LECTURE 5

Equilibrium Refinements in Extensive Form Games

Forward Induction

and

Signaling Games

MAIN POINTS OF PREVIOUS LECTURE

Out of equilibrium information sets

- In sequential games there are **equilibrium paths that do not reach some information sets**: these are the **out-of-equilibrium information sets**
- The optimality conditions of ***Nash equilibria does not constrain behavior at these nodes***, but
  - ***these information sets are out-of-equilibrium because of the actions the players are supposed to play at these nodes***
- In other words,
  - **reaching these nodes in equilibrium is a zero probability event,**
  - **but this probability is *endogeneous*, because is derived from the players' equilibrium behavior.**

Sequential Rationality

- An optimal strategy for a player should maximize his or her payoff, **conditional on every information set at which this player has the move**
- In other words, player i's strategy should specify an "optimal" action from each of player i's information sets, **even those that have zero endogenous probability to be reached**

- **Sequential rationality:**

- apply **some notion of rational behavior** any time you face a well defined decision situation, i.e. in any information set

- This implies that players make threats and promises that they do have an incentive (according to that notion of rational behavior) to carry out, once the information set is reached, even if it had ex ante zero probability.

Sequential rationality in imperfect information games

- ***The idea of Sequential Rationality*:**
  - **Every decision must be part of an optimal strategy for the remainder of the game**
- ***In games with imperfect information*:**
  - **At every decision situation (=information set) the player's subsequent strategy must be optimal with respect to some assessment of the probabilities of all uncertain events, including any preceding but unobserved choices made by other players (Bayesian rationality).**

Construction of a formal definition of sequential rationality: definitions - 1

- A ***system of beliefs*** μ is a specification $\mu_h(x)$ **for each information set h**, where
- **$\mu_h(x) \geq 0$ is the (conditional) probability player i assesses that a node $x \in h \in H_i$ has been reached, GIVEN $h \in H_i$.**
- Therefore $\sum_{x \in h} \mu_h(x) = 1 \qquad \forall h \in H$

Construction of a formal definition of sequential rationality: definitions - 2

- An ***assessment*** *is*

a beliefs-strategies pair (μ, π).

Definition of

SEQUENTIAL RATIONALITY

for imperfect information games

An assessment (μ,π) is ***sequentially rational*** if given the beliefs μ

- **no player *i* prefers at *any* information set $h \in H_i$ to change her strategy π_h^i**
  - In other words,

- **each player's behavior strategy is a best response at *any* information set $h \in H_i$, given her beliefs.**

Formal definition of

SEQUENTIAL RATIONALITY

An assessment (μ,π^*) is ***sequentially rational*** if

$$\forall i \in N, \quad \forall h \in H_i$$

$$\sum_{x \in h} \mu(x) \sum_{z \in Z(x)} v_i(z) P\left(z \mid \pi^*\right) \geq$$

$$\geq \sum_{x \in h} \mu(x) \sum_{z \in Z(x)} v_i(z) P\left(z \mid \pi_i', \pi_{-i}^*\right)$$

$$\forall \pi_i' \in \Pi_i$$

REMARK: sequential rationality requires players to use π* to evaluate the "continuation" probability

Definition of WEAK PERFECT BAYESIAN EQUILIBRIUM

A ***Weak Perfect Bayesian equilibrium*** is an assessment (μ,π) such that

1. Each player is sequentially rational, i.e. each player's **behavior strategy is a best response at *any* information set $h \in H_i$, given her beliefs and opponents' behavior**, i.e.

$$\text{for any } h \in H_i,\ \pi_i(h) \in BR_i(\mu_h, \pi_{-i})$$

2. The **beliefs are derived from the equilibrium strategies** through Bayes' rule whenever possible, i.e.

$$\forall h(x) \text{ such that } \Pr(h(x) \mid \pi) > 0$$

$$\mu_{h(x)}(x) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} \qquad \forall x \in h(x)$$

Effect of sequential rationality for imperfect information games

1. First, it eliminates strictly dominated actions from consideration off the equilibrium path.
2. Second, it **elevates beliefs to the importance of strategies**.

- **This provides a language — the language of beliefs — for discussing the merits of competing sequentially rational equilibria**.
  - ***Where these beliefs come from?***

beliefs are derived from the equilibrium strategies through Bayes' rule

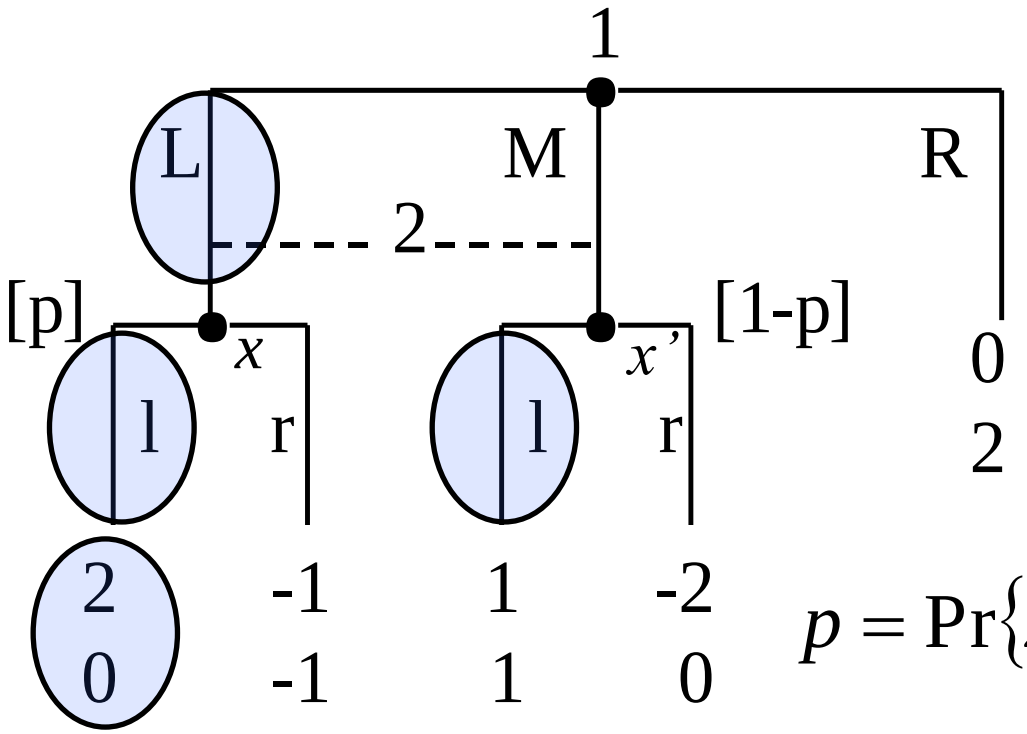
$$\forall h(x) \text{ such that } \Pr(h(x) \mid \pi) > 0$$

$$\mu_{h(x)}(x) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} \qquad \forall x \in h(x)$$

Game 1: how to calculate WPBE.

Start with the first possible NE:

1. $\pi^1(L)=1, \pi^2(l)=1$



[.] denotes a system of beliefs μ.

$$p = \Pr\{x \mid \{x, x'\}\} = \frac{\Pr\{x \mid \hat{\pi}\}}{\Pr\{\{x, x'\} \mid \hat{\pi}\}} =$$

$$= \frac{\pi^1(L)}{\pi^1(L) + \pi^1(M)} = \frac{1}{1+0} = 1$$

Calculus of WPBE in Game 1:

- Strategy l is sequentially rational for *the* system of **belief derived from equilibrium strategies** using Bayes rule:

$$Eu_2(l \mid p=1) = 0\times 1 + 1\times 0 = 0 > Eu_2(r \mid p) = -1\times 1 + 0\times 0 = -1$$

And L is a best reply for player 1 to l

$$Eu_1(L,l) = 2 > Eu_1(M,l) = 1$$

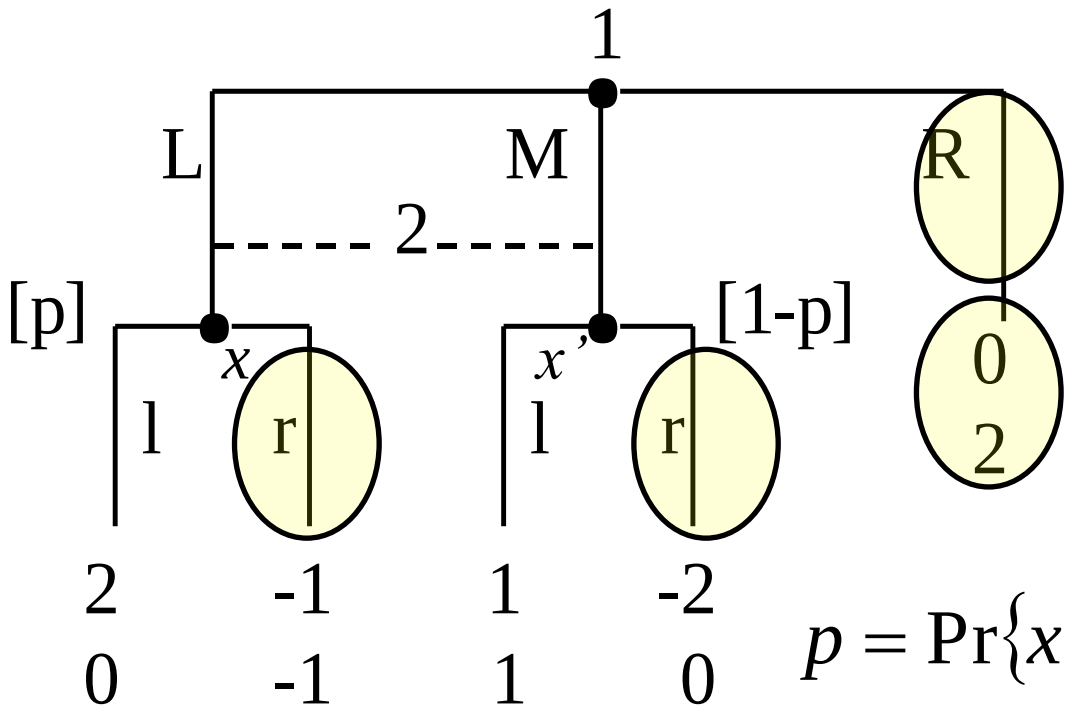
$$Eu_1(L,l) = 2 > Eu_1(R,l) = 0$$

Therefore (L,l),p=1 is a WPBE

Game 1: how to calculate WPBE.

Then consider the second possible NE:

2. $\pi^1(R)=1, \pi^2(r)=1$



[.] denotes a system of beliefs μ.

$$p = \Pr\{x \mid \{x, x'\}\} = \frac{\Pr\{x \mid \hat{\pi}\}}{\Pr\{\{x, x'\} \mid \hat{\pi}\}} =$$

$$= \frac{\pi^1(L)}{\pi^1(L) + \pi^1(M)} = \frac{0}{0+0} \in [0,1] = p$$

Game 1:

- Strategy r is not sequentially rational for ***any*** **possible system of belief:**

$$Eu_2(l \mid p) = 0 \times p + 1 \times (1-p) > -1 \times p + 0 \times (1-p) = Eu_2(r \mid p)$$

$$\Updownarrow$$

$$Eu_2(l \mid p) = (1-p) > -p = Eu_2(r \mid p) \Leftrightarrow Eu_2(l \mid p) = 1 > 0 = Eu_2(r \mid p)$$

- This is how weak perfect Bayesian equilibrium prevents strictly dominated strategies from being used as threats off the equilibrium path: they are not sequentially rational for any possible system of beliefs.

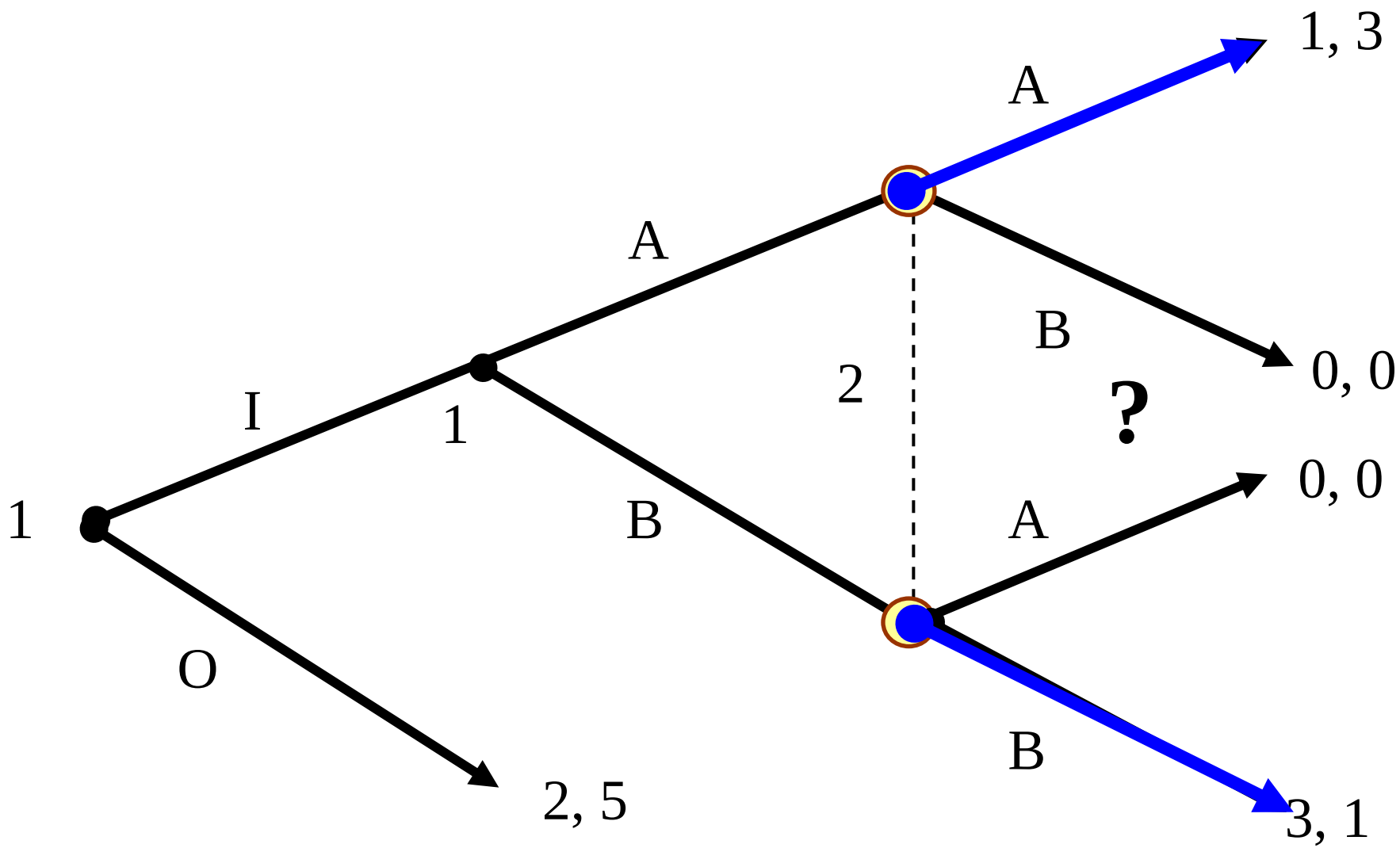
WPBE in Perfect Information Games: Backward Induction

Backward Induction

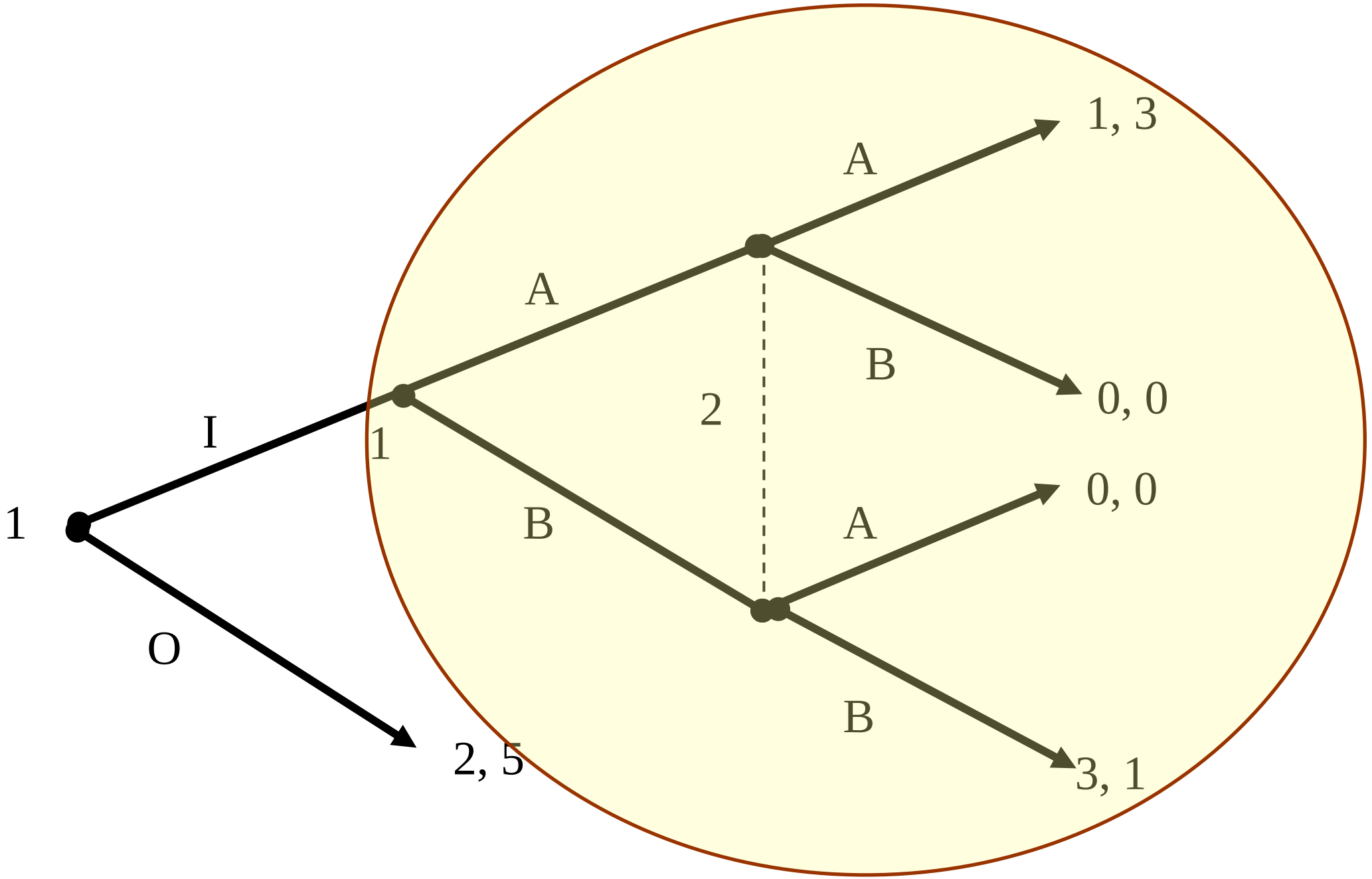
- **Backward Induction** if
  1. **Rationality means to avoid strictly dominated actions**, and
  2. **Sequential Rationality is common knowledge**
- Practically **Backward induction** is the process of analyzing a game from back to front, from information sets at the end of the tree to information sets at the beginning
- At each information set, one strikes from considerations actions that are dominated, **given the terminal nodes that can be reached and that will be reached according to backward induction.**

A PROBLEM WITH BACKWARD INDUCTION IN IMPERFECT INFORMATION GAMES

EXAMPLE WHERE BACKWARD INDUCTION DOESN'T WORK



WHEN BACKWARD INDUCTION DOESN'T WORK
USE SUBGAMES



Subgame Perfection - 1 (Selten, 1965)

- The concept of sequential rationality can be expanded to cover general extensive form games:
  - Apply Nash equilibrium any time you face a well defined strategic situation
  - The notion of subgame is the formal translation of "a well defined strategic situation"

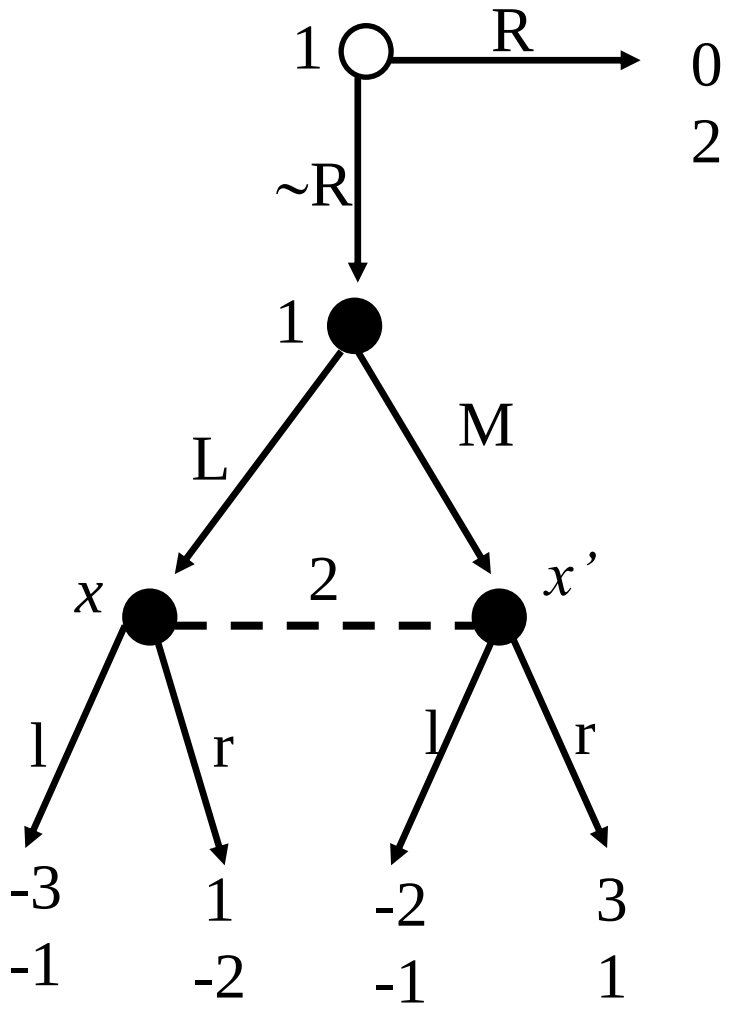
THE PROBLEMS WITH WPBE AND THE NOTION OF SEQUENTIAL EQUILIBRIUM

Game 1: comparing NE, SPE and WPBE

	l	r
RL	0, 2	0, 2
RM	0, 2	0, 2
~RL	-3, -1	1, -2
~RM	-2, -1	3, 1

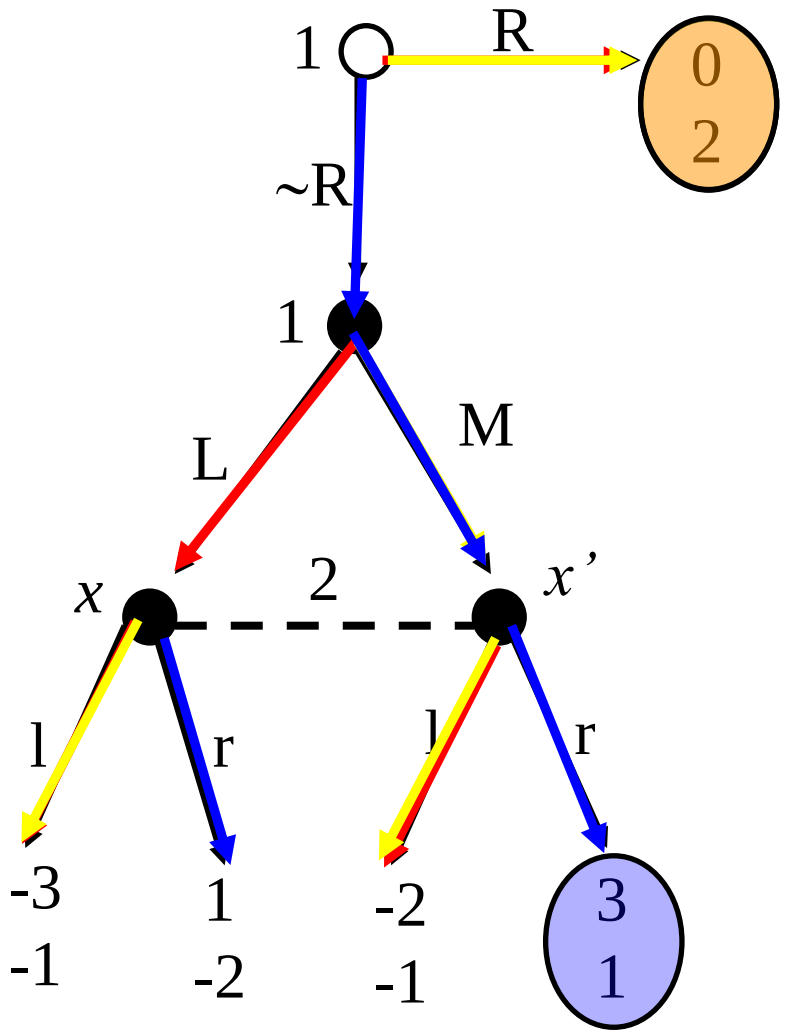
First,
the pure strategy
Nash Equilibria
of game 1

Three pure strategy Nash Equilibria:
(RL,l), (RM,l),
(~RM,r)



Game 1: comparing NE, SPE and WPBE

	l	r
RL	0, 2	0, 2
RM	0, 2	0, 2
~RL	-3, -1	1, -2
~RM	-2, -1	3, 1



Game 1: WPBE

Two WPBE:

1. (~RM,r),

$$\mu(x' \mid h(x)) = 1$$

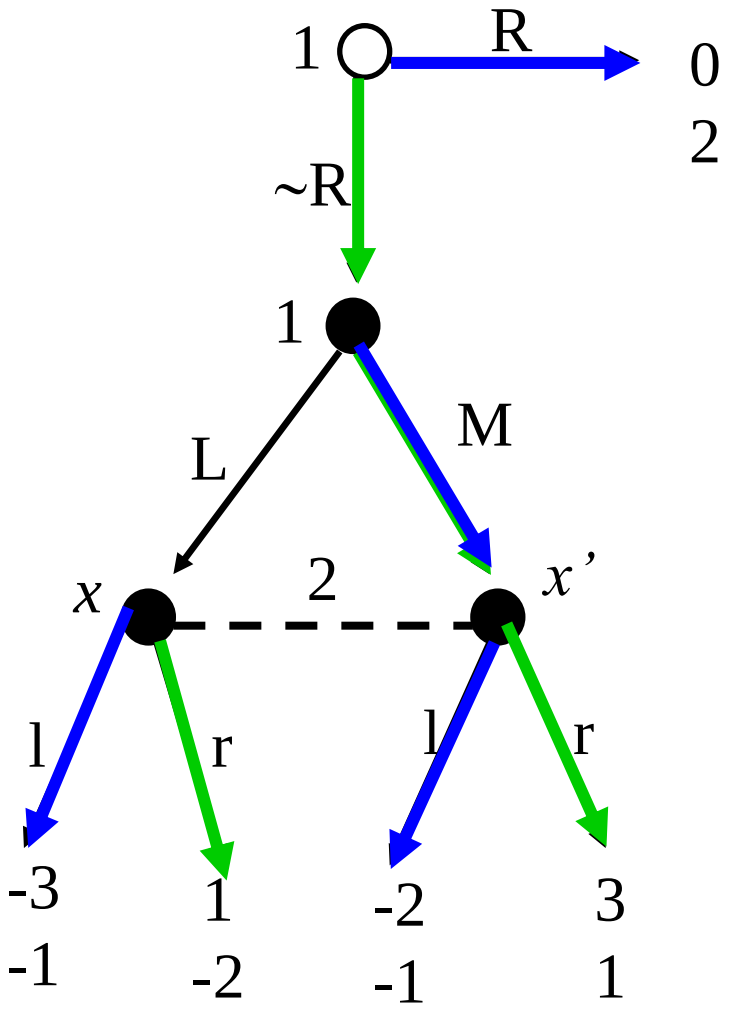
2. (RM,l)

$$\mu(x \mid h(x)) = 1$$

2. (RL,l) is not WPBE

Because L is not s

equentially rational



Game 1: calculating beliefs for WPBE

Deriving beliefs through

Bayesian rule from playing ¬R:

$$\mu(x \mid h(x)) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} =$$

$$= \frac{\pi_1(\neg R) \times \pi_1(L)}{\pi_1(\neg R) \times \pi_1(L) + \pi_1(\neg R) \times \pi_1(M)} =$$

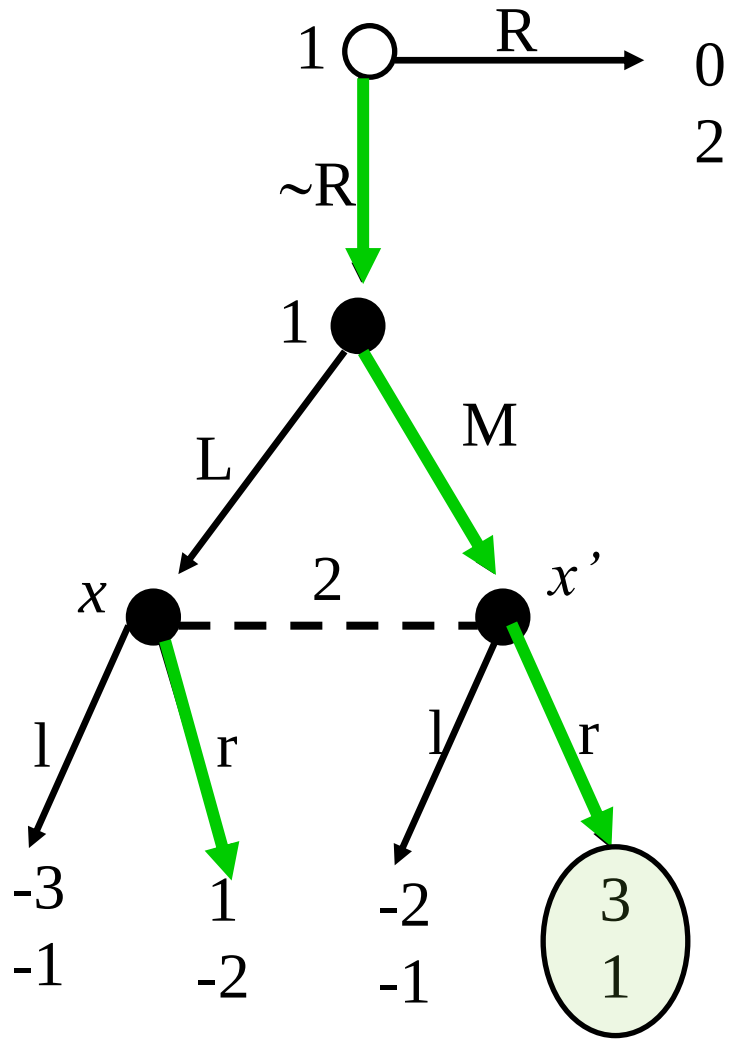
$$= \frac{1 \times 0}{1 \times 0 + 1 \times 1} = 0$$

$$\therefore \mu(x' \mid h(x)) = 1, \text{ then}$$

r is a best reply at $\{x, x'\}$

M is a best reply to r

and $\sim R$ is a best reply to M, r



Game 1: deriving beliefs for a WPBE

Deriving beliefs through

Bayesian rule from playing R:

$$\mu(x \mid h(x)) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} =$$

$$= \frac{\pi_1(\neg R) \times \pi_1(L)}{\pi_1(\neg R) \times \pi_1(L) + \pi_1(\neg R) \times \pi_1(M)} = \frac{0}{0}$$

$$\therefore \mu(x \mid h(x)) \in [0,1]$$

Note that we can't semplify the ratio for $\pi_1(\neg R)$

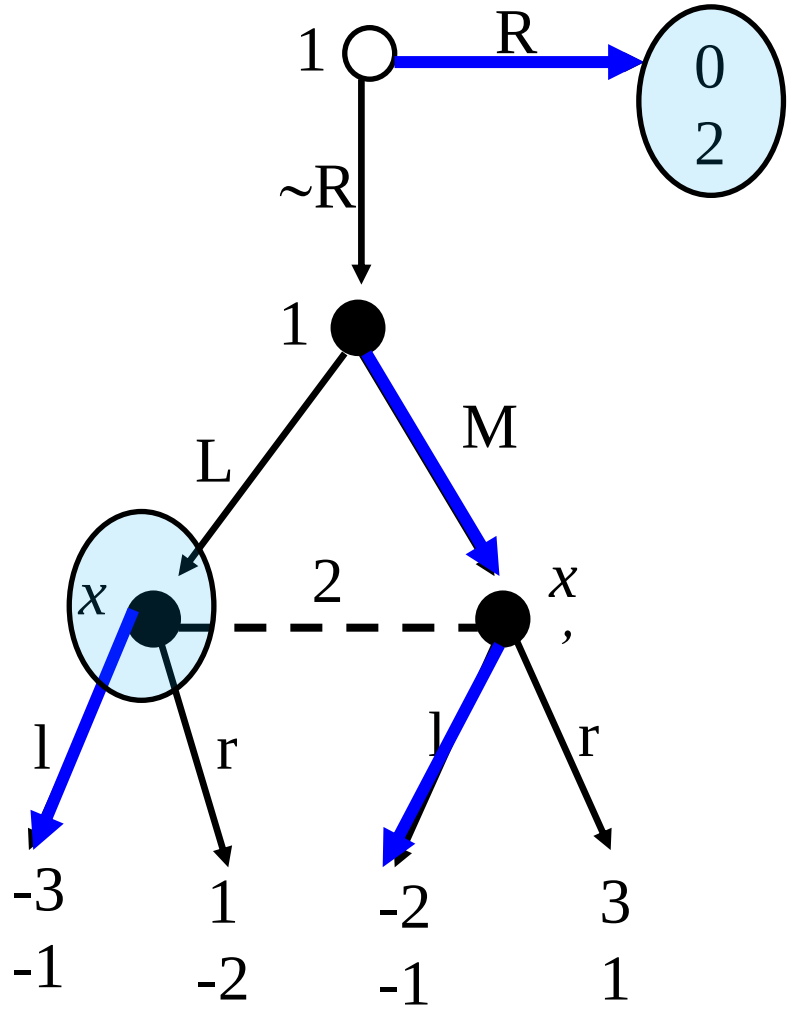
because $\pi_1(\neg R) = 0.$

Suppose $\mu(x \mid h(x)) = 1,$ then

l is a best replies

M is a best reply to l

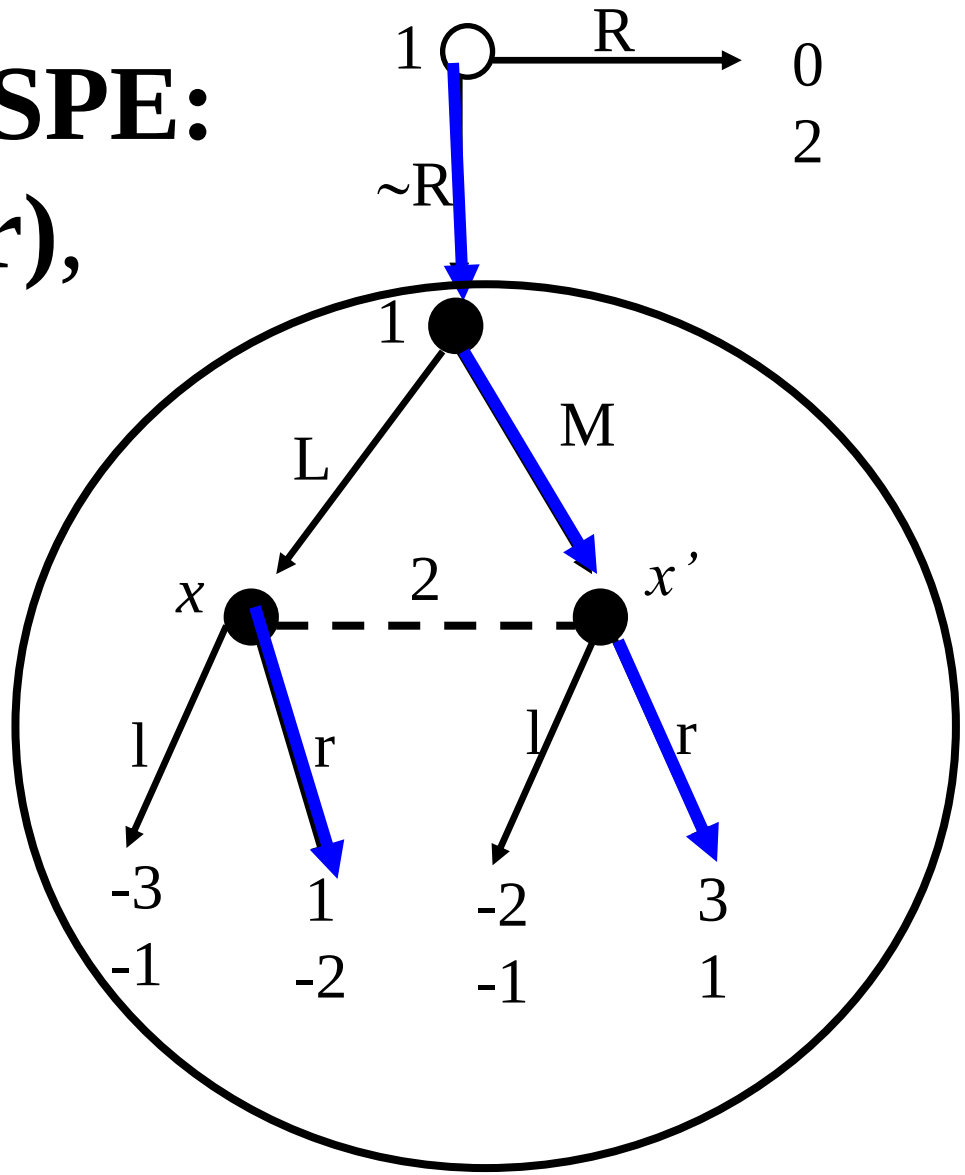
and R is a best reply to M, l



Game 1: applying SPE

A unique SPE: (~RM,r),

Problem:
A WPBE need not be subgame perfect



Game 1: discussing beliefs for a WPBE

Deriving beliefs through

Bayesian rule from playing R:

$$\mu(x \mid h(x)) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} =$$

$$= \frac{\pi_1(\neg R) \times \pi_1(L)}{\pi_1(\neg R) \times \pi_1(L) + \pi_1(\neg R) \times \pi_1(M)} = \frac{0}{0}$$

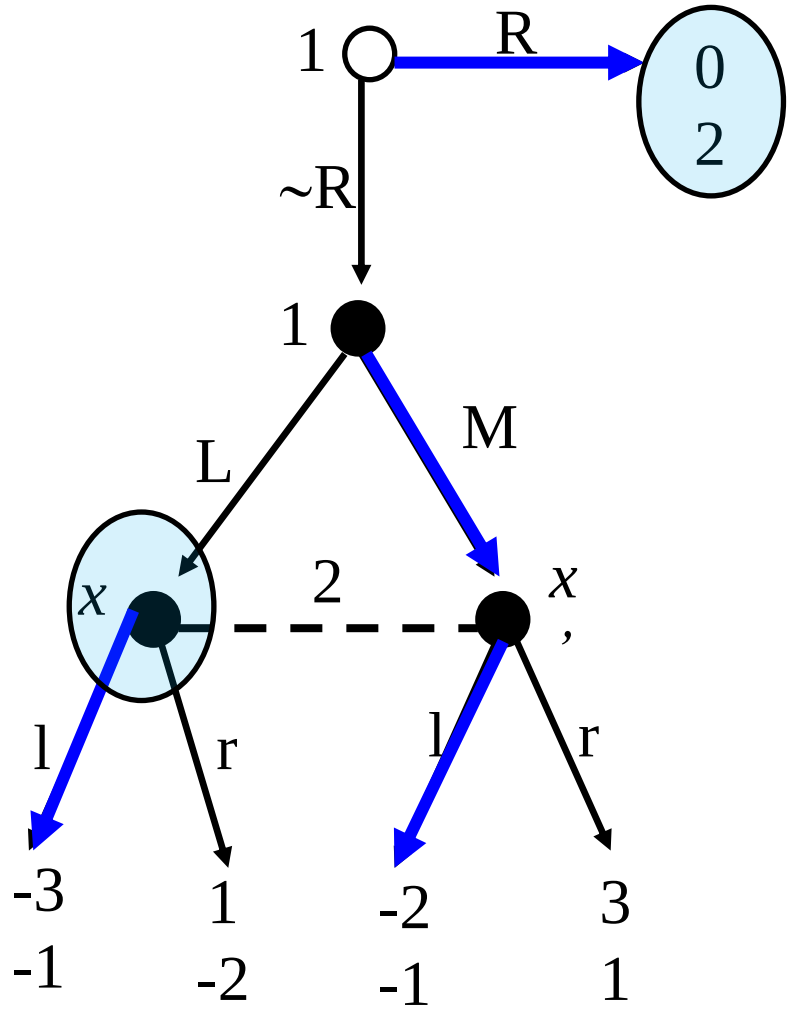
$$\therefore \mu(x \mid h(x)) \in [0,1]$$

Suppose $\mu(x \mid h(x)) = 1$, then

l is a best replies

M is a best reply to l

and R is a best reply to M, l



- What is the meaning of $\mu(x|h(x)) = 1$?
- It means that $\pi_1(\sim R) \times \pi_1(L)$ is infinitely more likely than $\pi_1(\sim R) \times \pi_1(M)$. Is it plausible?

Refining the notion of Weak Perfect Bayesian Equilibrium

- To solve the previous problem we try to refine the notion of WPBE, using **totally mixed strategies** and defining **SEQUENTIAL EQUILIBRIA.**
- A strategy profile π is *totally mixed* if it assigns strictly positive probability to each action $a \in A(h)$ for each information set $h \in H$.

Definition: **Consistency**

Definition:

- An assessment (μ,π) is ***consistent*** if

1. there exists a sequence of totally mixed behavioral strategies π_n and
2. corresponding beliefs μ_n derived from Bayes' rule

such that

$$\lim_{n\to\infty}(\mu_n,\pi_n)=(\mu,\pi).$$

Definition of **SEQUENTIAL EQUILIBRIUM**

- A ***sequential equilibrium*** is an assessment (μ,π) that is both

1. *sequentially rational* and
2. *consistent*.

Game 2: deriving beliefs with consistency

Deriving consistent beliefs through Bayesian rule from playing RM,l:

$$\mu(x \mid h(x)) = \frac{\Pr(x \mid \pi)}{\Pr(h(x) \mid \pi)} =$$

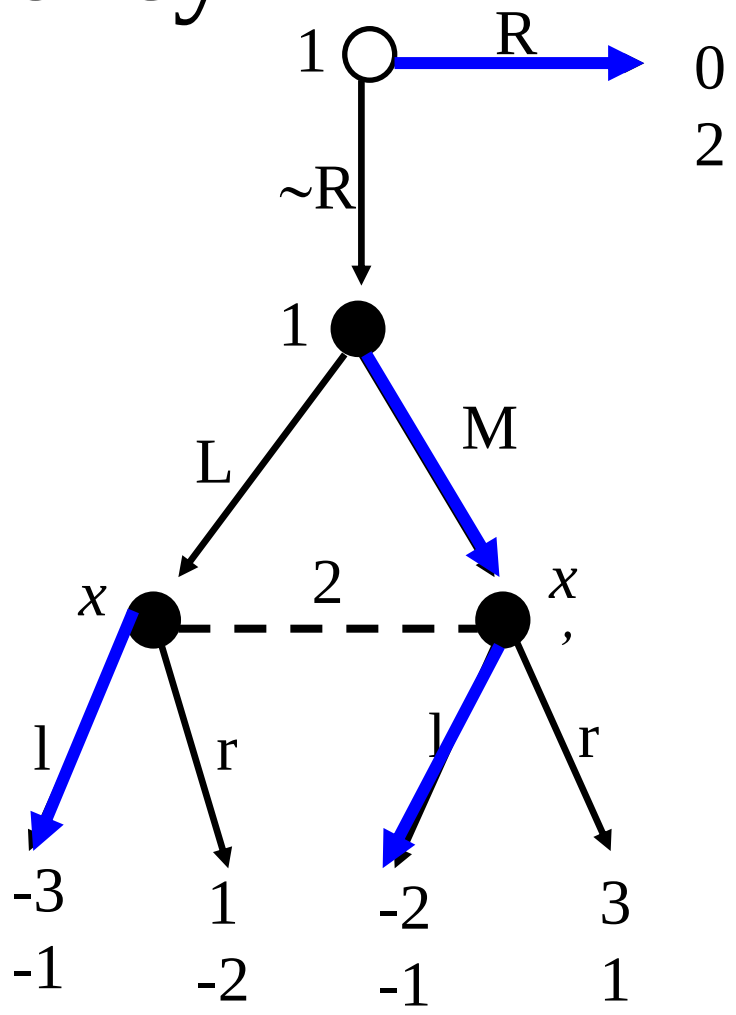
$$= \frac{\pi_1(\neg R) \times \pi_1(L)}{\pi_1(\neg R) \times \pi_1(L) + \pi_1(\neg R) \times \pi_1(M)} =$$

$$\frac{\varepsilon \times \eta}{\varepsilon \times \eta + \varepsilon \times (1-\eta)} = \frac{\eta}{\eta + 1 - \eta} \xrightarrow[\eta \to 0]{} 0$$

$$\therefore \mu(x \mid h(x)) = 0$$

then M, l are NOT best replies

the unique SE in pure strategies is $(\neg RM, r)$ which is Subgame Perfect



Meaning of SEQUENTIAL EQUILIBRIA

- In a SE any equilibrium strategy is approximated by a totally mixed strategy
- Because of this, any information set is reached with strictly positive probability possibly vanishing
- This means that out of equilibrium information sets are reached with small vanishing probabilities, i.e. **by mistakes**:

impossible events are explained as due to trembling hands.

Theorem

For every finite extensive-form game there exists at least one sequential equilibrium. Also, if (μ,π) is a sequential equilibrium then π is a subgame-perfect Nash equilibrium.

$$SE_{\pi} \subseteq WPBE_{\pi} \subseteq NE$$

Moreover

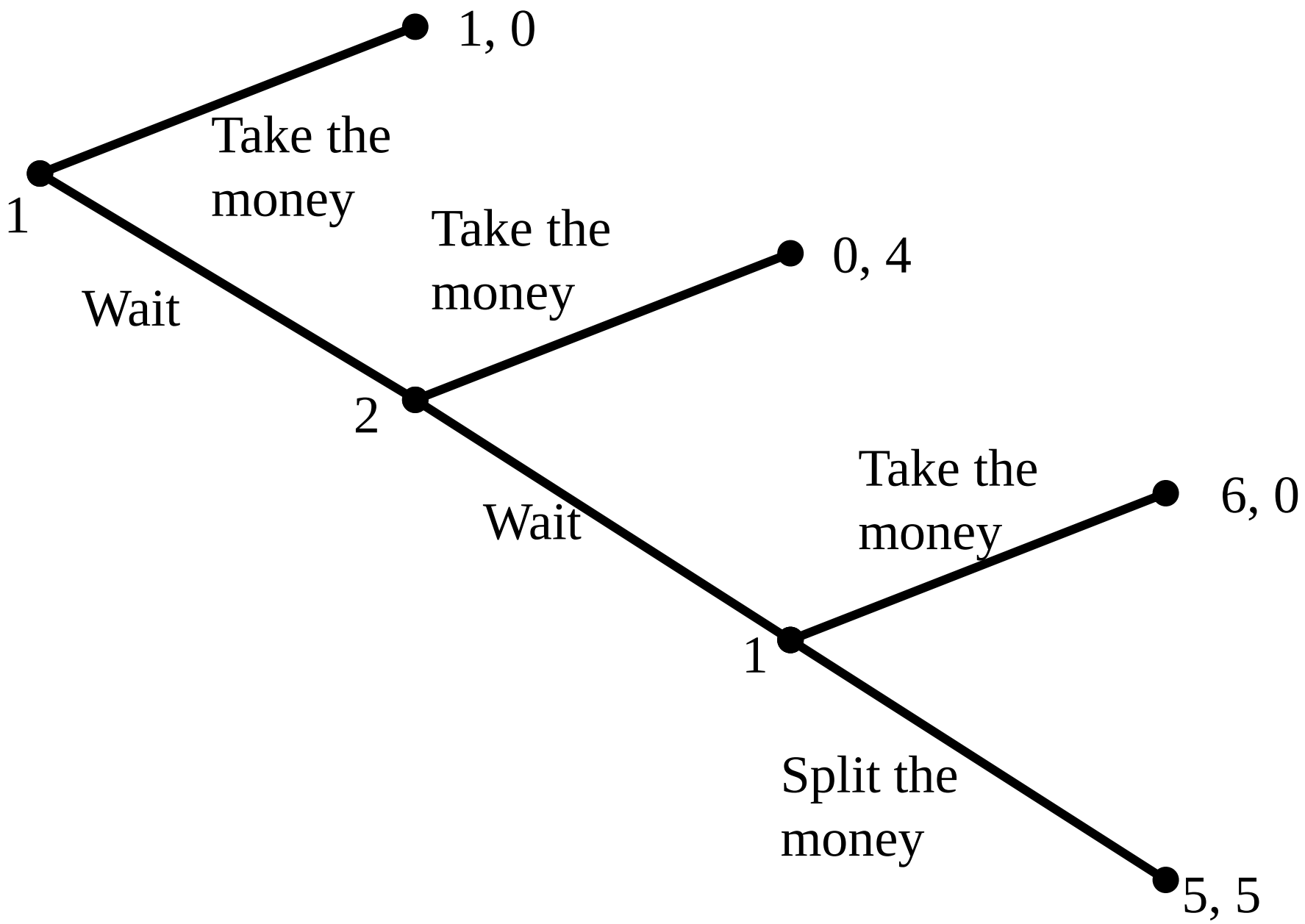
$$SE \neq \emptyset$$

NEW CONCEPTS TO MODEL STRATEGIC INTERACTION

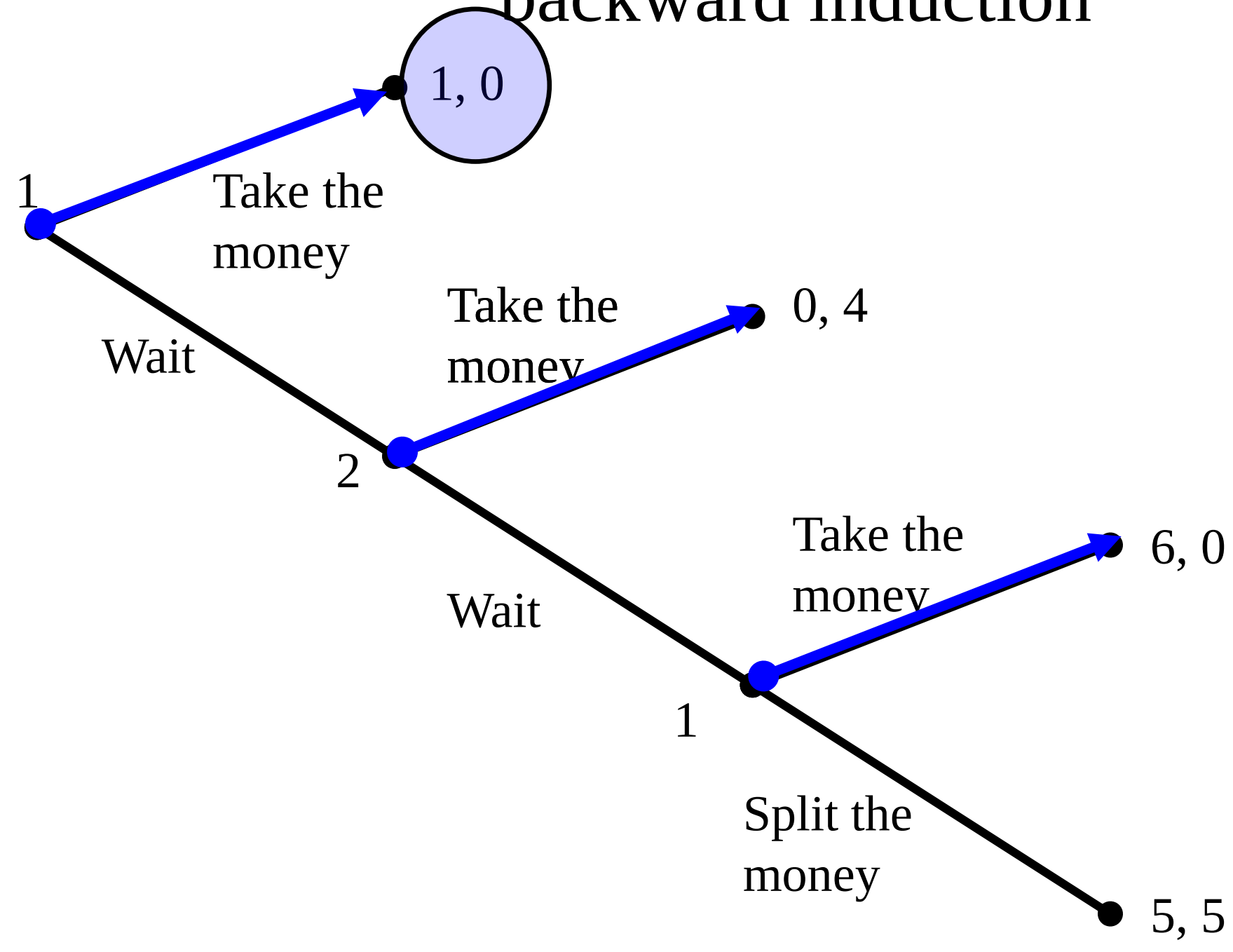
PROBLEMS WITH SEQUENTIAL RATIONALITY: THE CASE OF BACKWARD INDUCTION

Common knowledge of rationality at "irrational" information sets

Centipede: a simplified version



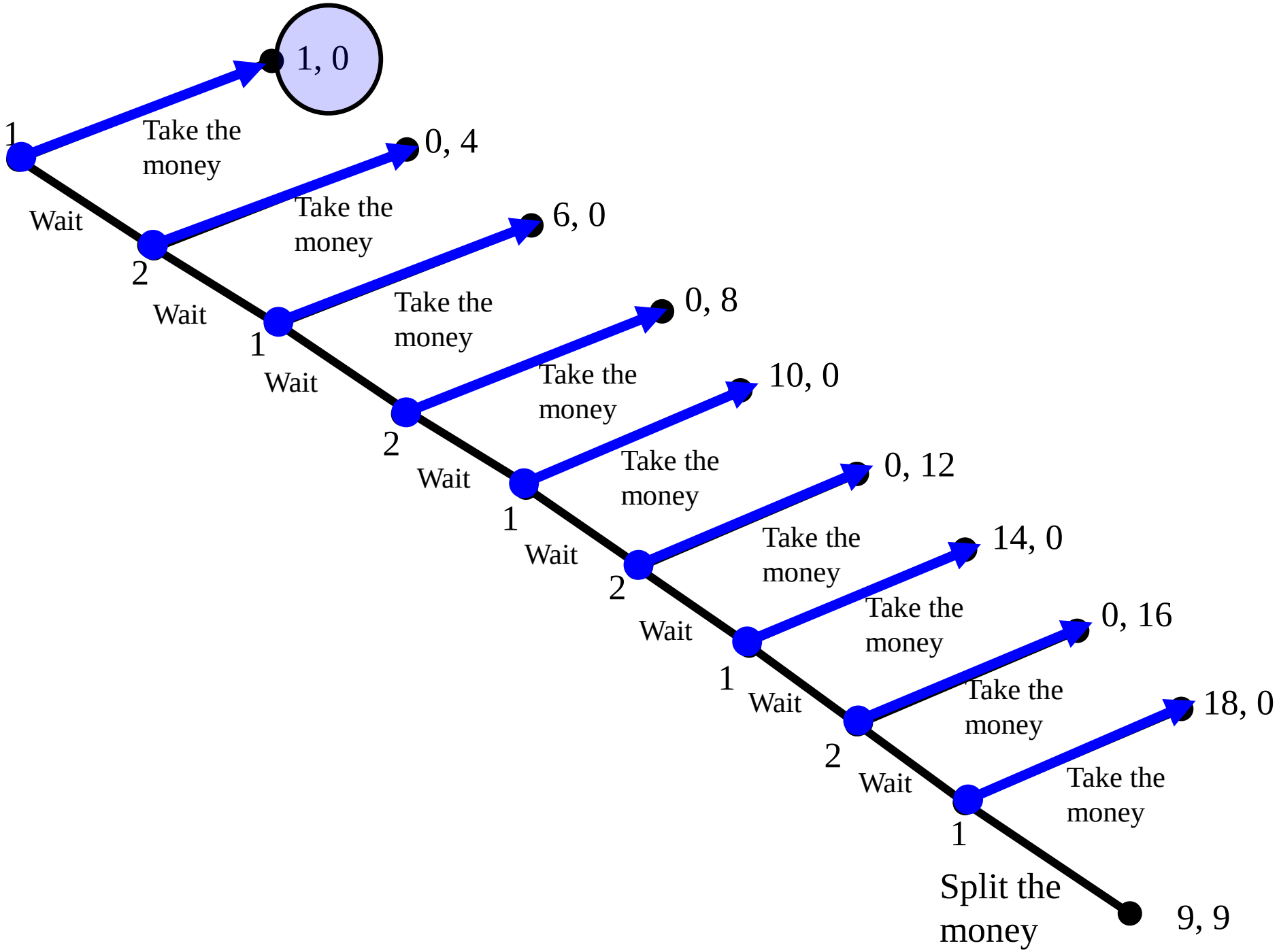
Centipede: the unique equilibrium by backward induction



PROBLEM AT "IRRATIONAL" INFORMATION SETS

- Suppose to be player 2 at your information set: if you are called to play, it means that player 1 has played WAIT,
- but you know that WAIT is **irrational for player 1.**
- **Maybe 1 has done a mistake …**
- **But then, if 1 has done an irrational move, why will you predict a rational moves at next 1's decision nodes?**
- **If you believe that 1 is "irrational", you can try to go for 5 instead of 3**
- **But then it is rational to play wait-wait …**

Centipede: the probability of mistakes



PROBLEM AT "IRRATIONAL" INFORMATION SETS

- Suppose to be player 2 at your first information set: if you are called to play, it means that player 1 has played WAIT,
- but you know that WAIT is irrational for player 1.
- Maybe 1 has done a mistake with probability ε
- Then you are player 1 at your second information set: if you are called to play, it means that you has done a mistake (probability ε) and 2 also, with probability δ
- Then you are player 2 at your second information set: if you are called to play, it means that 1 have done two mistakes (probability ε^2) and you also, with probability δ
- Then you are player 1 at your third information set: if you are called to play, it means that you have done two mistakes (probability ε^2) and 2 also, with probability δ^2
- Then you are player 2 at your third information set: if you are called to play, it means that 1 has done three mistakes (probability ε^3) and 2 also, with probability δ^2
- Then you are player 1 at your fourth information set: if you are called to play, it means that you have done three mistakes (probability ε^3) and 2 also, with probability δ^3
- Then you are player 2 at your fourth information set: if you are called to play, it means that 1 has done four mistakes (probability ε^4) and 2 also, with probability δ^3
- Then you are player 1 at your fifth information set: if you are called to play, it means that you have done four mistakes (probability ε^4) and 2 also, with probability δ^4
- ….

PROBLEM AT "IRRATIONAL" INFORMATION SETS

- **The logic of backward induction/subgame perfection/WPBE requires sequential rationality, i.e. *rationality to be common knowledge at every node of the extensive form*,**
- **but this might be counterintuitive because these nodes can be reached only because of IRRATIONAL moves**
- as shown in previous example
- Backward induction/subgame perfection/WPBE may be justified by rationality **and** small possibility of mistakes, **repeatedly at each information sets**
- **But**
- rationality **and** small possibility of mistakes, **repeatedly at each information sets might be seen as mutually contradictory**

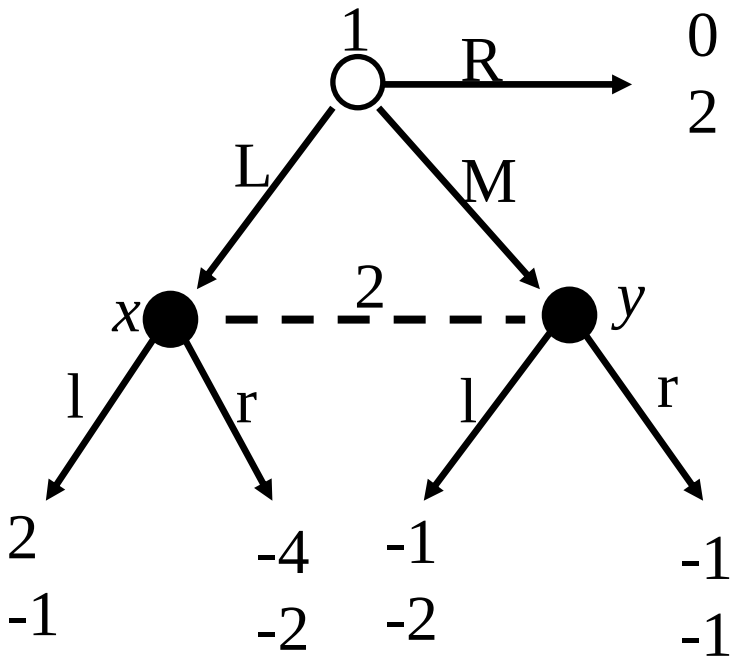
PROBLEMS WITH WPBE AND WITH SE: "IRRATIONAL" DEVIATION

Ordering of plausibility of equilibrium deviations:

Forward Induction

Game 1: plausible and implausible WPBE

First: calculate NE

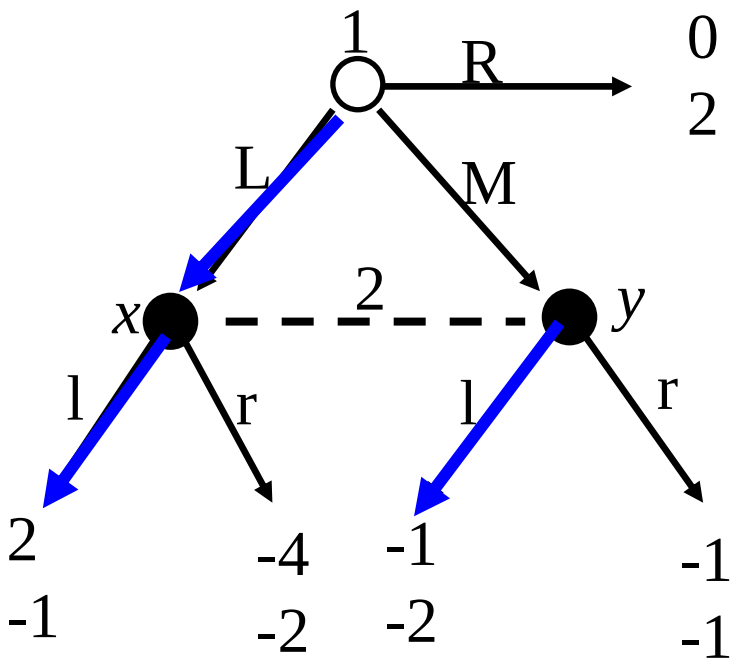


	l	r
L	2, -1	-4, -2
M	-1, -2	-1, -1
R	0, 2	0, 2

Two NE: (L, l) and (R,r)

Game 1: plausible and implausible WPBE

Check whether NE are WPBE:



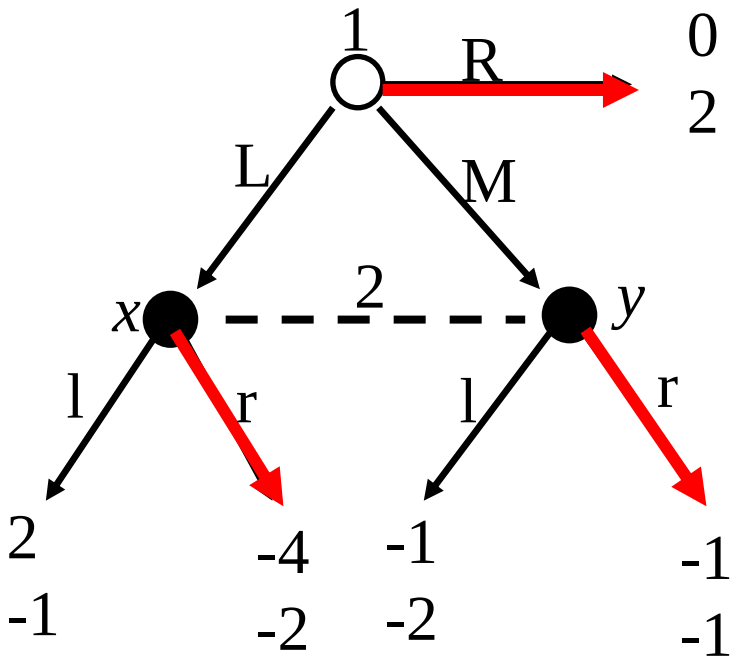
- First NE: (L,l)
- (L,l) and $\mu(x|\{x,y\})=1$ is a WPBE:
- Bayes rule and equilibrium strategies implies the beliefs at player 2's information set to be that player 1 played L with probability one.

$$\pi_1(L)=1,\ \ \pi_1(M)=\pi_1(R)=0$$

$$\mu(x\,|\,\{x,y\})=\frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)}=\frac{1}{1+0}=1$$

Game 1: plausible and implausible WPBE

Check whether NE are WPBE:



- Second NE (R,r):
- (R,r) and $\mu(y|\{x,y\})=1$ is a WPBE:
- Since player 2's information set is not reached in equilibrium, then 2's beliefs that 1 played M with probability $q \geq 1/2$ are possible.

$$\pi_1(L)=0,\ \ \pi_1(M)=0, \pi_1(R)=1$$

$$\mu(x\,|\,\{x,y\})=\frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)}=\frac{0}{0+0}\in[0,1]$$

$$\text{Hence}\ \ \mu(y\,|\,\{x,y\})=\frac{\pi_1(M)}{\pi_1(L)+\pi_1(M)}=1\Leftrightarrow \pi_1^C(M)=1$$

Game 1

Problem: **implausible beliefs**

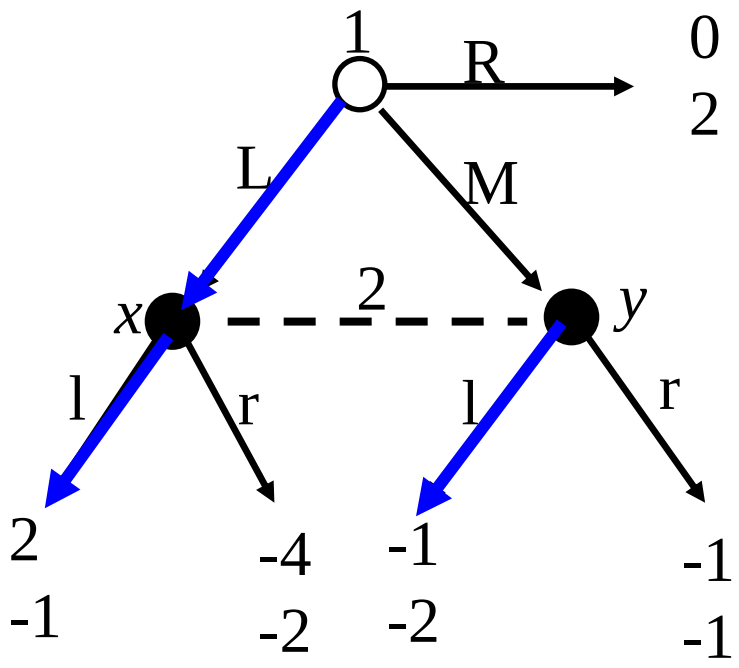
- (R,r) and $\mu(y|\{x,y\})=1$ is implausible, **since $\mu(y|\{x,y\})=1$ means that player 2 is conjecturing that 1 has played M, which is strictly dominated by R.**
- Player 2 is using *implausible beliefs* rather than *incredible actions* to threaten player 1, and the threat succeeds in making 1 play R.
- **The problem is that Bayes rule does not restrict beliefs out of the equilibrium path and the equilibrium path depends on beliefs**

Does Sequential equilibria solve this problem?

–To try to solve the previous problem apply **SEQUENTIAL EQUILIBRIA.**

Game 1: totally mixed strategy and beliefs

First WPBE: (L,l) and μ(x|{x,y})=1



Consider the totally mixed strategy specified below

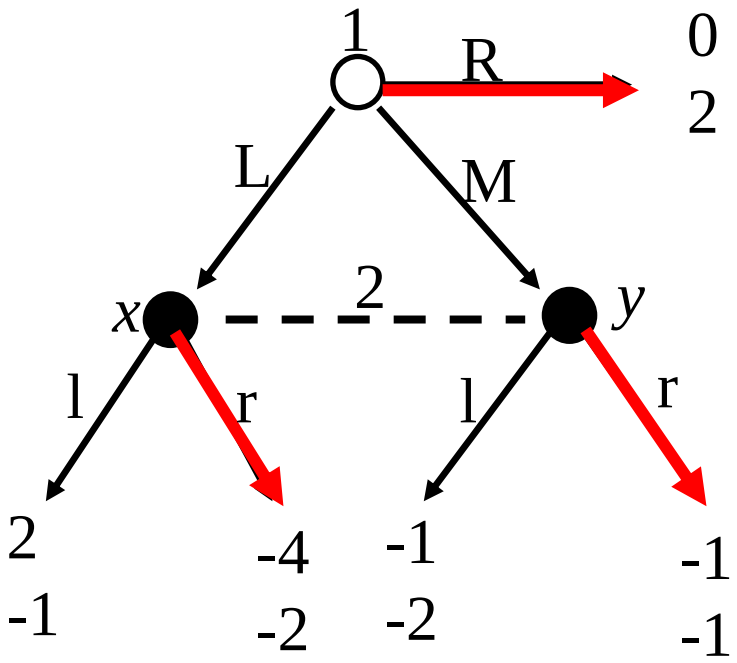
$$\pi_1(L) = 1-\varepsilon-\delta, \quad \pi_1(M) = \varepsilon, \quad \pi_1(R) = \delta$$

$$\mu\left(x \mid \{x,y\}\right) = \frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)} = \frac{1-\varepsilon-\delta}{1-\varepsilon-\delta+\varepsilon} \xrightarrow[\varepsilon\to 0,\delta\to 0]{} 1$$

Game 1: totally mixed strategy and beliefs

- Second WPBE: (R,r) and $\mu(y|\{x,y\})=1$

Consider the totally mixed strategy specified below, with $p \in [0;1]$



$$\pi_1(R) = 1 - \frac{1}{n};$$

$$\pi_1(L) = \varepsilon = \frac{p}{n};$$

$$\pi_1(M) = \delta = \frac{1-p}{n}$$

$$\mu\left(x \mid \{x,y\}\right) = \frac{\Pr\left(x \mid \pi_1\right)}{\Pr\left(\{x,y\} \mid \pi_1\right)} = \frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)} = \frac{\frac{p}{n}}{\frac{p}{n}+\frac{1-p}{n}} = p \in [0,1]$$

Game 1

Two sequential equilibria:

1. (L,l) and $\mu(x|\{x,y\})=1$: Bayes rule forces the beliefs at player 2's information set to be that player 1 played L with probability one.
2. (R,r), $\mu(x|\{x,y\})=0$: player 2's information set is not reached in equilibrium and we have shown that **even consistency doesn't bind in this game**

- The beliefs of this second SE means that 2 believes that 1 played M with probability $(1\text{-}p) \geq ½$, which might be considered not plausible.

Game 1: Forward Induction

$$\mu\left(x \mid \{x,y\}\right) = \frac{\Pr\left(x \mid \pi_1\right)}{\Pr\left(\{x,y\} \mid \pi_1\right)} = \frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)} = 0 \Leftrightarrow$$

$$\Leftrightarrow \pi_1(M) \text{ is more likely than } \pi_1(L)$$

- Note that the reason 2's beliefs are implausible is that
  1. **Player 1 gets a payoff of 1 from playing the equilibrium strategy R**
  2. **This is superior to any payoff 1 could receive by playing M. And yet when 2's information set is reached 2 believes with probability at least 1/2 that M has been played**.
- Instead, **2 *<u>might</u>* reason that 1 would not deviate unless there was something to be gained**, so L must have been played with probability close to one, in which case l rather than r is the BR.

DIFFERENT REFINEMENTS AND DIFFERENT EXPLANATIONS OF DEVIATIONS

SIMPLE MISTAKES

- The simplest explanations of a deviation from the equilibrium path is just a simple mistake:
  - **One holds to the hypothesis that all players intend to follow the prescription of the equilibrium, but that they sometimes fail**
- **One obtains useful restrictions on out-of-equilibrium beliefs only insofar as one is willing to attribute relative likelihood to particular mistakes.**

SIMPLE MISTAKES (1)

- Suppose one makes the following hypothesis:
  1. Mistakes are unlikely,
  2. every mistakes is possible
  3. the chances of mistakes are independent by
     1. different players and
     2. **information sets of the same player**

- then you get **trembling hand perfect equilibria in extensive form which is *almost* equivalent to Sequential Equilibria.**

SIMPLE MISTAKES (2)

- Suppose one makes the previous hypothesis:
  1. Mistakes are unlikely,
  2. every mistakes is possible
  3. the chances of mistakes are independent by different players, but
  4. **mistakes are made on entire strategies**

- then you get **trembling hand perfect equilibria in normal form.**

SIMPLE MISTAKES (3)

- Suppose one makes the previous hypothesis:
  1. Mistakes are unlikely,
  2. every mistakes is possible
  3. the chances of mistakes are independent by different players,
  4. **mistakes are made on entire strategies**
  5. **worse mistakes are less likely to be made than those that are less worse.**

- then you get **proper equilibria, which includes a sort of *Forward Induction* criterion, i.e. mistakes are ranked in terms of degree of rationality.**

MISTAKEN THEORIES (1)

- Deviations from equilibrium play may be explained by the fact that one or more players does not understand what is expected of him.
- One would then look for relatively likely alternative theories for how to play the game to explain
  1. Who has defected
  2. What has been the nature of defection
  3. What might be the consequences of that defection for later play.

- **Structural consistency** is a way of formalizing this type of reasoning, but …

MISTAKEN THEORIES (2)

- this reasoning can lead to direct attack to Sequential Equilibrium, in particular to the hypothesis that
- player countenance no further deviations from the equilibrium when evaluating what to do in the face of an apparent deviation,
- for example if after a deviation one believes that the error in theory may be one's own, then **deviations among different players may be thought to be correlated**.

CONSCIOUS SIGNALS

- Deviations might be **interpreted as the conscious attempt of a player to signal something to others**
- The credibility of this signal is based on comparisons with payoffs obtained in the purported equilibrium.
- This idea is behind the notion of **forward induction** and is the basic motivations of many effective refinements in the signaling contexts.

WAYS OF FORMALIZING THE FORWARD INDUCTION IDEA

FORWARD INDUCTION EQUILIBRIA:

BAD DEVIATIONS

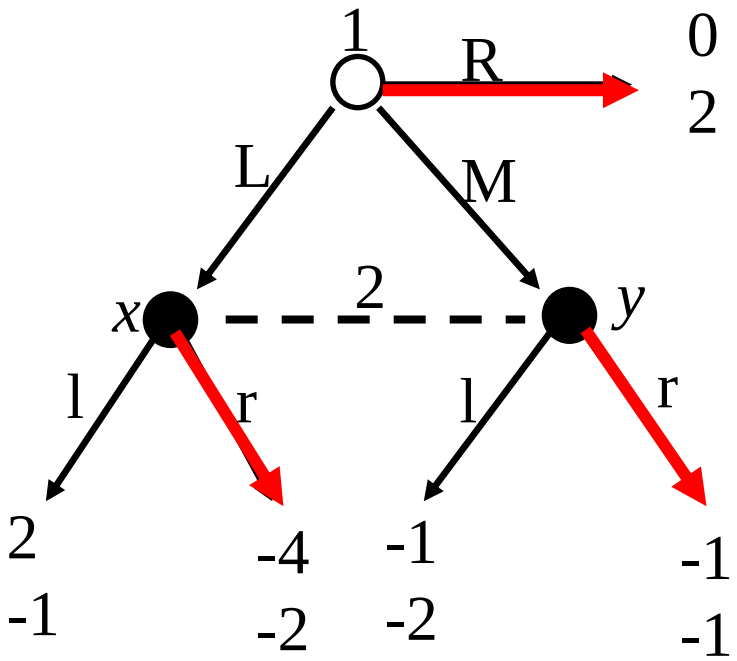
Forward Induction and Sequential Equilibria

- For a general game, a deviation from a specified equilibrium is said to be "bad" if it *always* yields the deviator less than her equilibrium payoff in *every* circumstance.

Game 1: totally mixed strategy and beliefs

- (R,r) and $\mu(y|\{x,y\})=p\in[0,1]$:
- Since player 2's information set is not reached in equilibrium, then 2's beliefs that 1 played M with probability $q \geq 1/2$ are possible.

Consider the totally mixed strategy specified below, with $p \in [0;1]$



$$\pi_1(R)=1-\frac{1}{n};\quad \pi_1(L)=\frac{p}{n};\quad \pi_1(M)=\frac{1-p}{n}$$

$$\mu\left(x|\{x,y\}\right)=\frac{\Pr\left(x|\pi_1\right)}{\Pr\left(\{x,y\}|\pi_1\right)}=\frac{\pi_1(L)}{\pi_1(L)+\pi_1(M)}=\frac{\frac{p}{n}}{\frac{p}{n}+\frac{1-p}{n}}=p\in[0,1]$$

Hence any p ∈ [0, 1] is a **consistent belief**

Forward Induction and Bad deviations

- The SE (R,r) seems unreasonable because it requires player 2 to believe with high probability that player 1 has made a **bad deviation** from the equilibrium.
  - M would be a bad deviation for 1.
- Thus, (M,r) is **not a forward induction equilibrium because it can be supported only by beliefs that assess positive probability that a**

bad deviation

has occurred.

Forward Induction and Bad Deviations

Limitations:

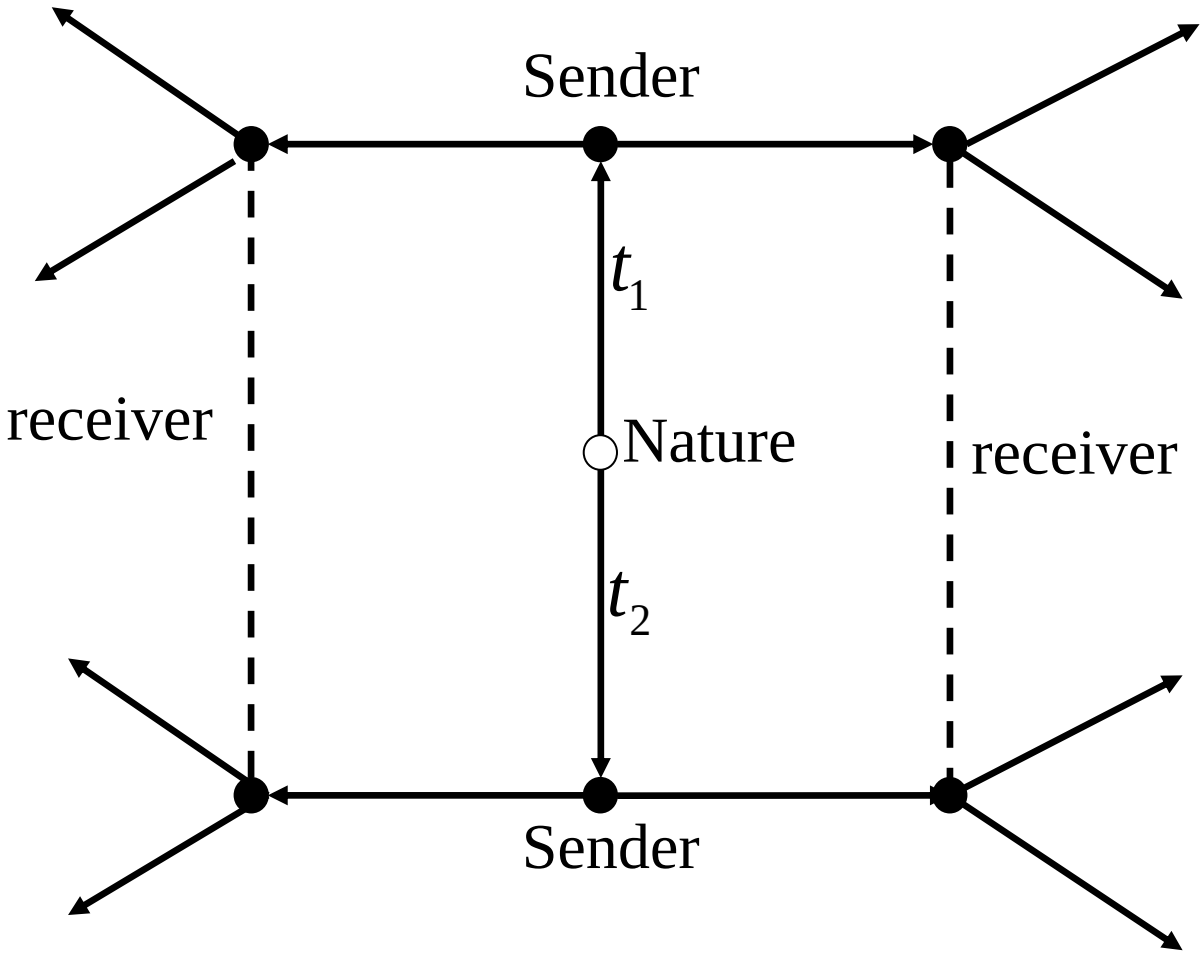
In more complex games the set of bad deviations often is empty, in which case every sequential equilibrium is a forward induction equilibrium, and we must resort once again to ad hoc arguments to capture forward induction.

SIGNALING GAMES

The general structure of Signaling Games

- Two players: a Sender (S) and a Receiver (R).
- The timing of the game is:
  - (1) nature draws a type for S, denoted $t \in T$, according to the commonly known probability distribution p(t);
  - (2) S privately observes the type t and then sends the message $m \in M$ to R; and
  - (3) R observes m and then takes the action $a \in A$.
- **SIMPLIFICATION:** T, M, and A are all finite.
- Payoffs are $U^S(t,m,a)$ and $U^R(t,m,a)$.
- Everything but t, is common knowledge.

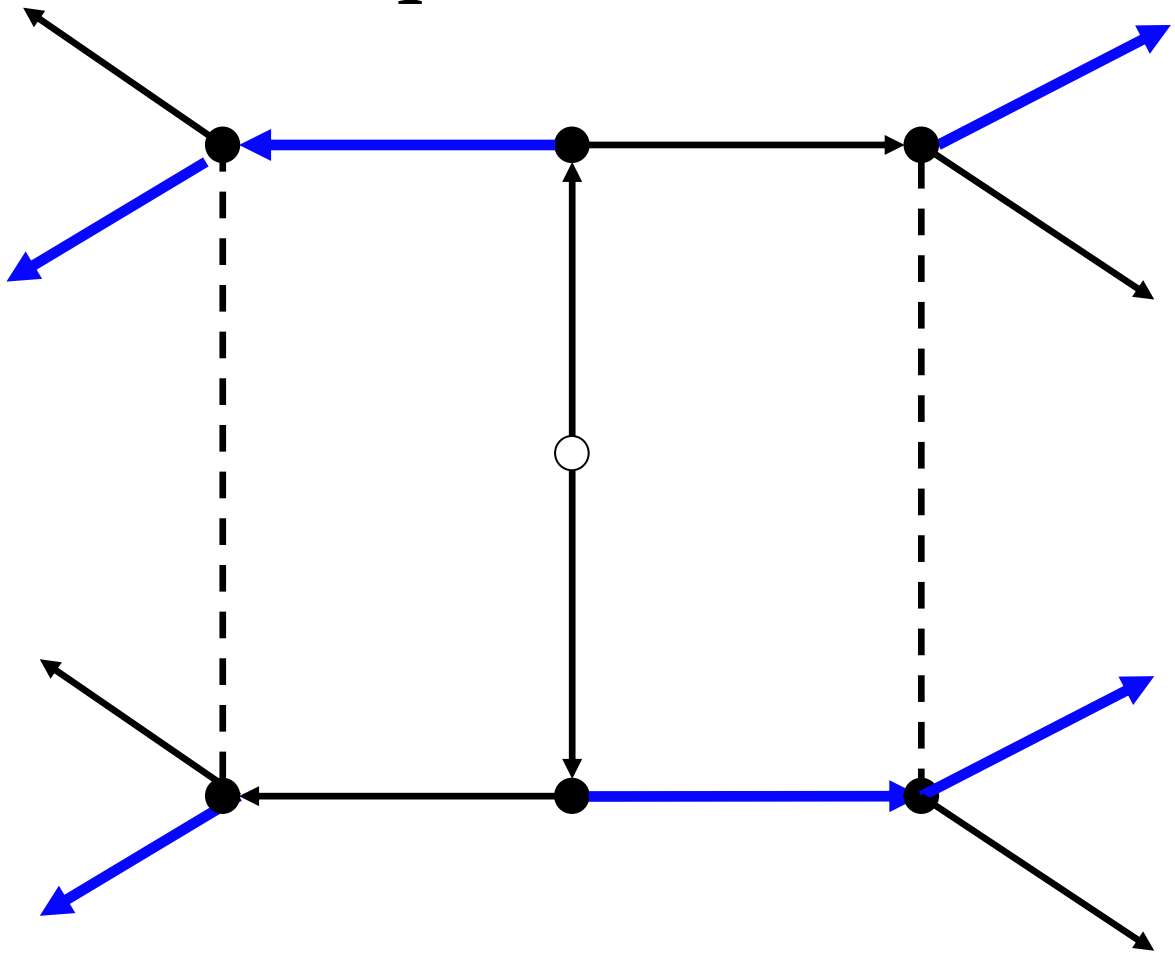
A possible game tree



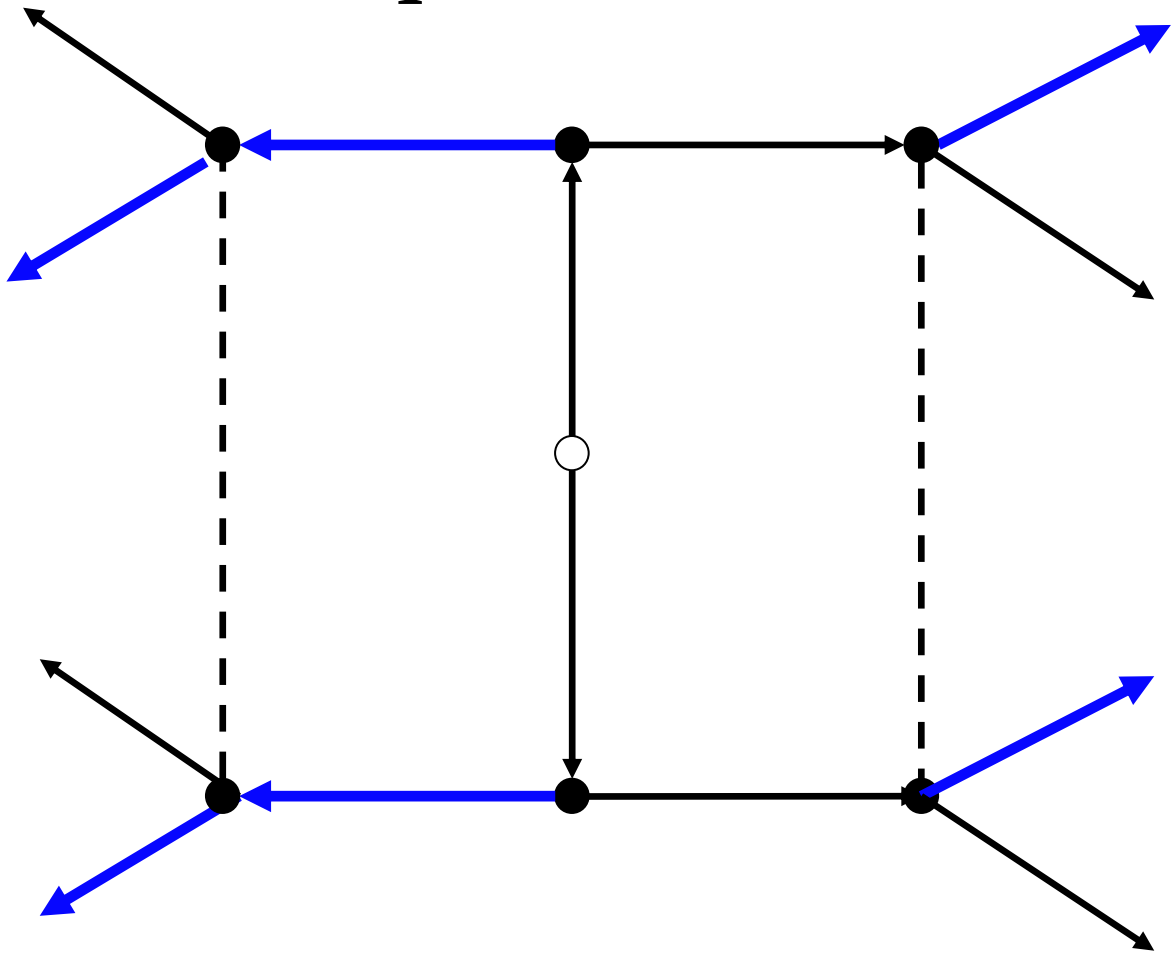
Types of equilibria

- An equilibrium where all types of informed players do the same thing is called a **pooling equilibrium**
- An equilibrium where all types of informed players do different thing is called a **separating equilibrium**
- Also **partially pooling or semiseparating equilibria** are possible

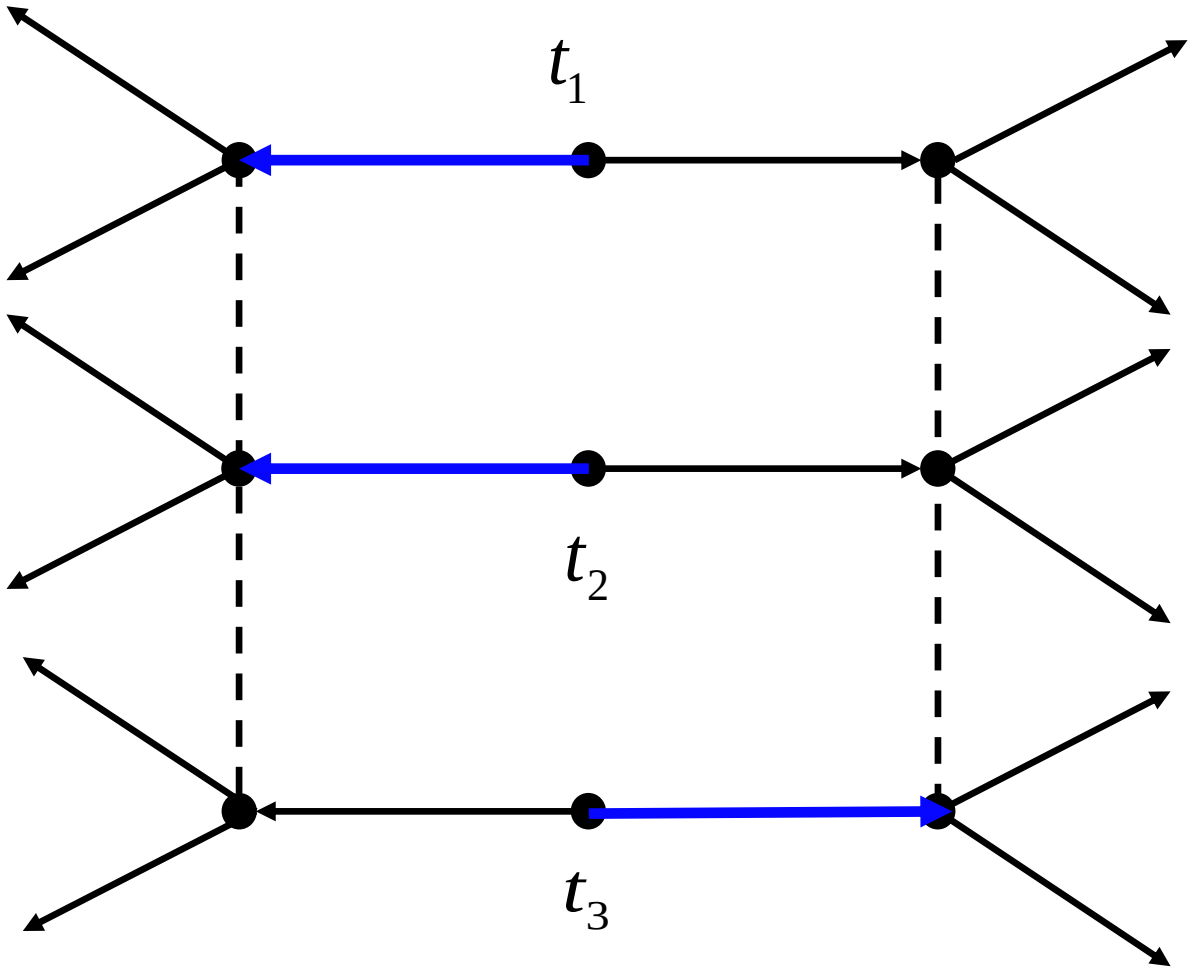
Example of possible separating equilibrium



Example of possible pooling equilibrium



Example of possible semiseparating equilibrium

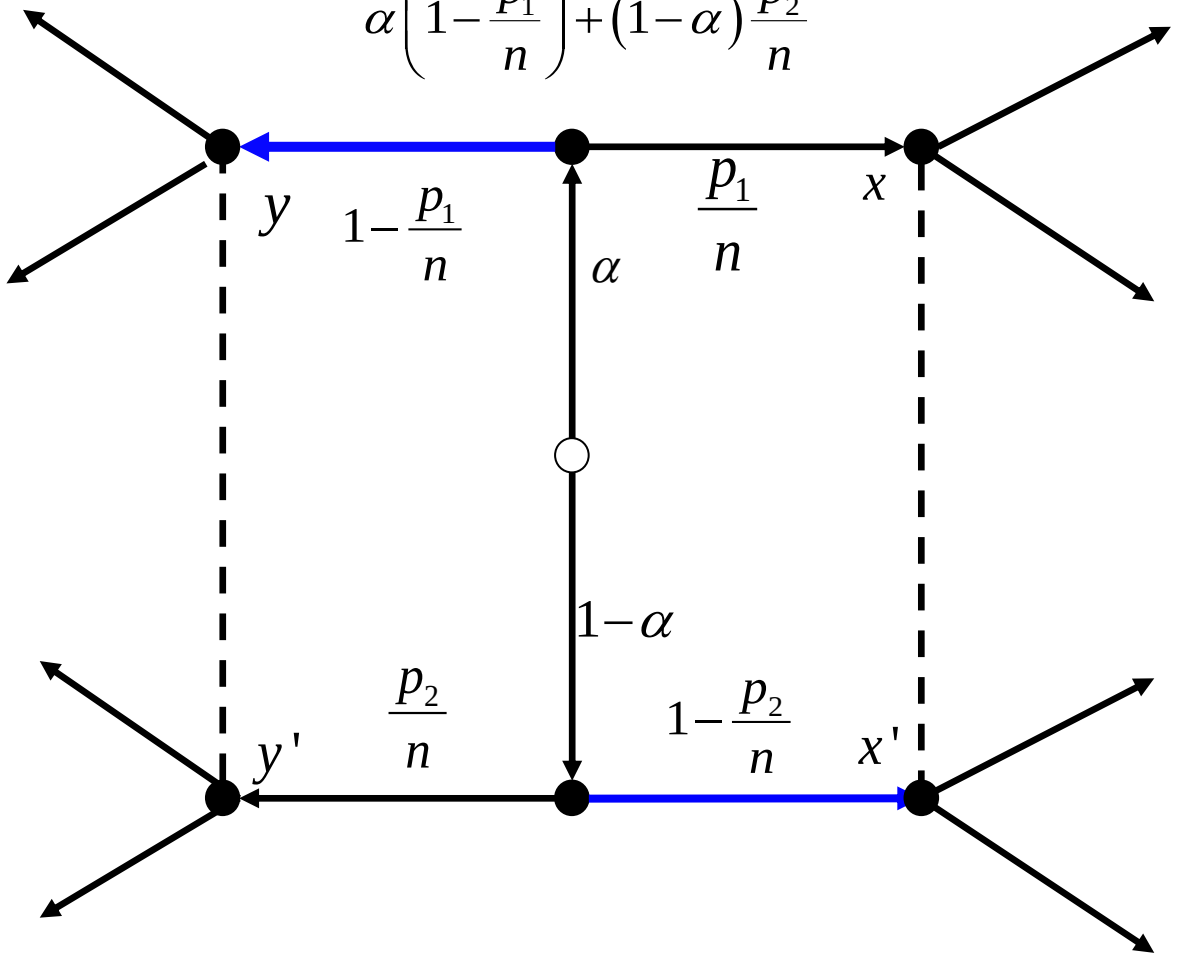


Weak Perfect Bayesian Equilibria and Sequential Equilibria

- In Signalling Games WPBE and SE coincide:
  - If we have (semi)**separating equilibria**, then there are no out-of-equilibrium information sets and thus no problems
  - If we have **pooling or partialling pooling equilibria**, then see next picture

Weak Perfect Bayesian Equilibria and Sequential Equilibria: **Separating Equilibria**

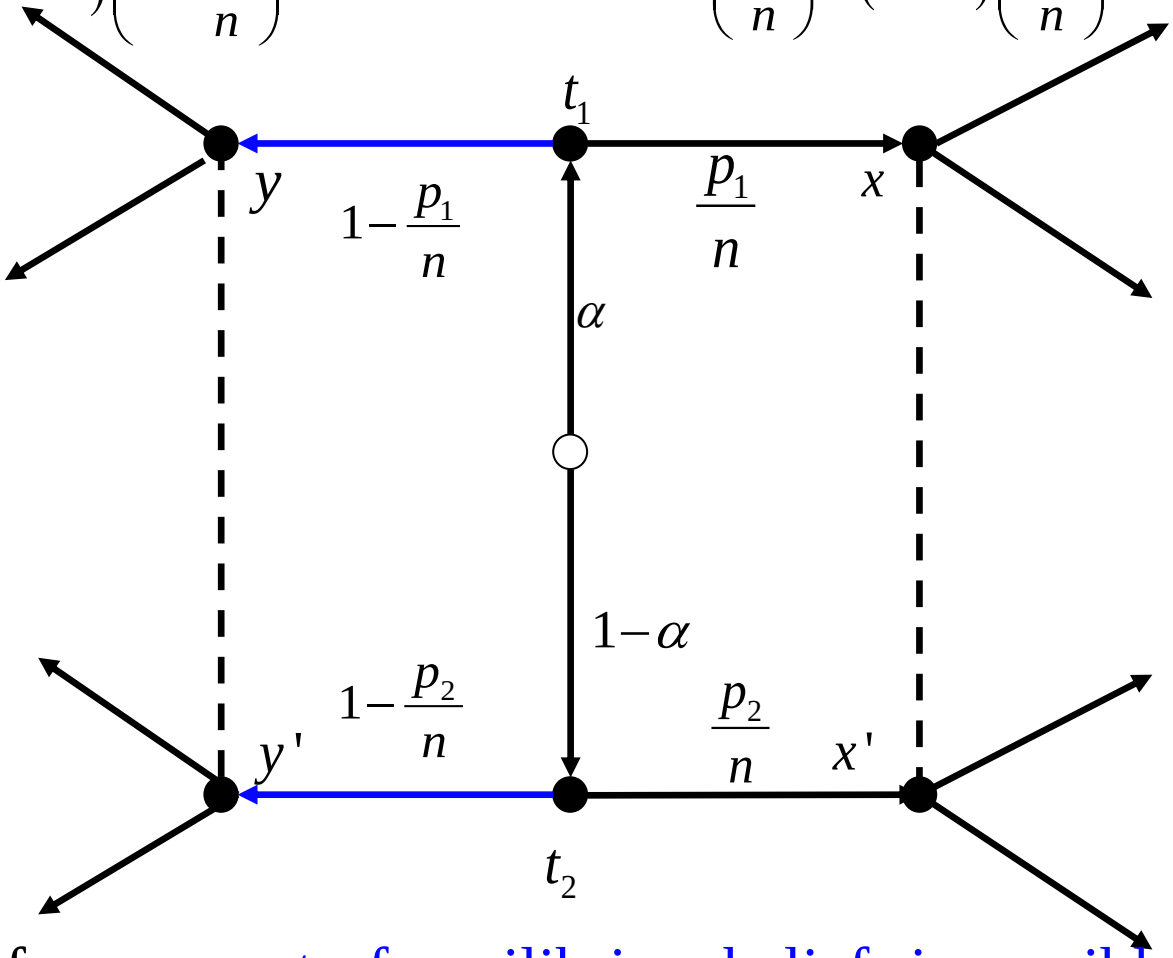
$$\mu(y) = \frac{\alpha\left(1-\frac{p_1}{n}\right)}{\alpha\left(1-\frac{p_1}{n}\right)+(1-\alpha)\frac{p_2}{n}} \xrightarrow[n\to\infty]{} 1$$



There are no out-of-equilibrium beliefs,
hence there is no problem with separating equilibria

Weak Perfect Bayesian Equilibria and Sequential Equilibria: **Pooling Equilibria**

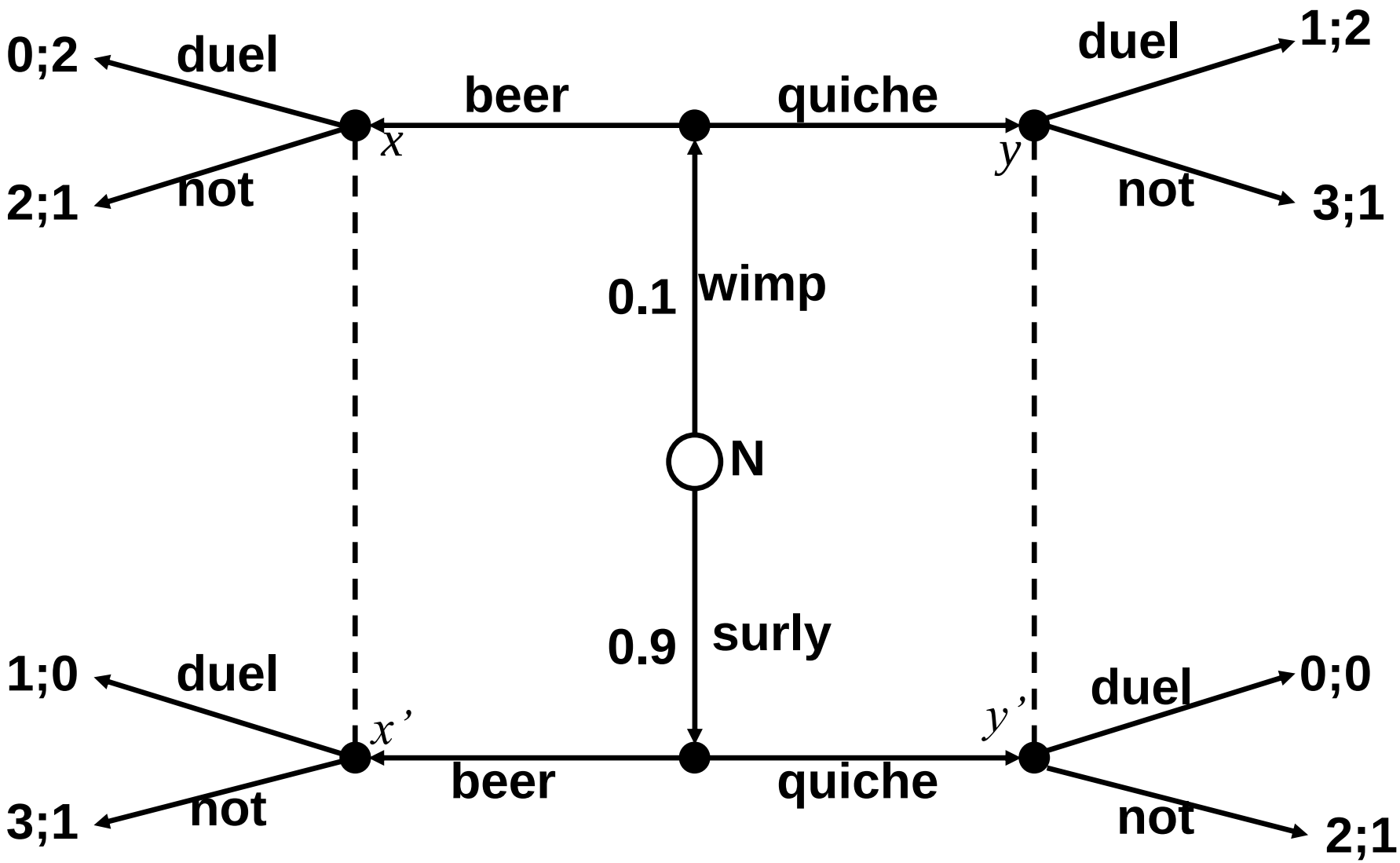
$$\mu(y)=\frac{\alpha\left(1-\frac{p_1}{n}\right)}{\alpha\left(1-\frac{p_1}{n}\right)+(1-\alpha)\left(1-\frac{p_2}{n}\right)}\xrightarrow{n\to\infty}\alpha \qquad \mu(x)=\frac{\alpha\left(\frac{p_1}{n}\right)}{\alpha\left(\frac{p_1}{n}\right)+(1-\alpha)\left(\frac{p_2}{n}\right)}=\frac{\alpha p_1}{\alpha p_1+(1-\alpha)p_2}\in[0,1]$$



Threfore any out-of-equilibrium beliefs is possible both with WPBE and with SE. ***The value of μ will depend on p¹ versus p², i.e. whether we believe is more likely that t¹ or t² has deviated from SE***

An example

Beer and Quiche: The Entry-Deterrence Problem



ANALYSIS OF THE STRATEGIC FORM TO FIND BAYES-NASH EQUILIBRIA AND THEN THE POSSIBLE WPBE

Beer and Quiche: the strategic form game

	DD	DN	ND	NN
BB	0.9;0.2	0.9;0.2	2.9;1	2.9;1
BQ	0;0.2	1.8;1.1	0.2;0.1	2;1
QB	1;0.2	1.2;0.1	2.8;1.1	3;1
QQ	0.1;0.2	2.1;1	0.1;0.2	2.1;1

Beer and Quiche: the Bayes-Nash equilibria

	DD	DN	ND	NN
BB	0.9;0.2	0.9;0.2	2.9;1	2.9;1
BQ	0;0.2	1.8;1.1	0.2;0.1	2;1
QB	1;0.2	1.2;0.1	2.8;1.1	3;1
QQ	0.1;0.2	2.1;1	0.1;0.2	2.1;1

Beer and Quiche: the Bayes-Nash equilibria and the possible WPBE

Two kinds of possible pooling WPBE:

1. (BB; ND): both types drink beer, and the entrant duels if quiche is observed but declines to duel if beer is observed. To find a WPBE we should derive the possible beliefs that makes such decisions sequentially rational
2. (QQ; DN): both types have quiche, the entrant duels if beer is observed but declines to duel if quiche is observed. To find a WPBE we should derive the possible beliefs that makes such decisions sequentially rational.

Beer and Quiche: the Bayes-Nash equilibria and the possible WPBE

The first possible pooling WPBE:

1. (BB; ND):

$$\mu(x\,|\,\{x,x'\}) = \mu(W\,|\,B) = \frac{\mu(W)\pi(B\,|\,W)}{\mu(W)\pi(B\,|\,W)+\mu(S)\pi(B\,|\,S)} = \frac{0.1\times 1}{0.1\times 1+0.9\times 1} = 0.1$$

$$\mu(y\,|\,\{y,y'\}) = \mu(W\,|\,Q) = \frac{\mu(W)\pi(Q\,|\,W)}{\mu(W)\pi(Q\,|\,W)+\mu(S)\pi(Q\,|\,S)} = \frac{0.1\times 0}{0.1\times 0+0.9\times 0} = \frac{0}{0} =: \alpha \in [0,1]$$

Hence ND should satisfy

$$Eu_2\big(N(\{x,x'\}\,|\,\mu) \geq Eu_2\big(D(\{x,x'\}\,|\,\mu) \Leftrightarrow 1\times 0.1+1\times 0.9 \geq 2\times 0.1+0\times 0.9 \text{ always satisfied}$$

$$Eu_2\big(D(\{y,y'\}\,|\,\mu) \geq Eu_2\big(N(\{y,y'\}\,|\,\mu) \Leftrightarrow 2\times\alpha+0\times(1-\alpha) \geq 1\times\alpha+1\times(1-\alpha)$$

$$\text{always satisfied for } \alpha = \geq 0.5$$

Then the WPBE is (BB; ND), $\mu(x|\{x,x'\}) = 0.1$, $\mu(y|\{y,y'\}) \geq 0.5$.

Beer and Quiche: the Bayes-Nash equilibria and the possible WPBE

The second possible pooling WPBE:

2. (QQ; DN):

$$\mu(x \mid \{x, x'\}) = \mu(W \mid B) = \frac{\mu(W)\pi(B \mid W)}{\mu(W)\pi(B \mid W) + \mu(S)\pi(B \mid S)} = \frac{0.1 \times 0}{0.1 \times 0 + 0.9 \times 0} = \frac{0}{0} =: \alpha \in [0,1]$$

$$\mu(y \mid \{y, y'\}) = \mu(W \mid Q) = \frac{\mu(W)\pi(Q \mid W)}{\mu(W)\pi(Q \mid W) + \mu(S)\pi(Q \mid S)} = \frac{0.1 \times 1}{0.1 \times 1 + 0.9 \times 1} = 0.1$$

Hence DN should satisfy

$$Eu_2\left(D(\{x, x'\} \mid \mu\right) \geq Eu_2\left(N(\{x, x'\} \mid \mu\right) \Leftrightarrow 2 \times \alpha + 0 \times (1 - \alpha) \geq 1 \times \alpha + 1 \times (1 - \alpha)$$

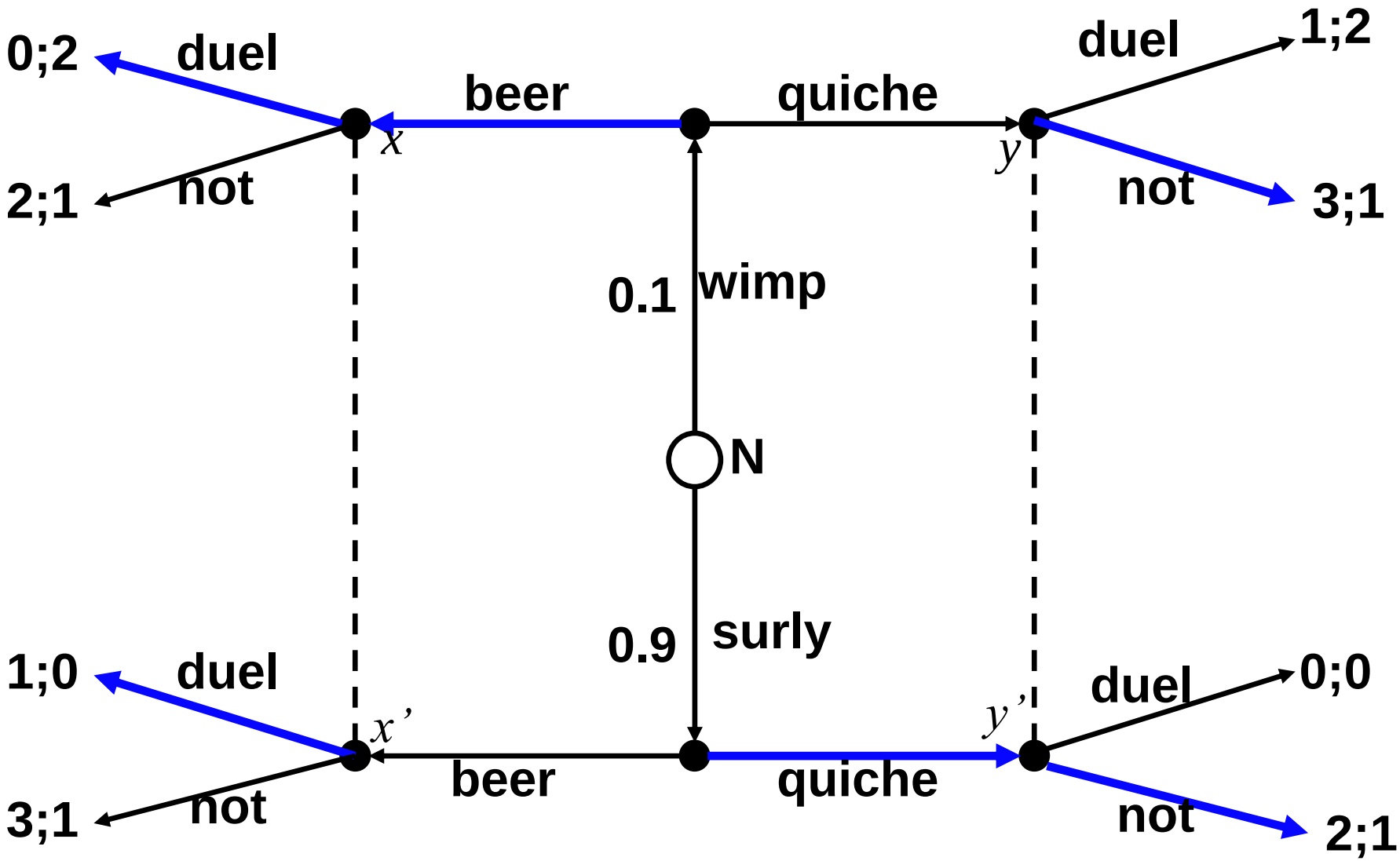
$$\text{always satisfied for } \alpha \geq 0.5$$

$$Eu_2\left(N(\{y, y'\} \mid \mu\right) \geq Eu_2\left(D(\{y, y'\} \mid \mu\right) \Leftrightarrow 1 \times 0.1 + 1 \times 0.9 \geq 2 \times 0.1 + 0 \times 0.9 \text{ always satisfied}$$

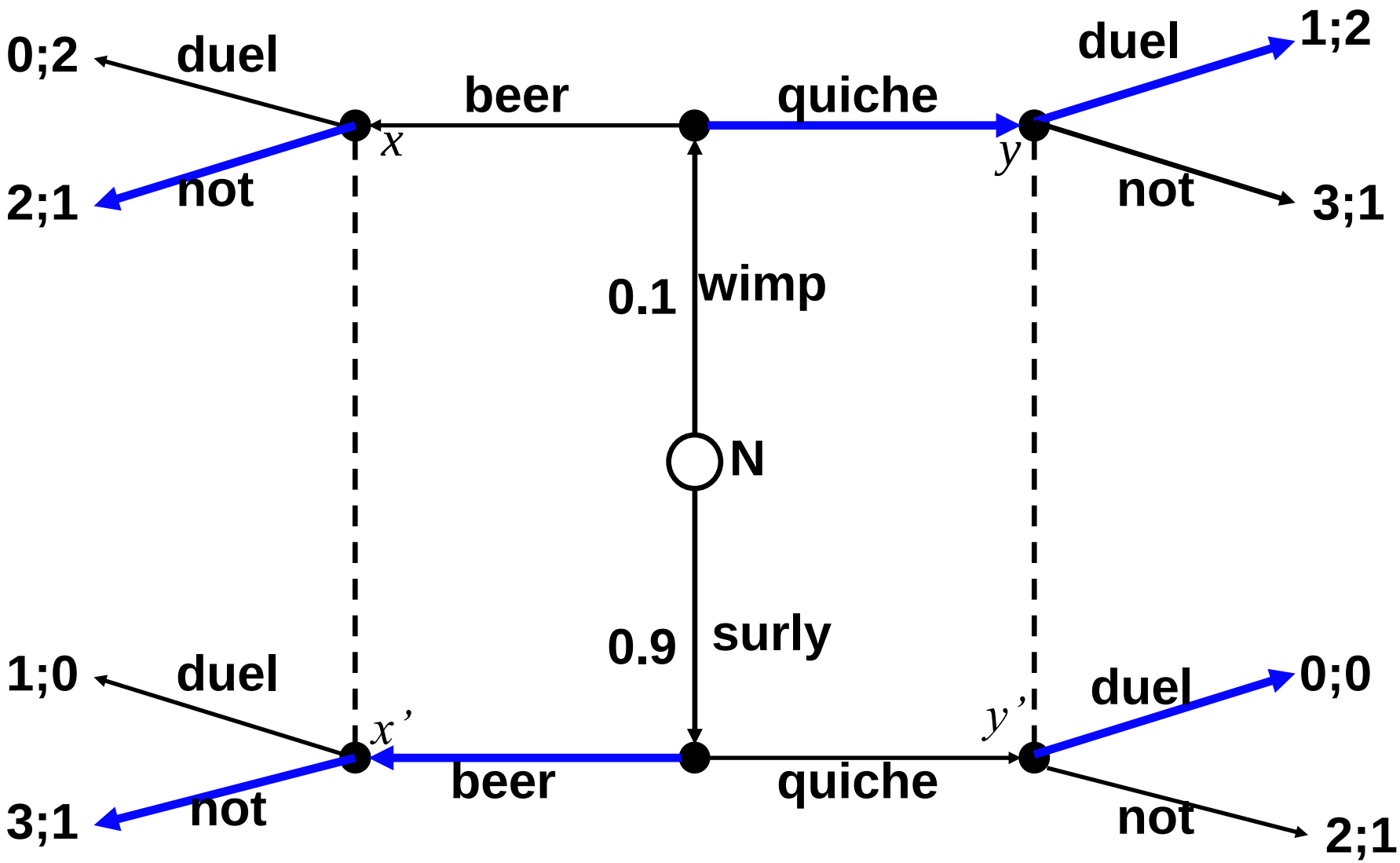
Then the WPBE is (QQ; DN), $\mu(x|\{x,x'\}) \geq 0.5$, $\mu(y|\{y,y'\}) = 0.1$.

ANALYSIS OF THE EXTENSIVE FORM TO FIND DIRECTLY THE POSSIBLE WPBE

First possible equilibrium:
beer-quiche, then μ(x)=1& μ(y')=1, which implies duel-not. In turn this implies that wimp deviates to quiche: it is not a WPBE

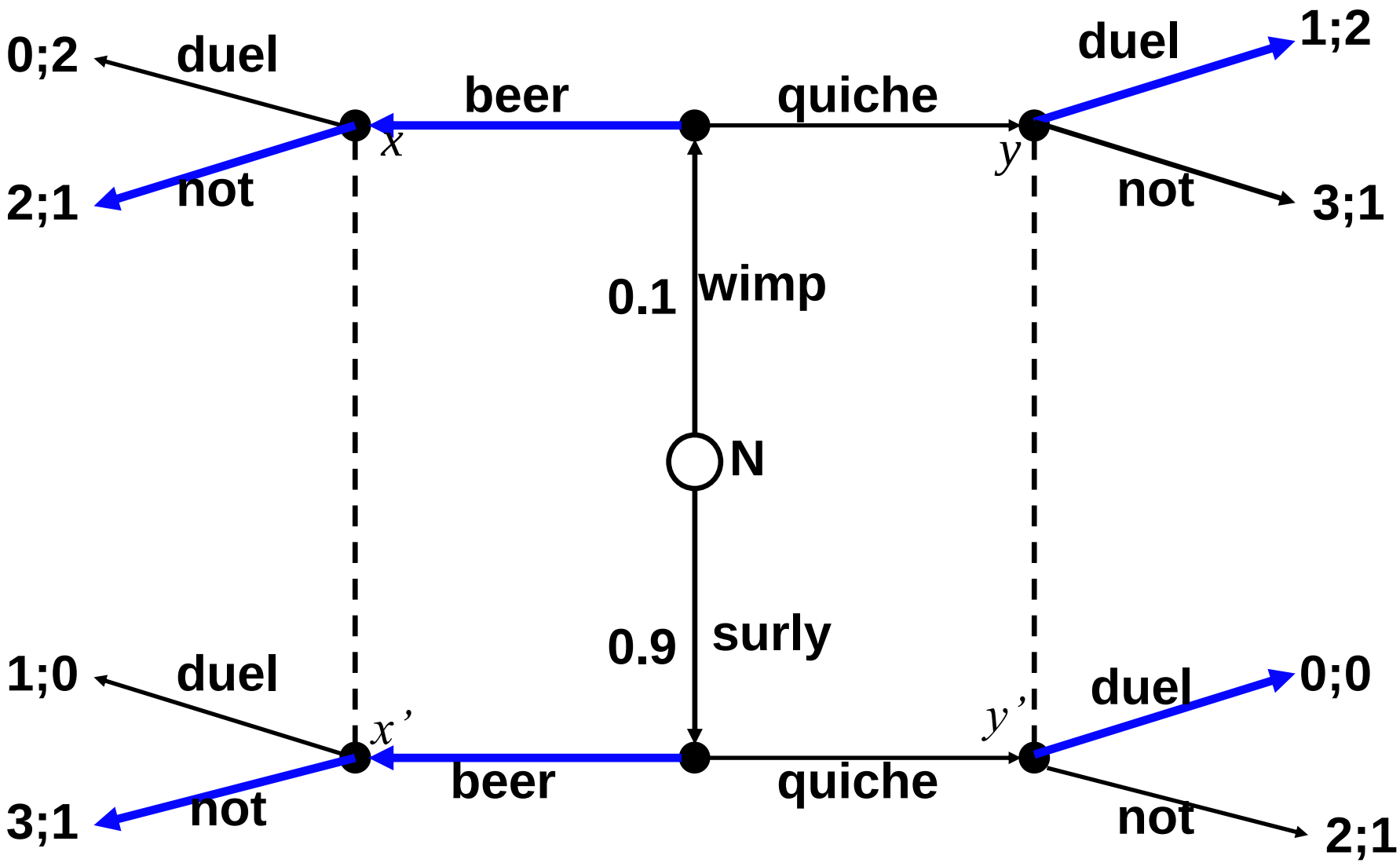


Second possible equilibrium:
quiche-beer, then μ(x')=1& μ(y)=1, which implies not-duel. In turn this implies that wimp deviates to beer: it is not a WPBE

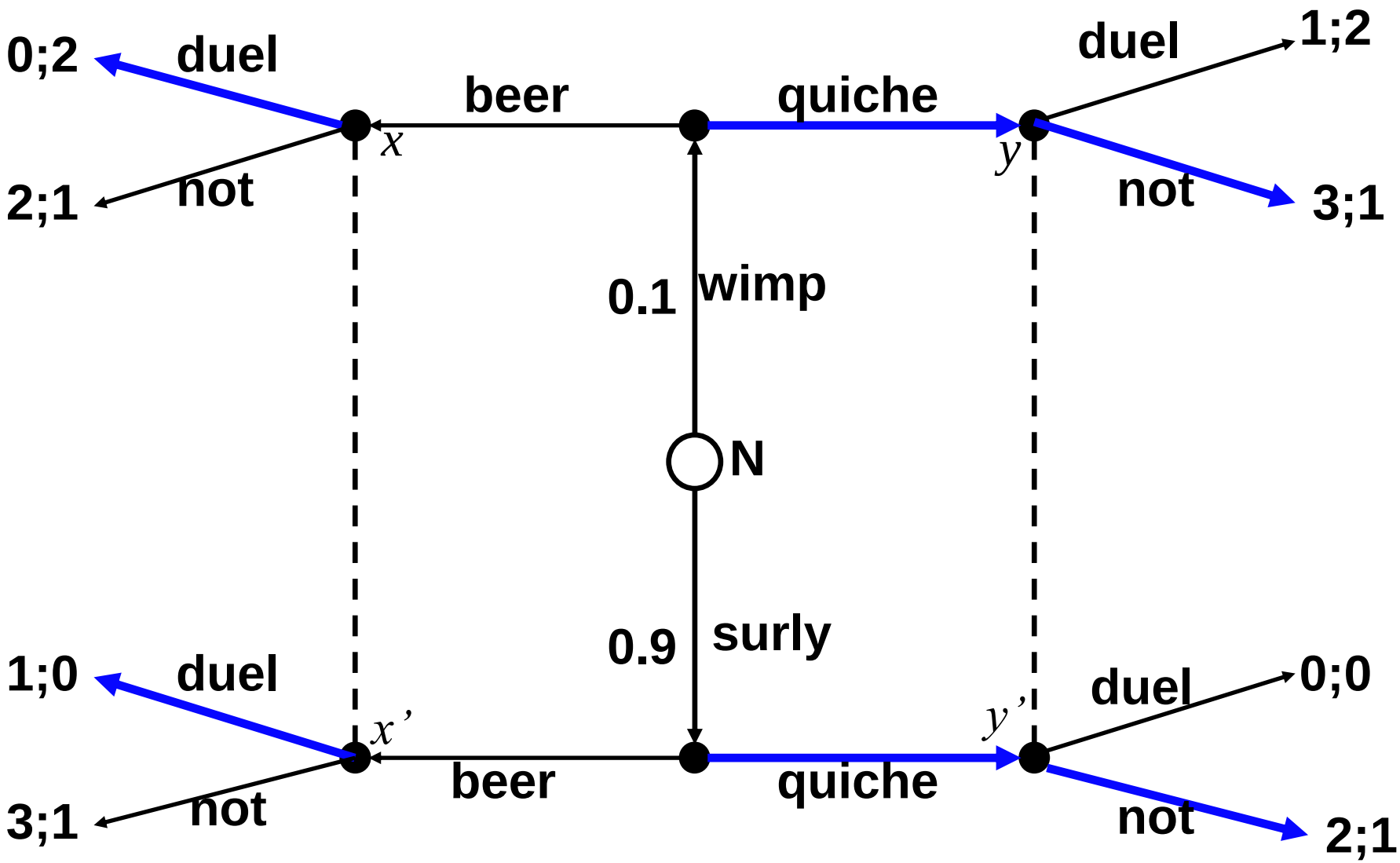


Third possible equilibrium:

beer-beer, then $\mu(x)=0.1$ & $\mu(y)\in[0,1]$, which implies not. In turn this implies that 1will not deviate iff 2 duel in $\{y,y'\}$, i.e. $\mu(y) \geq 1/2$



Fourth possible equilibrium:
quiche-quiche, then μ(y)=0.1& μ(x)∈[0,1], which implies not. In turn this implies that 1will not deviate iff 2 duel in *{x,x'}*, i.e. μ(x)≥1/2



The sequential equilibria of Beer and Quiche in pure strategies

- No separating equilibria:

1. q(w), b(s) imply d(q),n(b) and given this action of R, S would deviate

2. b(w), q(s) is not part of a SE, because R would play d(b),n(q), but then type w would deviate.

Interpretation of first Sequential Equilibrium

- Both types drink beer, and
- the entrant duels if quiche is observed but declines to duel if beer is observed.

In such an equilibrium, **the decision to duel following quiche is rationalized by any off-the-equilibrium-path belief that puts sufficiently high probability (at least 1/2) on the incumbent being wimpy.**

Note: **If Pr(wimp) > 1/2, then the only equilibrium is {beer, beer; duel, duel}**.

Interpretation of second Sequential Equilibrium

- Both types have quiche, and the entrant duels if beer is observed but declines to duel if quiche is observed.
- The beliefs that support the decision to duel are those that attach **high probability to the wimp type after deviation to beer.**
- But here **such beliefs seem unnatural**:
- **the prior belief is .9 that the incumbent is surly, but when conditioned on the observation of beer - which is preferred if surly but not if wimpy - the posterior belief is at least .5 that the incumbent is wimpy**.

The Intuitive Criterion and the Forward Induction Equilibrium in Signaling Games

Sequential Equilibria

How can we reject the second equilibrium?

- Using the **intuitive criterion** one can argue that surly will find it optimal to deviate from the proposed equilibrium (both eat quiche):

  If the entrant concludes that the beer-drinker is surly, then declining to duel is the optimal decision. This yields a payoff of 3 for surly, which is better than the 2 earned in equilibrium.

General approach to identify unreasonable Sequential Equilibria - 1

*Test for the **INTUITIVE CRITERION - 1***:

- Consider the game: T={t,t'} and M={m,m'}.
- Suppose a **pooling equilibrium**: t, t' send m with probability one.
- Then the message **m' is off the equilibrium** path, so R's beliefs after observing m' cannot be derived from Bayes' rule.
- By sequential rationality, the action R takes after observing m' must be optimal given R's beliefs. That is

$$a(m') \in \underset{a \in A}{\operatorname{argmax}} \sum_{t \in T} \mu(t|m') U^R(t, m', a).$$

General approach to identify unreasonable Sequential Equilibria - 2

*Test for the **INTUITIVE CRITERION - 2***:

Suppose:

- (1) ∀ beliefs, the action a(m') makes type **t** worse off than **t** is in the equilibrium, and
- (2) if R infers from m' that S is type **t'**, then R's optimal action will make **t'** better off than **t'** is in the equilibrium.

- Then, if S is type **t'**, the following implicit speech should be believed by R:

I am **t'**.

To prove this, I am sending m' instead of the equilibrium m.

Note that if I were type t, I would not want to do this, no matter what you might infer from m'.

And, as **t'**, *I have an incentive to do this*

***provided it convinces you that I am not* t.**

General approach to identify unreasonable Sequential Equilibria - 3

- Given (1) and (2), **t'** should deviate from the Sequential Equilibrium in which m is sent with probability one.
- On this ground, the INTUITIVE CRITERION (Cho & Kreps) reject the equilibrium.

Formalizing the "Intuitive Criterion“ for general signaling games

Notation - 1:

- After hearing $m \in M$, R's beliefs are $\mu(t|m)$.
- Sequential rationality requires that R's subsequent action a(m) maximize the expectation of $U^R(t,m,a)$ with respect to these beliefs.
- Define **the set of such best responses of the Responder** as

$$BR(\mu, m) \equiv \underset{a \in A}{\operatorname{argmax}} \sum_{t \in T} \mu(t|m) U^R(t, m, a).$$

- Then R's (behavior) strategy $\pi^R(a|m)$ is greater than zero only if $a \in BR(\mu, m)$.

Formalizing the "Intuitive Criterion"

Notation - 2:

- For subsets I of T, let BR(I,m) denote **the set of best responses for R to beliefs concentrated on I**:

$$BR(I,m) \equiv \bigcup_{\{\mu:\mu(I)=1\}} BR(\mu,m).$$

- Given the equilibrium strategies

$$\pi = \{\pi^S(m|t),\ \pi^R(a|m)\},$$

the **equilibrium payoff to a Sender of type t** is

$$U^*(t) \equiv \sum_{a\in A}\sum_{m\in M} \pi^R(a|m)\pi^S(m|t)U^S(t,m,a).$$

Intuitive Criterion

- **An equilibrium**

$$\{\pi^S(m|t),\ \pi^R(a|m)\}$$

***fails* to satisfy the intuitive criterion** if there exist

- an **unsent message** $m' \in M$ (i.e., $\pi^S(m'|t) = 0$ for all $t \in T$) and a
- a subset J of T such that

(1) for all $t \in J$, for all $a \in BR(T, m')$,

$$\mathbf{U^*(t) > U^S(t,m',a)}$$

(2) there exists a type $t' \in T \setminus J$ such that

for all $a \in BR(T \setminus J, m')$,

$$\mathbf{U^*(t') < U^S(t',m',a)}.$$

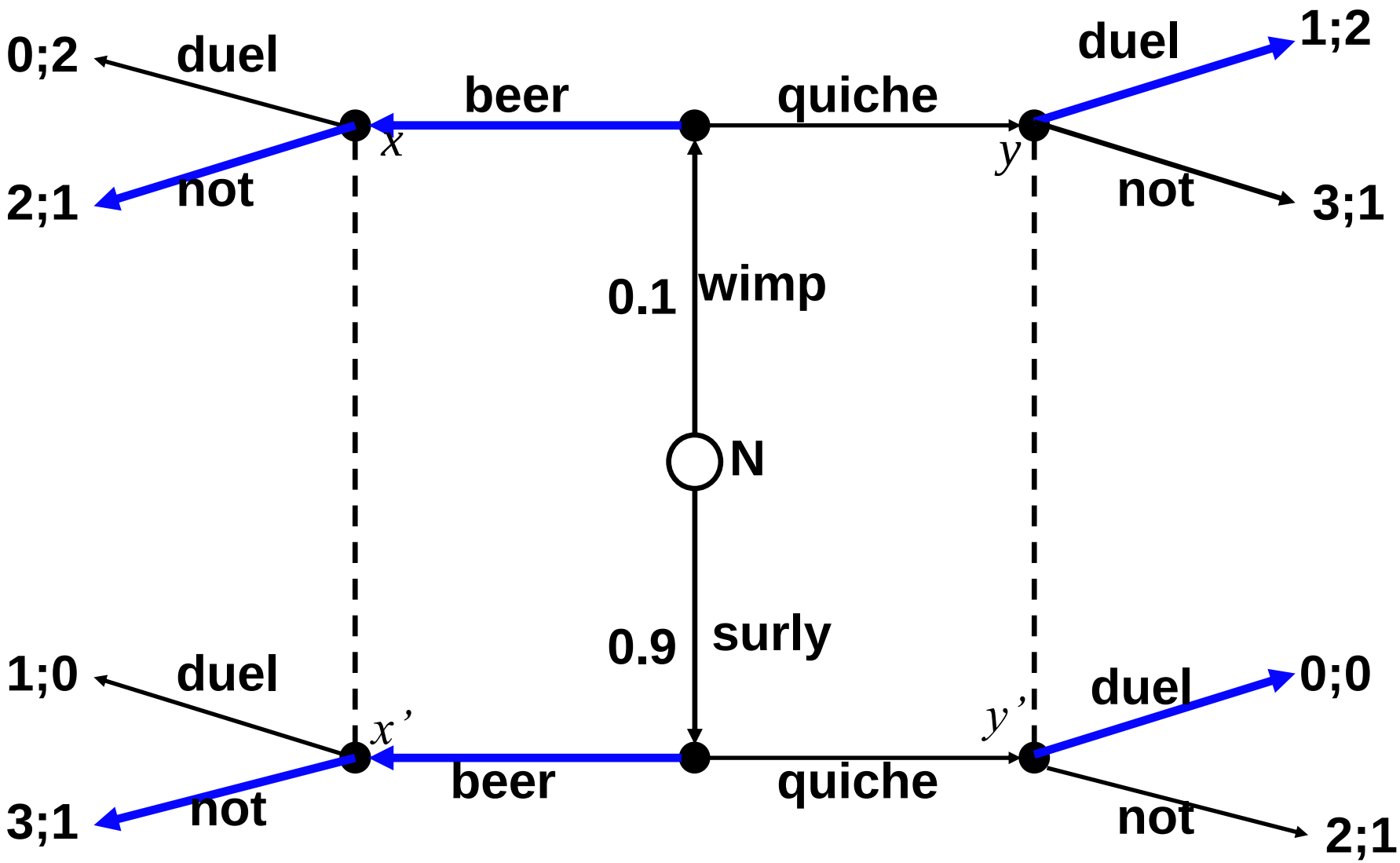
Intuitive meaning of the intuitive criterion

- *Step 1*: Condition (1) suggests that R's out of equilibrium beliefs $\mu(t|m')$ should put no probability on types $t \in J$: **reasonable μ (.|m') should be concentrated on T \ J.**
- *Step 2*: Given this new plausible beliefs, if there is a type t' satisfying (2), then this type t' has an incentive to deviate from the proposed equilibrium, since t' is better off deviating no matter what reasonable belief R will hold, **if these beliefs are concentrated on T \ J**.

ANALYSIS OF THE SEQUENTIAL EQUILIBRIA OF THE BEER-QUICHE GAME USING THE INTUITIVE CRITERION

First pooling equilibrium:

beer-beer, then μ(x)=0.1& μ(y)∈[0,1], which implies not. In turn this implies that 1will not deviate iff 2 duel in *{y,y'}*, i.e. μ(y) ≥ 1/2



First sequential equilibrium - 1

- BB, N({x,x'}) D({y,y'}),
- $\mu(x|\{x,x'\})=0.1$ & $\mu(y|\{y,y'\})\geq 0.5$

- The equilibrium payoffs for both types of the sender are

$U^*(W)=2$ and $U^*(S)=3$

- Now consider the unsent message Q

and

(a) a possible subset J of T, e.g. J={W} so that

BR({W,S},Q)={D,N}; BR({W},Q)={D}; BR({S},Q)={N};

(b) a type $t' \in T \sim J=\{S\}$

check whether the two conditions to reject a sequential equilibrium as failing the intuitive criterion are satisfied

First sequential equilibrium - 2

- **Intuitive criterion**: reject any sequential equilibrium satisfying the following conditions:

  there exists an unsent message m' and a subset of types J such that

  (1) for all $t \in J$, for all $a \in BR(T,m')$, $U^*(t) > U^S(t,m',a)$, and

  (2) there exists $t' \in T \sim J$ such that for all $a \in BR(T \sim J,m')$, $U^*(t') < U^S(t',m',a)$,

  where $U^*(t)$ is t's expected payoff in the equilibrium under consideration.

First sequential equilibrium - 3

- unsent message Q and J={W} such that

$$(1)\ U^{*}(W) = 2 < U^{S}(W,Q,N) = 3 \text{ where}$$
$$N \in BR(\{W,S\},Q)=\{D,N\}$$

Therefore we can not reject this equilibrium using the intuitive criterion if we take J={W}.

First sequential equilibrium - 4

- Now consider the other possible subset J of T, J={S} so that

BR({W,S},Q)={D,N}; BR({W},Q)={D};
BR({S},Q)={N};

(b) a type t' $\in$ T~J={W}

check whether the two conditions to reject a sequential equilibrium as failing the intuitive criterion are satisfied

First sequential equilibrium - 5

- unsent message Q and J={S} such that

$$(1)\ U^{*}(S) = 3 > U^{S}(S,Q,a)$$

for any $a \in BR(\{W,S\},Q)=\{D,N\}$

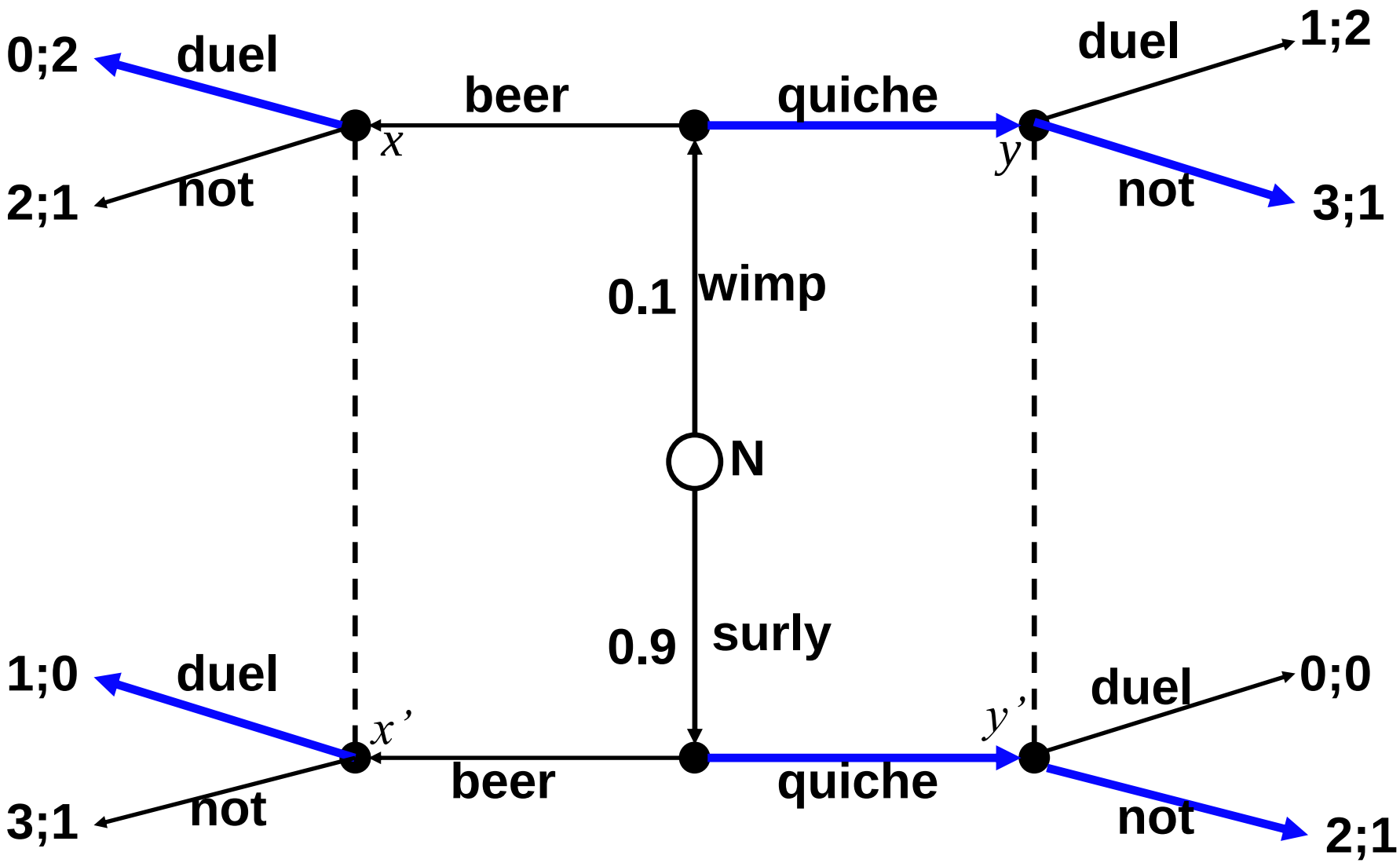
$$(2)\ U^{*}(W) = 2 > U^{S}(W,Q,D) = 1$$ where

$D \in BR(\{W\},Q)=\{D\}$

Therefore we can not reject this equilibrium using the intuitive criterion if we take J={S}.

THEREFORE WE SHOULD ACCEPT THIS EQUILIBRIUM AS CONSISTENT WITH THE INTUITIVE CRITERION

Second pooling equilibrium:
quiche-quiche, then μ(y)=0.1& μ(x)∈[0,1], which implies not. In turn this implies that 1will not deviate iff 2 duel in *{x,x'}*, i.e. μ(x)≥1/2



Second sequential equilibrium - 1

- QQ, D({x,x'})N({y,y'}), $\mu(y|\{y,y'\})=0.1$ & $\mu(x|\{x,x'\})\geq 0.5$
- The equilibrium payoffs for both types of the sender are

  $U^*(W)=3$ and $U^*(S)=2$
- Now consider the unsent message B and

  (a) a possible subset J of T, J={W} so that

BR({W,S},B)={D,N}; BR({W},B)={D}; BR({S},B)={N};

(b) a type $t' \in T \sim J=\{S\}$

and check whether the two conditions to reject a sequential equilibrium as failing the intuitive criterion are satisfied

Second sequential equilibrium - 2

- **Intuitive criterion**: reject any sequential equilibrium satisfying the following conditions

  there exists an unsent message m' and a subset of types J such that

  (1) for all $t \in J$, for all $a \in BR(T,m')$, $U^*(t) > U^S(t,m',a)$, and

  (2) there exists $t' \in T \sim J$ such that for all $a \in BR(T \sim J, m')$, $U^*(t') < U^S(t',m',a)$,

  where $U^*(t)$ is t expected payoff in the equilibrium under consideration.

Second sequential equilibrium - 3

- unsent message B and J={W} such that

$$(1)\ U^{*}(W) = 3 > U^{S}(W,B,a)$$

for any $a \in BR(\{W,S\},B)=\{D,N\}$

$$(2)\ U^{*}(S) = 2 < U^{S}(S,B,N) = 3 \text{ where}$$

$$N \in BR(\{S\},B)=\{N\}$$

Therefore we reject this equilibrium using the intuitive criterion if we take J={W}.

THEREFORE THIS EQUILIBRIUM IS INCONSISTENT WITH THE INTUITIVE CRITERION