# Formal Analysis

## Hilary Term, Uncertainty and PBE

Thomas Bräuninger

University of Mannheim
thomas.braeuninger@uni-mannheim.de

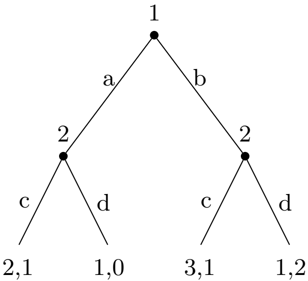
February 15, 2022

# Agenda

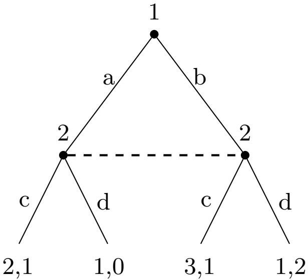
1. Information, uncertainty and Perfect Bayesian equilibria
2. Type 1: Games where uninformed player moves first
3. Type 2: Games where both players are uninformed
4. Type 3: Games where uninformed player moves second

# Perfect information

An extensive game

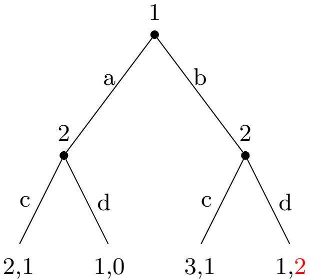


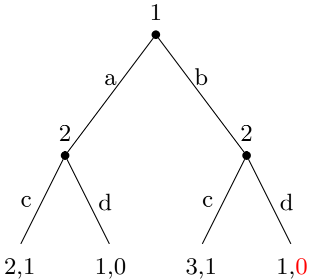
# Uncertainty I




- 2 does not observe 1's move
- 2 does not know whether she is at left or right node
- a game with *imperfect information*

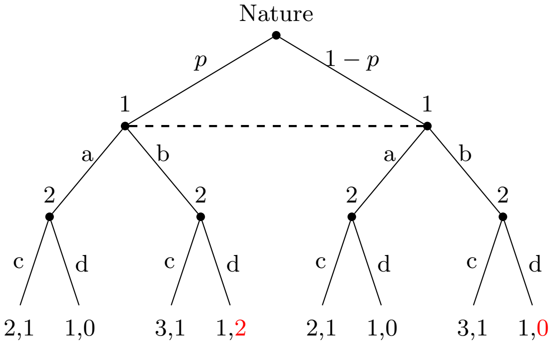
# Uncertainty II




- 1 does not know 2's payoff (or, utility function of player 2, the *type* of player 2, the *state of the world*, ...)
- 1 does not know whether they are playing left or right game
- 'incomplete' description, not really a game ...

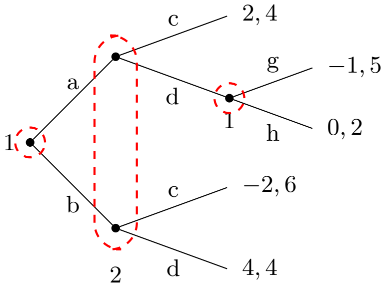
# Uncertainty II

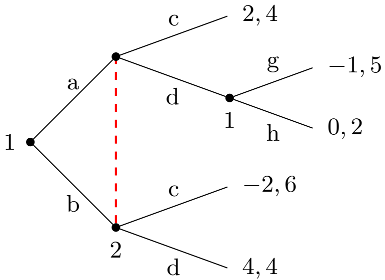



- add *Nature* as an additional, non-strategic player selecting *types* or *states* according to some (known) probability distribution
- a game with *incomplete information*

# Information sets

**Definition (Information set).** An *information set* of a player is a set of nodes of the game tree among which she cannot distinguish. The player does not "know" where she is. A *singleton* is an information set that consists of only one node.



# Perfect and complete information

*Perfect information*

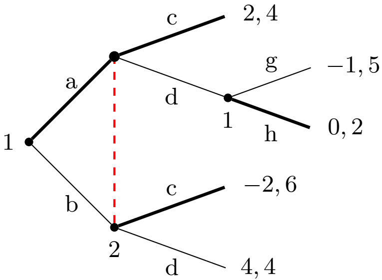
- in a game with perfect information every information set is a singleton
- otherwise it's a game with imperfect information

*Complete information*

- in a game with *incomplete information* "Nature" moves first and is not observed by at least one player
- otherwise the game is referred to as one of complete information

# Beliefs

The previous example is instructive




- we kind of know that $(ah, c)$ is the equilibrium
- so while 2 cannot be certain where she is, the only reasonable *belief* is: “Hey, I am at the upper node!”

# Beliefs

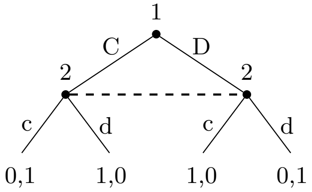
We shall define this term properly

**Definition (Belief).** A belief $b$ at an information set $I$ is a probability distribution over all elements of $I$.

In the previous example, we would write $b = (1, 0)$.

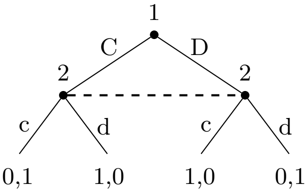
# Another example (and more on normal form vs extensive games)

What is 2's belief?




- with *perfect* information, $c|C$ and $d|D$ would be best responses
- whereas with imperfect information ...
- players move simultaneously ... A normal form game!

# Another example (and more on normal form vs extensive games)




$\Leftrightarrow$

| | | 2 | |
|---|---|---|---|
| | | c | d |
| 1 | C | 0, 1 | 1, 0 |
| | D | 1, 0 | 0, 1 |

- matching pennies has a unique NE (in mixed strategies): $s = \left((\frac{1}{2}C, \frac{1}{2}D), (\frac{1}{2}c, \frac{1}{2}d)\right)$
- hence, given $s$, 2's belief must be $b = (\frac{1}{2}, \frac{1}{2})$
- ... and given $b$, $s_2 = (\frac{1}{2}c, \frac{1}{2}d)$ indeed is a best response

# Sequentially rational strategies and weakly consistent beliefs

Two more definitions

**Definition (Sequential rational).** Given a profile of beliefs $b$, a strategy profile $s$ is *sequentially rational* if at any information set, $s$ is a best response given $b$

**Definition (Weakly consistent).** A belief profile $b$ is *weakly consistent* relative to strategy $s$ if the beliefs are formed "according to Bayes' rule whenever possible".

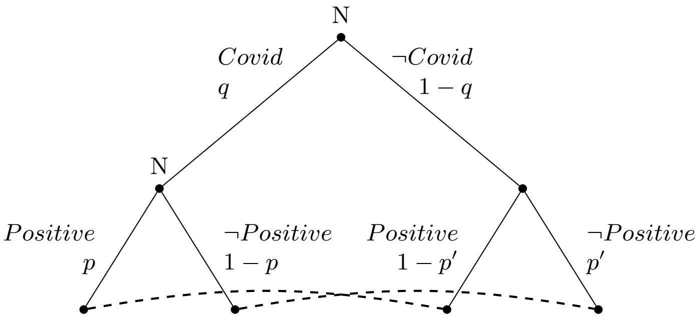
# Bayes' rule

We still have to define what is meant by "according to Bayes' rule whenever possible"

- Recall Bayes' rule

$$
\begin{aligned}
\Pr(A|B) &= \frac{\Pr(B|A)\Pr(A)}{\Pr(B)} \\
&= \frac{\Pr(B|A)\Pr(A)}{\Pr(B|A)\Pr(A) + \Pr(B|notA)\Pr(notA)}
\end{aligned}
$$

# Belief updating using Bayes' rule




What is a consistent belief for the first information set?

$$\begin{aligned}\Pr(C \mid P) &= \frac{\Pr(P \mid C)\Pr(C)}{\Pr(P \mid C)\Pr(C) + \Pr(P \mid \neg C)\Pr(\neg C)} \\ &= \frac{pq}{pq + (1-p')(1-q)}\end{aligned}$$

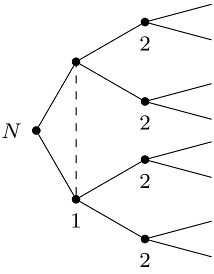
# Perfect Bayesian equilibrium

**Definition (Perfect Bayesian equilibrium).** In a sequential game with imperfect information the pair $(s, b)$ of strategy profile $s$ and belief profile $b$ is called a *perfect Bayesian equilibrium* if

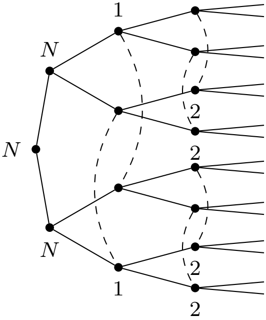
- $s$ is sequentially rational with regard to $b$, and
- $b$ is weakly consistent with regard to $s$

Note: if an information set is not reached under the strategy then Bayes' rule does not work and **all** beliefs are weakly consistent (these are called "out-of-equilibrium-beliefs")

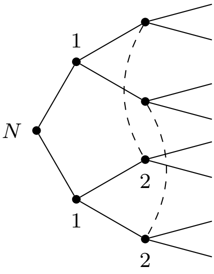
# Typical information structures in incomplete information games




type 1

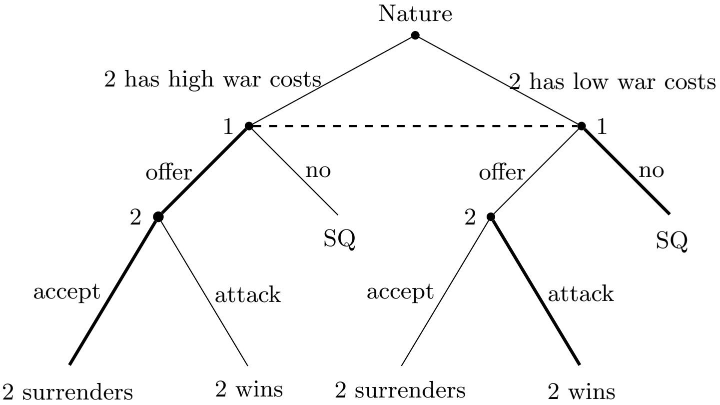



type 2

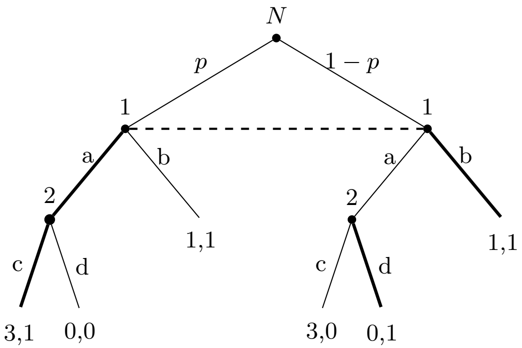



type 3

# Type 1: Simplified bargaining with uncertainty over costs



# Same story with parameters ...




- 1's belief simply is $b_1 = (p, 1-p)$
- expected utility from $a$ and $b$:

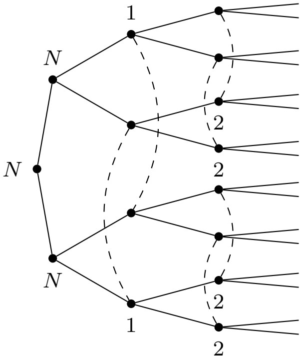
$$EU_1(a) = 3p + 0(1-p) = 3p$$
$$EU_2(b) = 1p + 1(1-p) = 1$$

# Type 1

$$EU_1(a) = 3p$$
$$EU_2(b) = 1$$

- $a$ is best response iff $p \geq \frac{1}{3}$, $b$ is best response iff $p \leq \frac{1}{3}$
- thus $s = (a, cd)$ is sequentially rational with regard to $b_1$ iff $p \geq \frac{1}{3}$, $s = (b, cd)$ is sequentially rational with regard to $b_1$ iff $p \leq \frac{1}{3}$
- $b_1$ is weakly consistent with regard to $s$
- $PBE = \{(a, cd; (p, 1-p))$ if $p \geq \frac{1}{3}, (b, cd; (p, 1-p))$ if $p \leq \frac{1}{3}\}$

# Type 2: Both are uninformed about other player's type



# Preventive war as an assurance game

| | | 2 | |
|---|---|---|---|
| | | $Not$ | $Attack$ |
| 1 | $Not$ | $s_1, s_2$ <br> $(1, 1)$ | $p_1^s - c_1, p_2^f - c_2$ <br> $(.1, .9)$ |
| | $Attack$ | $p_1^f - c_1, p_2^s - c_2$ <br> $(.9, .1)$ | $p_1 - c_1, p_2 - c_2$ <br> $(.5, .5)$ |

- $p$'s are winning probabilities, $c$'s are costs of war
- assumption: $p_i^{second} < p_i < p_i^{first}$
- if $s_i \geq p_i^f - c_i$: two NE, one is Pareto-superior

# Preventive war as an assurance game when players are either greedy ($G$) or status quo ($S$) types

$q_1$, $q_2$:

| | $N$ | $A$ |
|---|---|---|
| $N$ | $s_1^S, s_2^S$ | $p_1^s - c_1$, $p_2^f - c_2$ |
| $A$ | $p_1^f - c_1$, $p_2^s - c_2$ | $p_1 - c_1$, $p_2 - c_2$ |

$q_1$, $1 - q_2$:

| | $N$ | $A$ |
|---|---|---|
| $N$ | $s_1^S, s_2^G$ | $p_1^s - c_1$, $p_2^f - c_2$ |
| $A$ | $p_1^f - c_1$, $p_2^s - c_2$ | $p_1 - c_1$, $p_2 - c_2$ |

$1 - q_1$, $q_2$:

| | $N$ | $A$ |
|---|---|---|
| $N$ | $s_1^G, s_2^S$ | $p_1^s - c_1$, $p_2^f - c_2$ |
| $A$ | $p_1^f - c_1$, $p_2^s - c_2$ | $p_1 - c_1$, $p_2 - c_2$ |

$1 - q_1$, $1 - q_2$:

| | $N$ | $A$ |
|---|---|---|
| $N$ | $s_1^G, s_2^G$ | $p_1^s - c_1$, $p_2^f - c_2$ |
| $A$ | $p_1^f - c_1$, $p_2^s - c_2$ | $p_1 - c_1$, $p_2 - c_2$ |

- assume $s_i^G < p_i^f - c_i < s_i^S = 1$

## Preventive war as an assurance game when players are either greedy ($G$) or status quo ($S$) types

- for greedy types: dominant strategy is to attack (check!)
- for status quo types:

$$EU_i(N) = q_{-i}s_i^S + (1 - q_{-i})(p_i^s - c_i)$$
$$EU_i(A) = q_{-i}(p_i^f - c_i) + (1 - q_{-i})(p_i - c_i)$$

Solving $EU_i(N) \geq EU_i(A)$ leads to

$$q_{-i} \geq \frac{1}{\frac{1-(p_i^f - c_i)}{p_i - p_i^s} + 1}$$

in plain words: $S$-types will not attack if the prob. that the opponent is an $S$-type is large enough (though they will be attacked with prob. $q_{-i}$)

# Preventive war as an assurance game when players are either greedy ($G$) or status quo ($S$) types

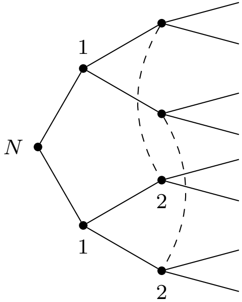
Let

$$T = \frac{1}{\frac{1-(p_i^f - c_i)}{p_i - p_i^s} + 1}.$$

As $\frac{\partial T}{\partial p_i^f} > 0$, $\frac{\partial T}{\partial c_i} < 0$, we have

- if the level of "trust" (that the opponent is an $S$-type) falls below $T$, conflict can occur *even if both* are $S$-types (the problem of mistrust)
- if first strike advantage (winning prob.) is large, or costs of war are small, more trust is needed

# Type 3: Uninformed player moves second




- (probably) most interesting case: can 2 *learn* something about state from 1's move?
- cp. PA framework: (informed) agent moves first, (uninformed) principal second

# Interest group lobbying

- sequence of moves
  - $N$ chooses state of the world $A$ or $B$ (e.g. 'policy $a$ is effective in reducing unemployment' vs 'policy $b$ is effective in reducing unemployment'
  - 1 (lobbyist) sends a message $\alpha$ or $\beta$ ($\alpha$='$a$ is effective in reducing unemployment', $\beta$='$b$ is effective in reducing unemployment')
  - 2 (policymaker) chooses policy $a$ or $b$
- payoffs
  - policymaker wants to match policy to state of the world

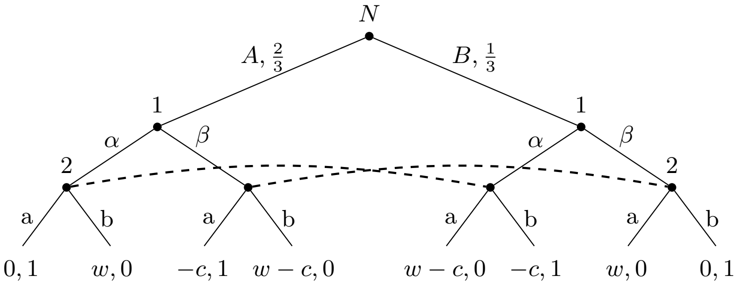
$$u_2(a, A), u_2(b, B) \geq u_2(a, B), u_2(b, A)$$

  - lobbyist has opposed preferences

$$w = u_1(a, B), u_1(b, A) \geq u_1(a, A), u_1(b, B) = 0$$

but incurs costs $c > 0$ if caught lying

# Interest group lobbying



# Finding equilibria

Before we start

- focus on pure strategy equilibria (no mixed strategies for now)
- 1 has four pure strategies

$$(\alpha|A, \beta|B), (\alpha|A, \alpha|B), (\beta|A, \beta|B), (\beta|A, \alpha|B)$$

- 2 has four pure strategies

$$(a|\alpha, b|\beta), (a|\alpha, a|\beta), (b|\alpha, b|\beta), (b|\alpha, a|\beta)$$

- to find equilibria we simply posit one and then check whether someone has an incentive to defect

# Truthful reporting?

- *suppose*

  $$s_1 = (\alpha|A, \beta|B)$$

  is an equilibrium strategy
- then $\Pr(A|\alpha) = \Pr(B|\beta) = 1$, $\Pr(B|\alpha) = \Pr(A|\beta) = 0$ i.e. 2's weakly consistent belief is

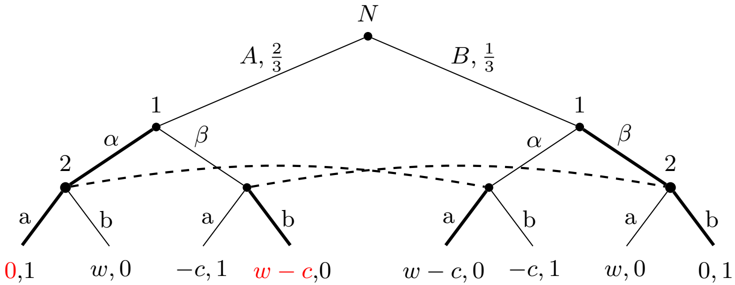
  $$b_2 = \big((1,0),(0,1)\big)$$

- hence 2's sequentially rational strategy is

  $$s_2 = (a|\alpha, b|\beta)$$

- now, is $s_1$ really a best response?

# Truthful reporting?




- if $0 \geq w - c$ then 1 has no incentive to switch to $(\beta|A, \beta|B)$
- same holds for $(\alpha|A, \alpha|B)$ and $(\beta|A, \alpha|B)$
- thus $\big((\alpha|A, \beta|B), (a|\alpha, b|\beta); ((1,0),(0,1))\big)$ is a PBE if $w \leq c$
- this is a *separating* equilibrium

## Same message?

- *assume*

$$s_1 = (\alpha|A, \alpha|B)$$

- then

$$\begin{aligned}
\Pr(A|\alpha) &= \tfrac{2}{3} \ \checkmark \\
\Pr(B|\alpha) &= \tfrac{1}{3} \ \checkmark \\
\Pr(A|\beta) &= \Pr(\beta|A)\Pr(A)/\Pr(\beta) = \tfrac{0p}{0} \ \lightning \\
\Pr(B|\beta) &= \Pr(\beta|B)\Pr(B)/\Pr(\beta) = \tfrac{0(1-p)}{0} \ \lightning
\end{aligned}$$

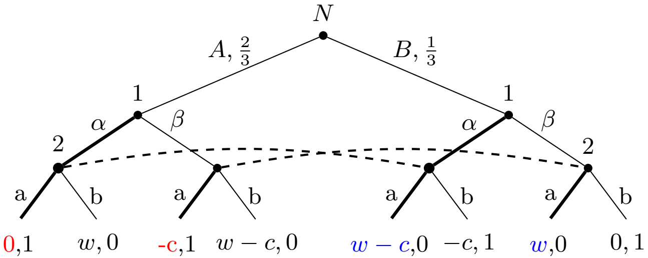
so *one* (of many) weakly consistent belief is

$$b_2 = \left((\tfrac{2}{3}, \tfrac{1}{3}), (\tfrac{2}{3}, \tfrac{1}{3})\right)$$

- then 2's sequentially rational strategy is

$$s_2 = (a|\alpha, a|\beta)$$

# Same message?




- as $0 \geq -c$, 1 has no incentive to switch to $(\beta|A, \alpha|B)$ ✓
- if $w - c \geq w$, 1 has no incentive to switch to $(\alpha|A, \beta|B)$ or $(\beta|A, \beta|B)$
- as $w - c \geq w \iff c \leq 0$
  $\big((\alpha|A, \alpha|B), (a|\alpha, a|\beta); ((\frac{2}{3}, \frac{1}{3}), (\frac{2}{3}, \frac{1}{3}))\big)$ is a PBE if $c = 0$
- this is a *pooling* equilibrium

# Note on readings

- Kydd, chap. 6.2 is slightly more complex version (using continuous strategy spaces) of type-1 games
- Kydd, chap. 6.3 and 6.5 are versions of type-2 games
- none of them actually requires PBE, using subgame perfection is perfectly fine
- type-3 games *do* require PBE; our example foreshadows Kydd, chap. 9 (see week 7 & 8)