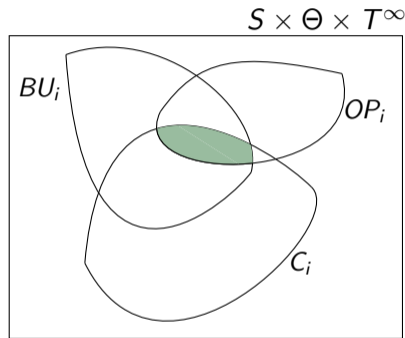
A sketch of our approach

- Start with a (large!) set of states of the world.
- Impose **coherence** (beliefs of different orders agree).
- Impose “correct” belief updating.
- Impose **optimal planning**.
- End up with (a standard notion of) **rationality**.
- **Behavioral predictions** = projection on $S \times \Theta$.



A sketch of our approach

- Start with a (large!) set of states of the world.
- Impose **coherence** (beliefs of different orders agree).
- Impose “correct” **belief updating**.
- Impose **optimal planning**.
- End up with (a standard notion of) **rationality**.
- **Behavioral predictions** = projection on $S \times \Theta$.



Why useful? We (and our players) can think about the white areas (where the **cognitive failures** are)!

Rationality

Rationality

- Starting point: **rationality**.
 - **Cognitive side**: beliefs are coherent and updated in a Bayesian way.
 - **Behavioral side**: course of action is optimally planned and executed.

Rationality

- Starting point: **rationality**.
 - **Cognitive side**: beliefs are coherent and updated in a Bayesian way.
 - **Behavioral side**: course of action is optimally planned and executed.
- Next step: **rationality and common (strong) belief in rationality** (RCSBR).
 - A **rationalizability**-like solution concept captures the behavioral implications of RCSBR.
 - Entailed **forward-induction** reasoning → players make sense of what they observe in a way consistent (if possible) with others being rational and strategically sophisticated.

How our solution looks like: “Buy me an ice-cream”

How our solution looks like: “Buy me an ice-cream”

- 1 Child Does not make sense to say *No* after *Homework*.

How our solution looks like: “Buy me an ice-cream”

① **Child** Does not make sense to say *No* after *Homework*.

Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).

How our solution looks like: “Buy me an ice-cream”

- 1 **Child** Does not make sense to say *No* after *Homework*.
Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).
- 2 **Child** Consider step 1: if he blushes, Mom will *Not* buy the ice-cream → no confidence → optimal action after *Videogames* is to say *No* (always optimal to tell the truth).

How our solution looks like: “Buy me an ice-cream”

① **Child** Does not make sense to say *No* after *Homework*.

Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).

② **Child** Consider step 1: if he blushes, Mom will *Not* buy the ice-cream → no confidence → optimal action after *Videogames* is to say *No* (always optimal to tell the truth).

Mom Consider step 1: if Child says *No*, he must have played *Videogames* → Does *Not* buy the ice-cream if (*No, no blush*).

How our solution looks like: “Buy me an ice-cream”

- ① **Child** Does not make sense to say *No* after *Homework*.
Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).
- ② **Child** Consider step 1: if he blushes, Mom will *Not* buy the ice-cream → no confidence → optimal action after *Videogames* is to say *No* (always optimal to tell the truth).
Mom Consider step 1: if Child says *No*, he must have played *Videogames* → Does *Not* buy the ice-cream if (*No, no blush*).
- ③ **Mom** Consider step 2: Child tells the truth → *Buys* the ice-cream if (*Yes, no blush*).

How our solution looks like: “Buy me an ice-cream”

- 1 **Child** Does not make sense to say *No* after *Homework*.
Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).
- 2 **Child** Consider step 1: if he blushes, Mom will *Not* buy the ice-cream → no confidence → optimal action after *Videogames* is to say *No* (always optimal to tell the truth).
Mom Consider step 1: if Child says *No*, he must have played *Videogames* → Does *Not* buy the ice-cream if (*No, no blush*).
- 3 **Mom** Consider step 2: Child tells the truth → *Buys* the ice-cream if (*Yes, no blush*).
- 4 **Child** Consider step 3: Child chooses between *Homework* plus ice-cream and *Videogames* without ice-cream (based on his personal trait).

How our solution looks like: “Buy me an ice-cream”

① **Child** Does not make sense to say *No* after *Homework*.

Mom If Child blushes, he must have played *Videogames* → she does *Not* buy the ice-cream if (*Yes, blush*).

② **Child** Consider step 1: if he blushes, Mom will *Not* buy the ice-cream → no confidence → optimal action after *Videogames* is to say *No* (always optimal to tell the truth).

Mom Consider step 1: if Child says *No*, he must have played *Videogames* → Does *Not* buy the ice-cream if (*No, no blush*).

③ **Mom** Consider step 2: Child tells the truth → *Buys* the ice-cream if (*Yes, no blush*).

④ **Child** Consider step 3: Child chooses between *Homework* plus ice-cream and *Videogames* without ice-cream (based on his personal trait).

★ **Takeaway** = signals about emotions and image concerns yield “full disclosure”.

Roadmap

- ① Introduction
- ② Framework
- ③ Rationality and behavioral predictions
- ④ Conclusion

What's next?

What's next?

- **Applications:**

- More likely to accept unfair offers if the proposer smiles?
- More conciliatory when negotiating with angry counterparts?

What's next?

- **Applications:**
 - More likely to accept unfair offers if the proposer smiles?
 - More conciliatory when negotiating with angry counterparts?
- **Behavioral implications** of different sets of assumptions:
 - Here, standard notion of rationality, very sophisticated (especially Child!).
 - But the framework allows us to focus on specific failures of rationality and derive corresponding predictions.

Thank you

`nicolo.generoso@yale.edu`