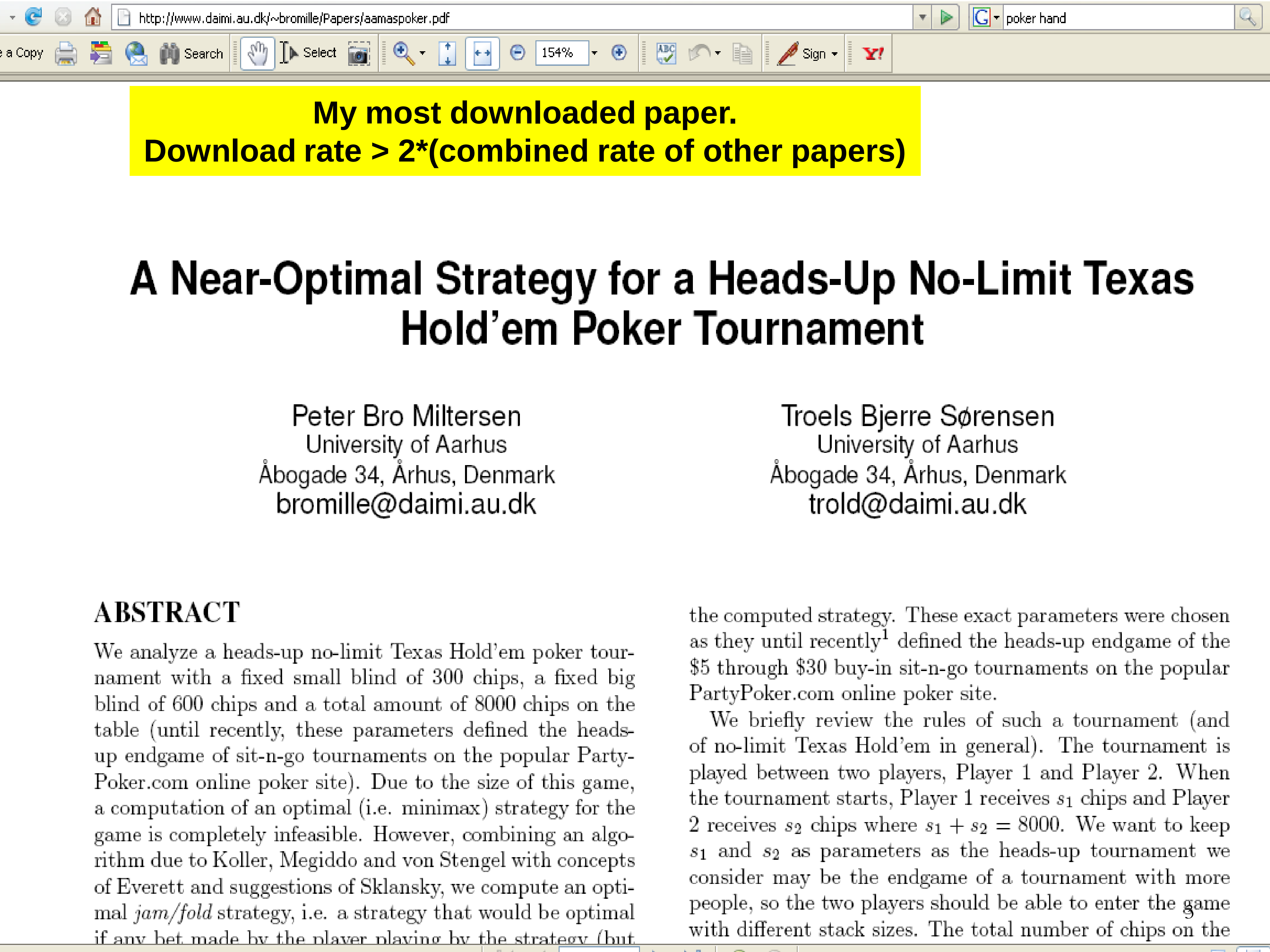
Equilibrium refinements in computational game theory

Peter Bro Miltersen,
Aarhus University

Center for
Algorithmic Game Theory

Computational game theory in AI: The challenge of poker….

My most downloaded paper.
Download rate > 2*(combined rate of other papers)



A Near-Optimal Strategy for a Heads-Up No-Limit Texas Hold'em Poker Tournament

Peter Bro Miltersen
University of Aarhus
Åbogade 34, Århus, Denmark
bromille@daimi.au.dk

Troels Bjerre Sørensen
University of Aarhus
Åbogade 34, Århus, Denmark
trold@daimi.au.dk

ABSTRACT

We analyze a heads-up no-limit Texas Hold'em poker tournament with a fixed small blind of 300 chips, a fixed big blind of 600 chips and a total amount of 8000 chips on the table (until recently, these parameters defined the heads-up endgame of sit-n-go tournaments on the popular PartyPoker.com online poker site). Due to the size of this game, a computation of an optimal (i.e. minimax) strategy for the game is completely infeasible. However, combining an algorithm due to Koller, Megiddo and von Stengel with concepts of Everett and suggestions of Sklansky, we compute an optimal *jam/fold* strategy, i.e. a strategy that would be optimal if any bet made by the player playing by the strategy (but

the computed strategy. These exact parameters were chosen as they until recently[1] defined the heads-up endgame of the $5 through $30 buy-in sit-n-go tournaments on the popular PartyPoker.com online poker site.

We briefly review the rules of such a tournament (and of no-limit Texas Hold'em in general). The tournament is played between two players, Player 1 and Player 2. When the tournament starts, Player 1 receives s_1 chips and Player 2 receives s_2 chips where $s_1 + s_2 = 8000$. We want to keep s_1 and s_2 as parameters as the heads-up tournament we consider may be the endgame of a tournament with more people, so the two players should be able to enter the game with different stack sizes. The total number of chips on the

Game theory in
(most of) Economics**:**

Descriptive

What **is** the outcome when
rational agents interact?

Stability concept**:**

Computational game theory in
(most of) CAV and (some of) AI:

Prescriptive

What **should** we do
to **win**?

Guarantee concept**:**

**For
2-player
0-sum
games**

↓

Nash equilibrium = Maximin/Minimax

Refined stability notions:

Sequential equilibrium

Trembling hand perfection → **Stronger guarantees?**

Quasiperfect equilibrium

Proper equilibrium

**Most of
this
morning**

Computational game theory in CAV vs. Computational game theory in AI

- Main challenge in CAV***: Infinite duration***
- Main challenge in AI: ***Imperfect information***

Plan

- ***Representing*** finite-duration, imperfect information, two-player zero-sum games and ***computing*** minimax strategies.
- Issues with minimax strategies.
- Equilibrium refinements (a crash course) and how refinements resolve the issues, and how to modify the algorithms to compute refinements.
- (If time) Beyond the two-player, zero-sum case.

(Comp.Sci.) References

- D. Koller, N. Megiddo, B. von Stengel. Fast algorithms for finding randomized strategies in game trees. *STOC'94*. doi:10.1145/195058.195451

- P.B. Miltersen and T.B. Sørensen. Computing a quasi-perfect equilibrium of a two-player game. *Economic Theory* **42**. doi:10.1007/s00199-009-0440-6

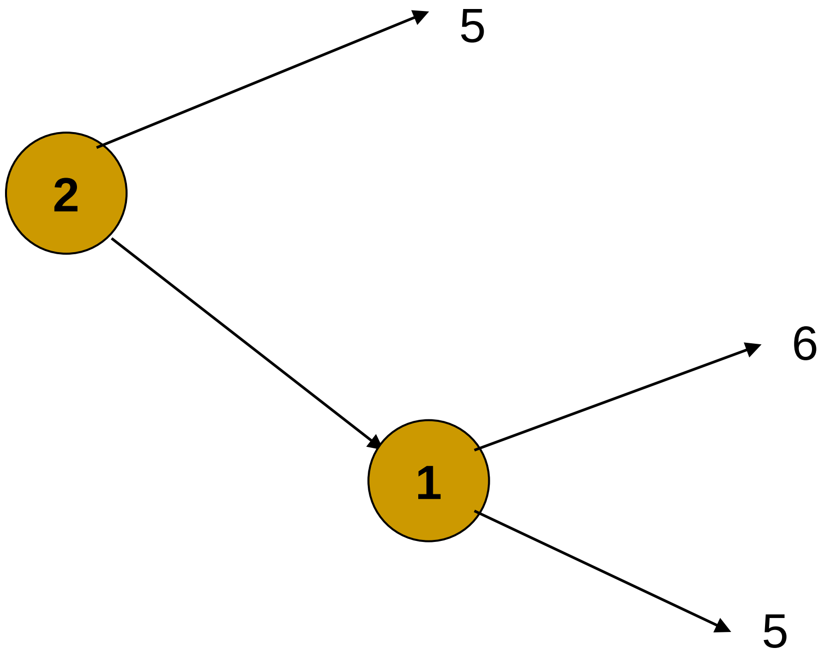
- P.B. Miltersen and T.B. Sørensen. Fast algorithms for finding proper strategies in game trees. *SODA'08*. doi:10.1145/1347082.1347178

- P.B. Miltersen. Trembling hand-perfection is NP-hard. arXiv:0812.0492v1

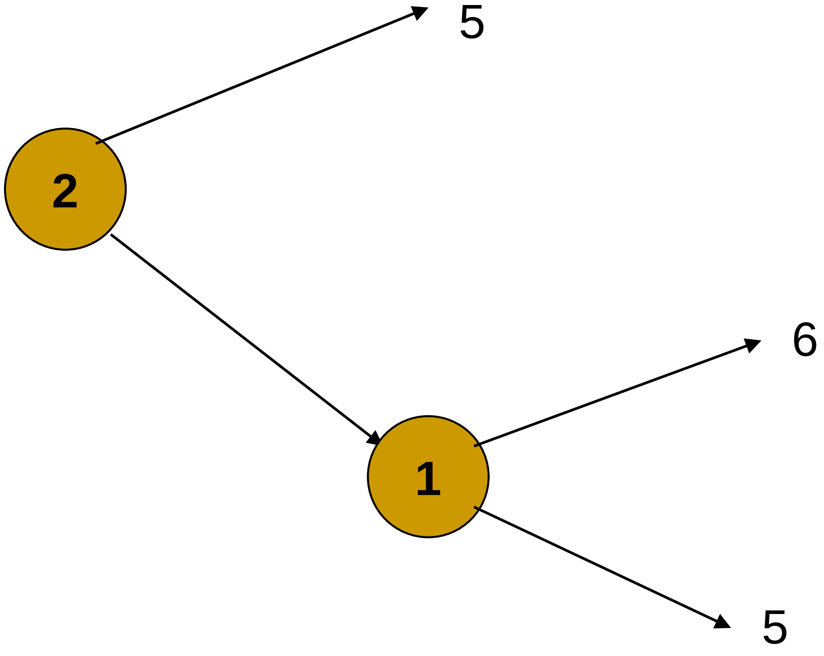
How to make a (2-player) poker bot?

- How to represent and solve two-player, zero-sum games?
- Two well known examples:
    - Perfect information games
    - Matrix games

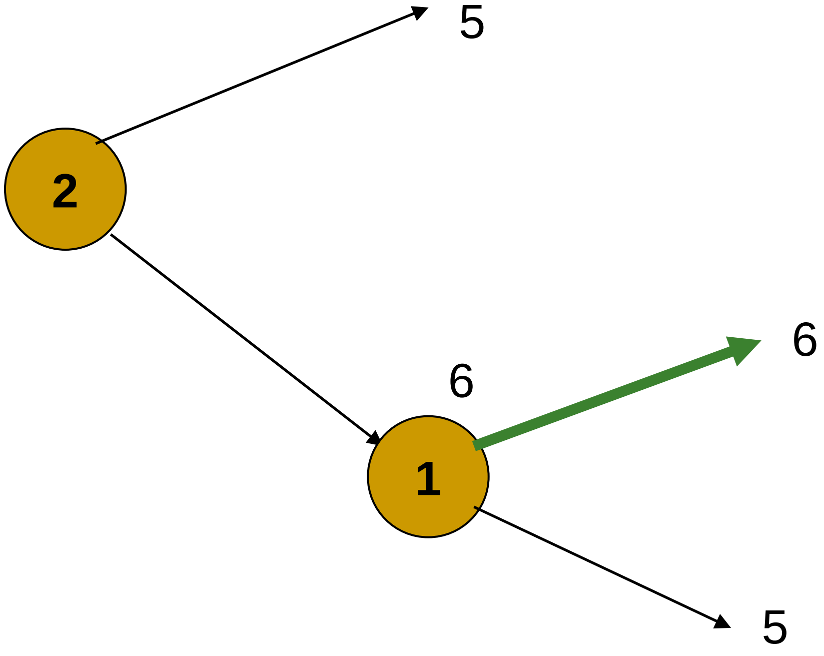
Perfect Information Game (Game tree)



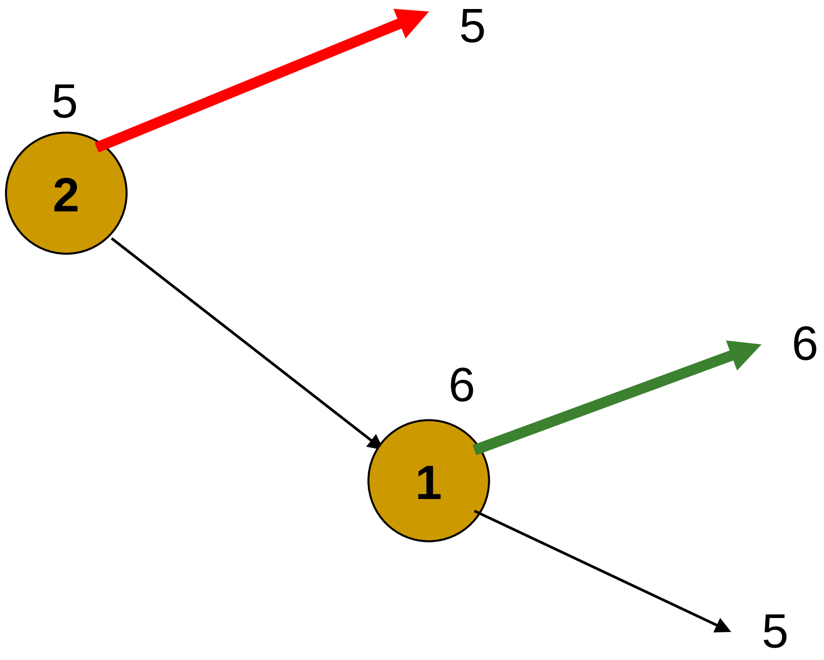
Backwards induction (Minimax evaluation)



Backwards induction (Minimax evaluation)



Backwards induction (minimax evaluation)



The stated strategies are *minimax:* They assure the *best possible* payoff against a *worst case* opponent. Also they are *Nash*: They are best responses against each other.

Matrix games

Matching Pennies:

	Guess head up	**Guess tails up**
Hide heads up	-1	0
Hide tails up	0	-1

Solving matrix games

Matching Pennies:

	Guess head up	**Guess tails up**	
Hide heads up	-1	0	**1/2**
Hide tails up	0	-1	**1/2**
	1/2	**1/2**	

Mixed strategies

The stated strategies are *minimax:* They assure the *best possible* payoff against a *worst case* opponent. Also they are *Nash*: They are best responses against each other.

Solving matrix games

- Minimax mixed strategies for matrix games are found using linear programming.

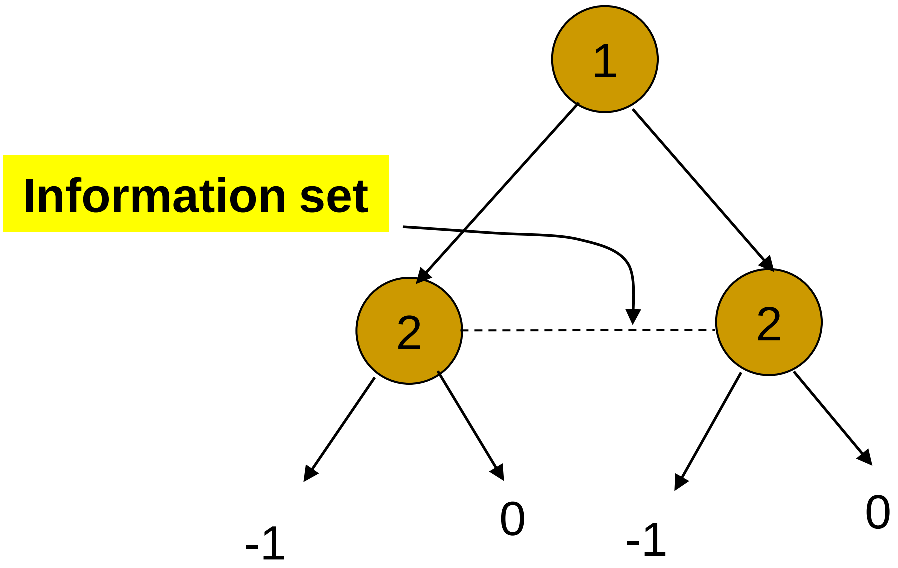
- **Von Neuman’s minmax theorem**: Pairs of minimax strategies are exactly the ***Nash equililbria*** of a matrix games.

How to make a (2-player) poker bot?

- Unlike chess, poker is a game of ***imperfect*** information.
- Unlike matching pennies, poker is an ***extensive*** (or sequential) game.

Can one combine the two very different algorithms (backwards induction and linear programming) to solve such games?

Matching pennies in *extensive form*

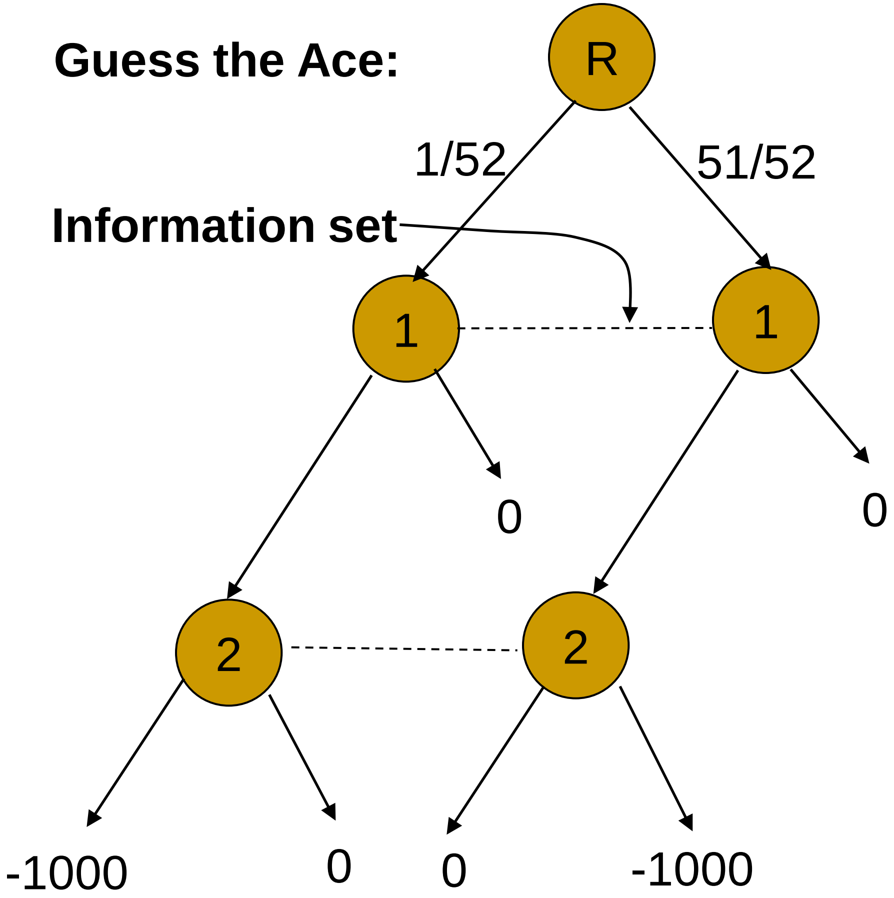


- Player 1 hides a penny either heads up or tails up.
- Player 2 does not know if the penny is heads of or tails up, but guesses which is the case.
- If he guesses correctly, he gets the penny.

Strategies must select the same (possibly mixed) action for each node in the information set.

Extensive form games

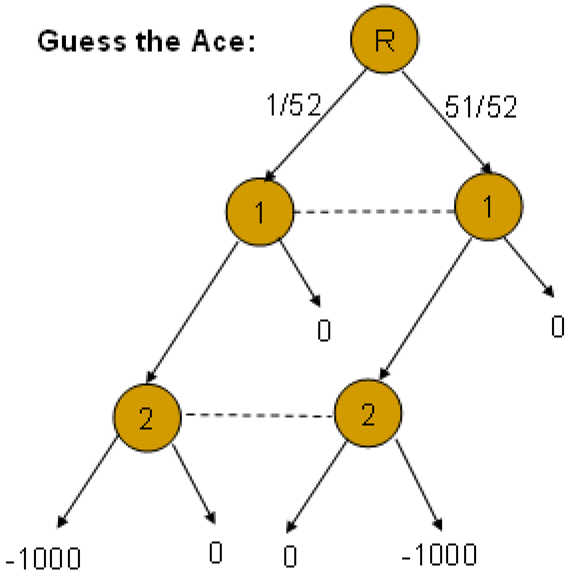
Guess the Ace:



- A deck of card is shuffled
- Either A is the top card or not
- Player 1 does not know if A is the top card or not.
- He can choose to end the game.
- If he does, no money is exchanged.
- Player 2 should now guess if A is the top card or not (he cannot see it).
- If he guesses correctly, Player 1 pays him $1000.

How should Player 2 play this game?

How to solve?



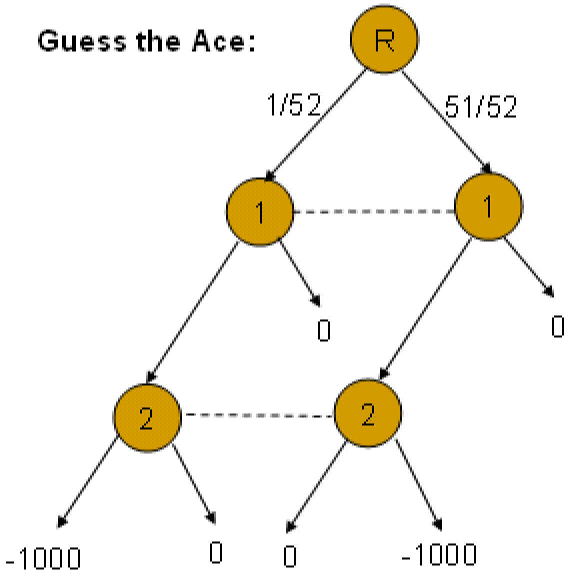
	Guess A	Guess Other
Stop	0	0
Play	-19.23	-980.77

Extensive form games can be converted into matrix games!

The rows and columns

- A pure strategy for a player (row or column in matrix) is a vector consisting of one designated action to make in *each* information set belonging to him.
- A mixed strategy is a distribution over pure strategies.

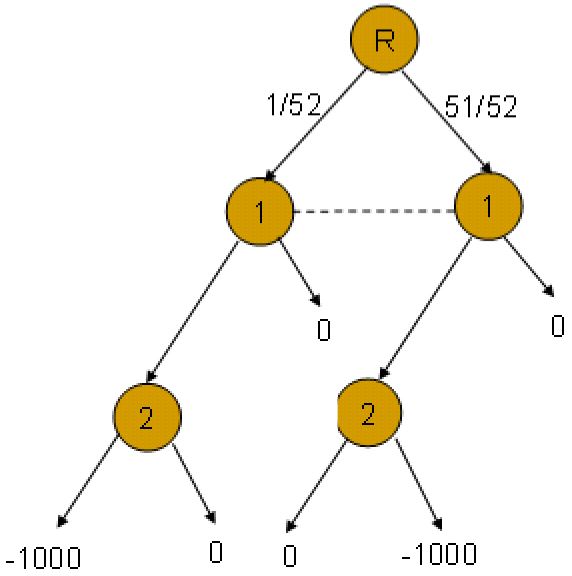
Done?

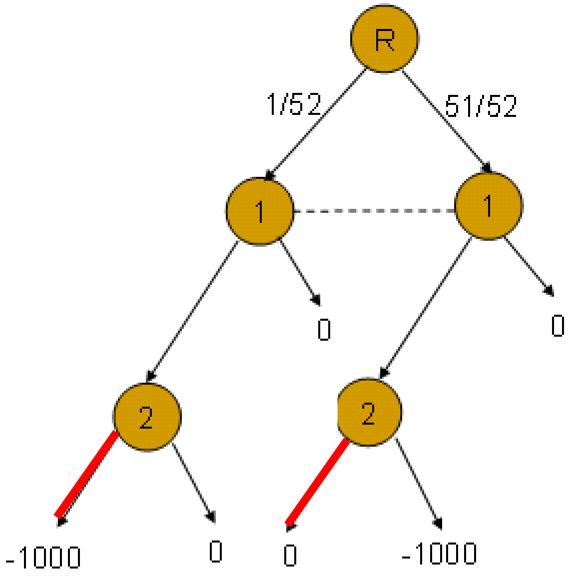


	Guess A	Guess Other
Stop	0	0
Play	-19.23	-980.77

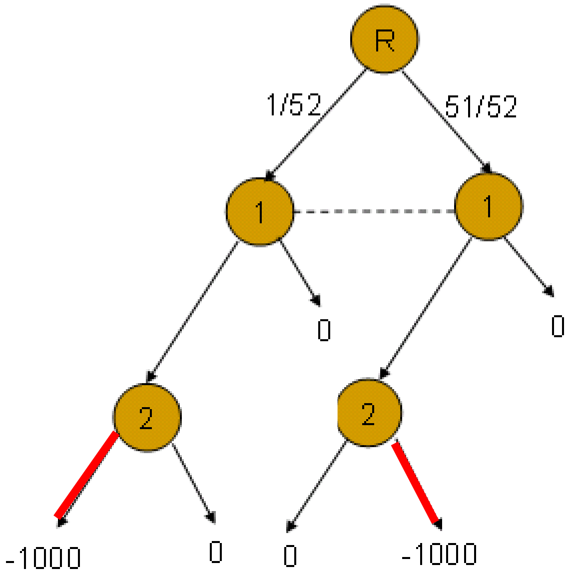
Exponential blowup in size!

Extensive form games can be converted into matrix games!

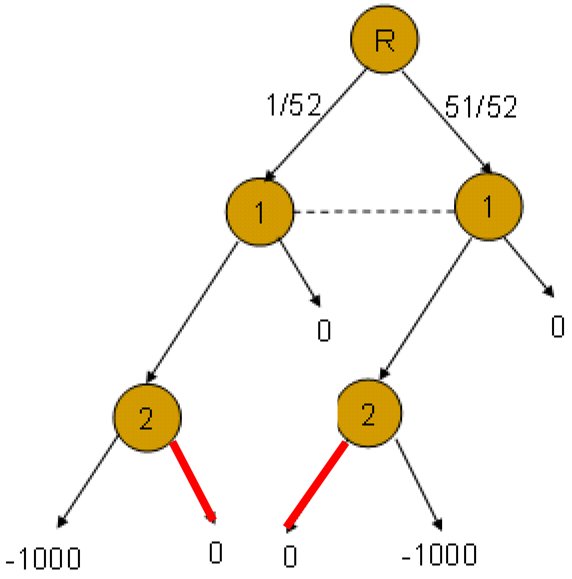

R
1/52
51/52
1
1
0
0
2
2
-1000
0
0
-1000



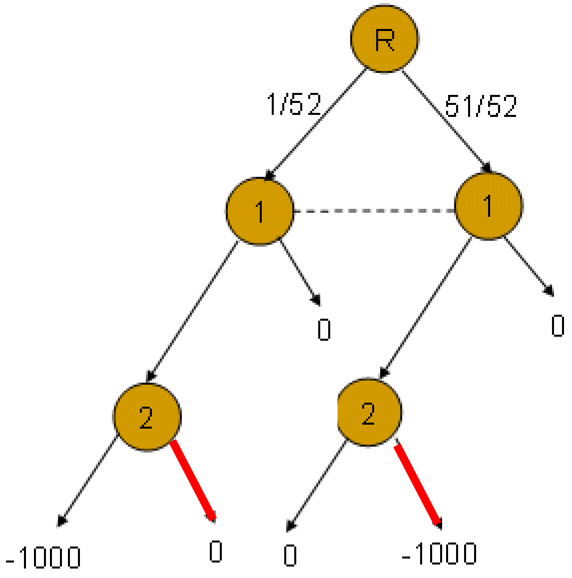
LL



LL **LR**



LL LR RL



LL LR RL RR

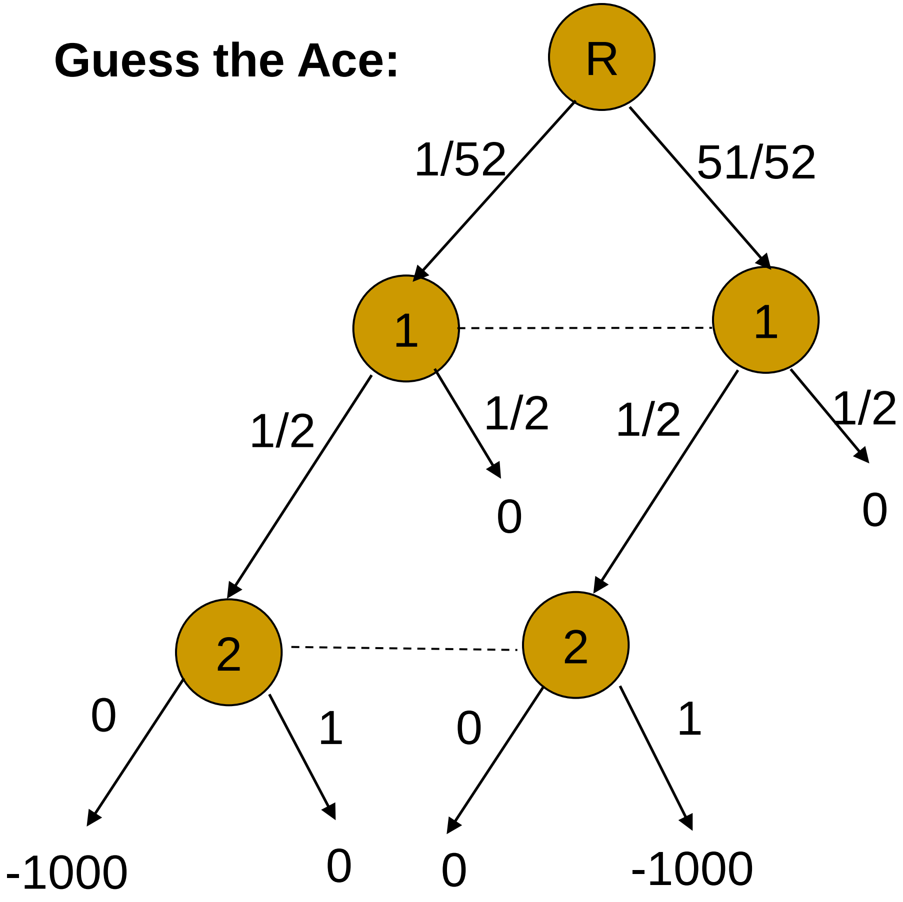
n information sets each with binary choice → 2^n columns

Behavior strategies (Kuhn, 1952)

- A behavior strategy for a player is a family of probability distributions, one for each information set, the distribution being over the ***actions*** one can make there.

Behavior strategies

Guess the Ace:



Behavior strategies

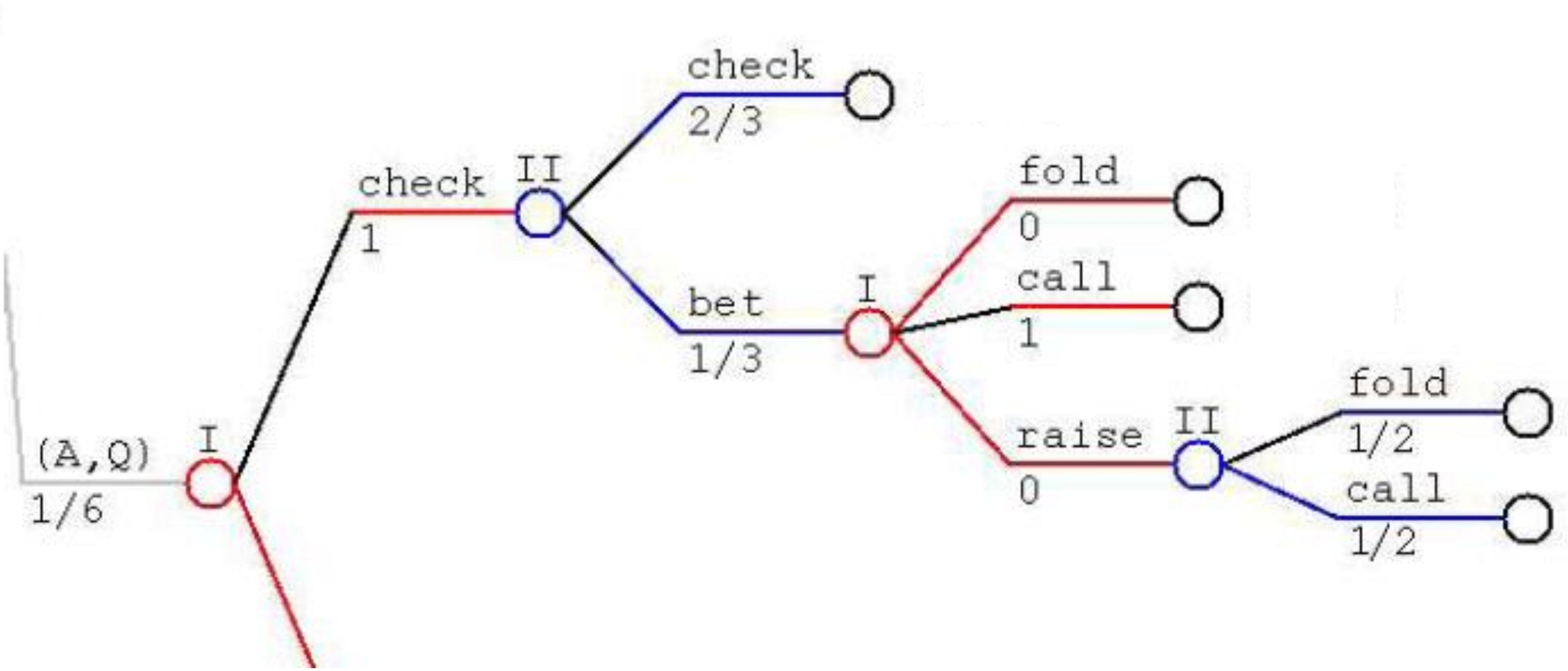
- Unlike mixed strategies, behavior strategies are compact objects.
- For games of **perfect recall**, behavior strategies and mixed strategies are equivalent (Kuhn, 1952).
- Can we find minimax behavior strategies efficiently?
- **Problem:** The minimax condition is no longer described by a linear program!

Realization plans (sequence form)

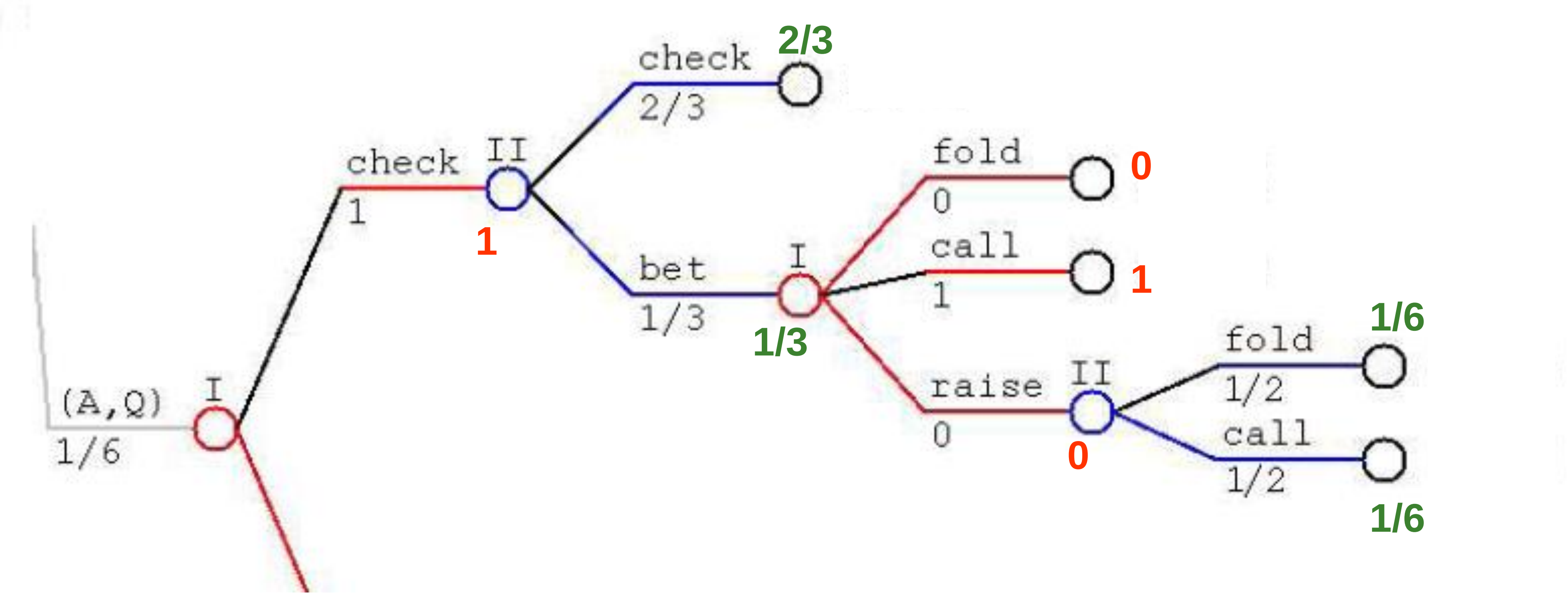
(Koller-Megiddo-von Stengel, 1994)

- Given a behavior strategy for a player, the **realization weight** of a ***sequence*** of moves is the **product** of probabilities assigned by the strategy to the moves in the sequence.
- If we have the realization weights for all sequences (a **realization plan**), we can deduce the corresponding behavior strategy (and vice versa).

Behavior strategies



Realization plans



(1,0,1,0,….) is a ***realization plan*** for Player I

(2/3, 1/3, 1/6, 1/6, …) is a ***realization plan*** for Player II

Crucial observation (Koller-Megiddo-von Stengel 1994)

- The set of valid realization plans for each of the two players (for games of perfect recall) is definable by a set of linear equations and positivity.

- The expected outcome of the game if Player 1 playing using realization plan $\mathbf{x}$ and Player 2 is playing using realization plan $\mathbf{y}$ is given by a bilinear form $\mathbf{x}^T A \mathbf{y}$.

- This implies that minimax realization plans can be found efficiently using linear programming!

Optimal response to fixed **x**.

- If MAXs plan is ***fixed*** to x, the best response by MIN is given by:

- Minimize $(\mathbf{x}^T A)\mathbf{y}$ so that $F\mathbf{y} = f$, $\mathbf{y} \geq 0$.

  ($F\mathbf{x} = f$, $\mathbf{y} \geq 0$ expressing that **y** is a realization plan.)

- The dual of this program is:

  Maximize $f^T \mathbf{q}$ so that $F^T \mathbf{q} \leq \mathbf{x}^T A$.

What should MAX do?

- If MAX plays **x** he should assume that MIN plays so that he obtains the value given by
  Maximize $f^T \mathbf{q}$ so that $F^T \mathbf{q} \leq \mathbf{x}^T A$.

- MAX wants to minimize this value, so his maximin strategy **y** is given by
  Maximize $f^T\mathbf{q}$ so that $F^T \mathbf{q} \leq \mathbf{x}^T A$, $E\mathbf{x} = e$, $\mathbf{x} \geq 0$.

(E**x** = e, **x** $\geq$ 0 expressing that **x** is a realization plan)

KMvS linear program

$$
\begin{aligned}
\max_{x,q} \quad & f^\top q \\
\text{s.t.} \quad -A^\top x + F^\top q &\leq 0 \\
Ex &= e \\
x &\geq \mathbf{0}
\end{aligned}
$$

One constraint for each action (sequence) of player 2

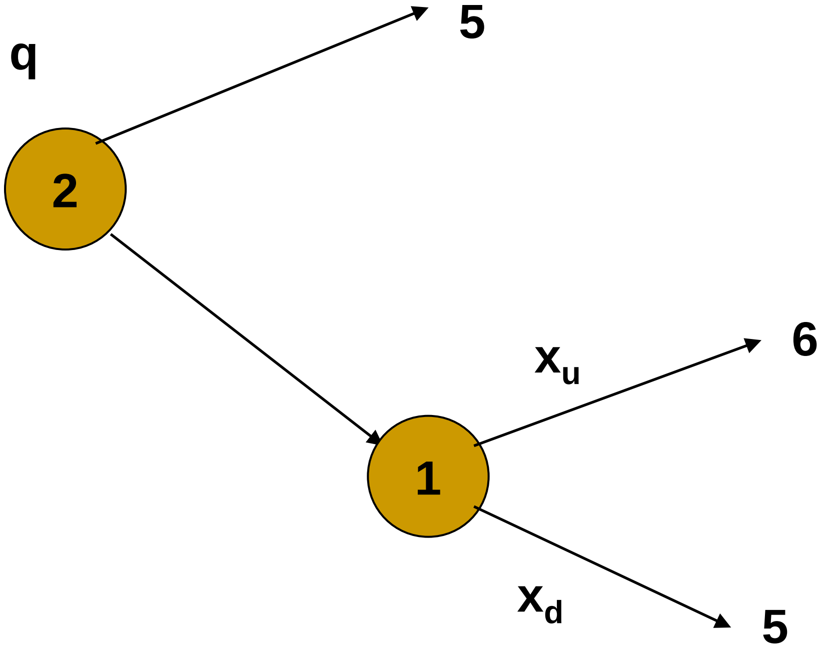
x is valid realization plan

x **– Realization plan for Player 1**

q – a “value” for each information set of Player 2

Up or down?

$q=5, \quad x_u = 1, x_d = 0$



Max q

$$q \leq 5$$

$$q \leq 6\,x_u + 5\,x_d$$

$$x_u + x_d = 1$$

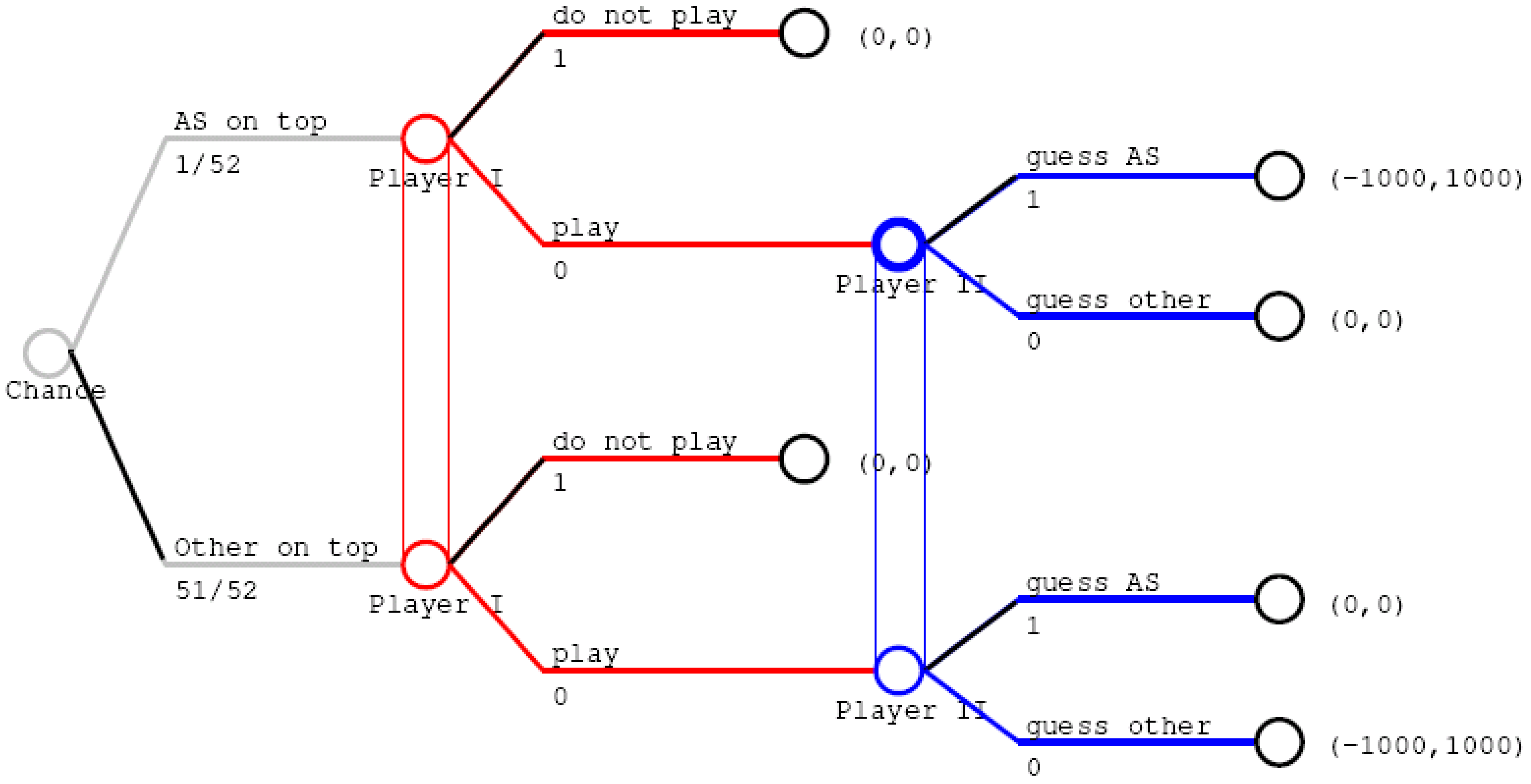
$$x_u, x_d \geq 0$$

Intuition:Left hand side of inequality in solution is what Player 2 *could* achieve, right hand side is what he *actually* achieves by this action.

KMvS algorithm in action

- Billings *et al.*, 2003: Solve abstraction of heads-up limit Texas Hold’Em.
- Gilpin and Sandholm 2005-2006: Fully solve limit Rhode Island Hold’Em. Better abstraction for limit Texas Hold’Em.
- Miltersen and Sørensen 2006: Rigorous approximation to optimal solution of no-limit Texas Hold’Em tournament.
- Gilpin, Sandholm and Sørensen 2007: Applied to 15 GB abstraction of limit Texas Hold’Em.
- It is included in the tool GAMBIT. Let’s try the GAMBIT implementation on Guess The Ace….

Guess-the-Ace, Nash equilibrium found by Gambit by KMvS algorithm



Complaints!

- [..] the strategies are not guaranteed to take advantage of mistakes when they become apparent. This can lead to very counterintuitive behavior. For example, assume that player 1 is guaranteed to win \$1 against an optimal player 2. But now, player 2 makes a mistake which allows player 1 to immediately win \$10000. It is perfectly consistent for the ‘optimal’ (maximin) strategy to continue playing so as to win the \$1 that was the original goal.

  Koller and Pfeffer, 1997.

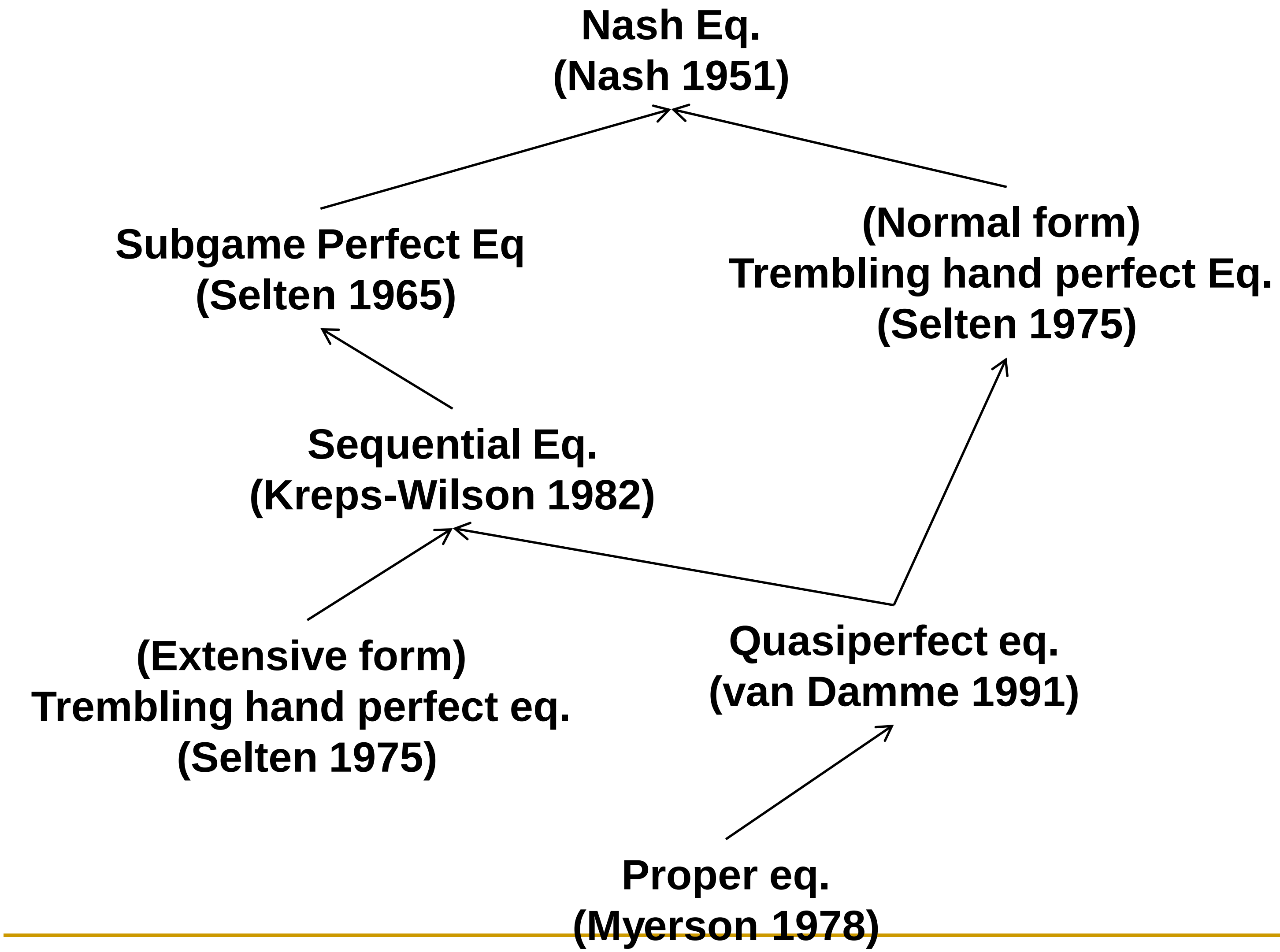
- If you run an=1 bl=1 it tells you that you should fold some hands (e.g. 42s) when the small blind has only called, so the big blind could have checked it out for a free showdown but decides to muck his hand. Why is this not necessarily a bug? (This had me worried before I realized what was happening).

  Selby, 1999.

Plan

- ***Representing*** finite-duration, imperfect information, two-player zero-sum games and ***computing*** minimax strategies. ✓
- Issues with minimax strategies. ✓
- Equilibrium refinements (a crash course) and how refinements resolve the issues, and how to modify the algorithms to compute refinements.
- (If time) Beyond the two-player, zero-sum case.

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

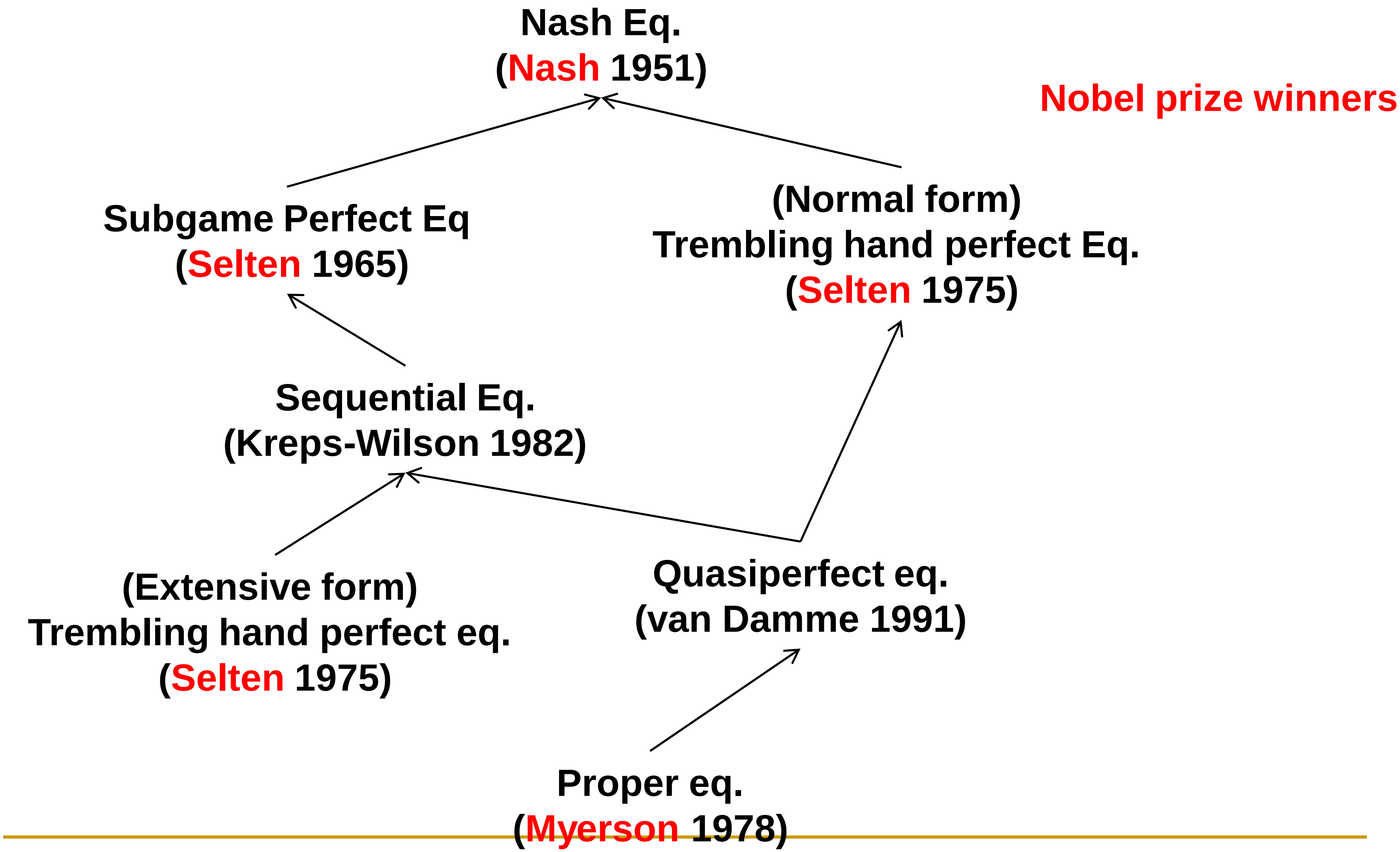
Sequential Eq.
(Kreps-Wilson 1982)

(Extensive form)
Trembling hand perfect eq.
(Selten 1975)

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Nobel prize winners

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

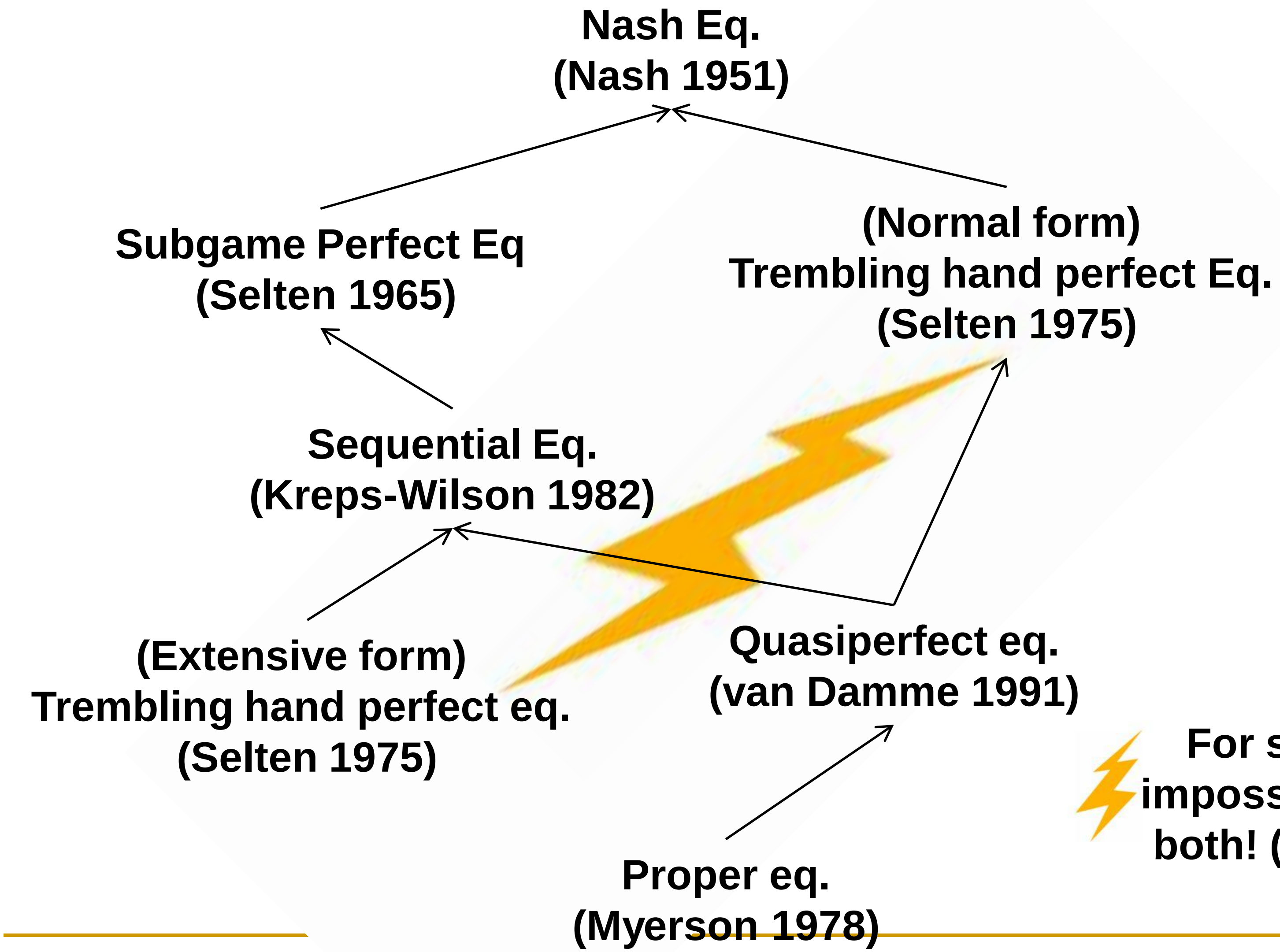
Sequential Eq.
(Kreps-Wilson 1982)

(Extensive form)
Trembling hand perfect eq.
(Selten 1975)

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

(Extensive form)
Trembling hand perfect eq.
(Selten 1975)

Quasiperfect eq.
(van Damme 1991)

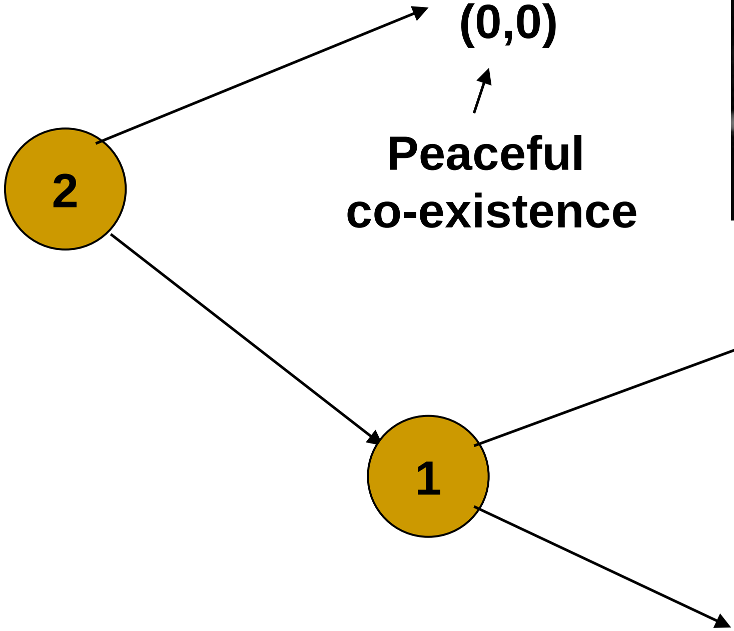
Proper eq.
(Myerson 1978)

⚡ **For some games, impossible to achieve both! (Mertens 1995)**

Subgame perfection (Selten '1965)

- First attempt at capturing ***sequential rationality***.
- An equilibrium is *subgame perfect* if it induces an equilibrium in all subgames.
- A *subgame* is a subtree of the extensive form that does not break any information sets.

Doomsday Game



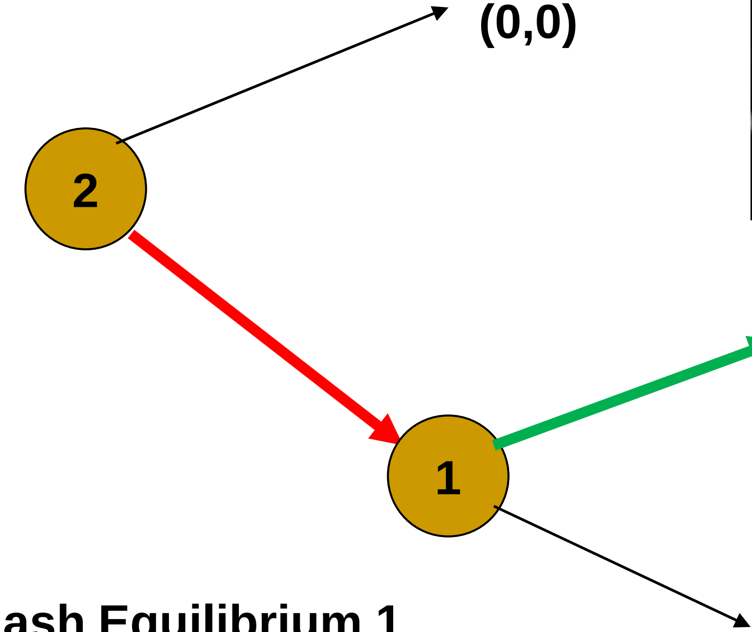
(0,0)

Peaceful co-existence

(-1,1) ← Invasion and surrender

(-100,-100) ←

Doomsday Game



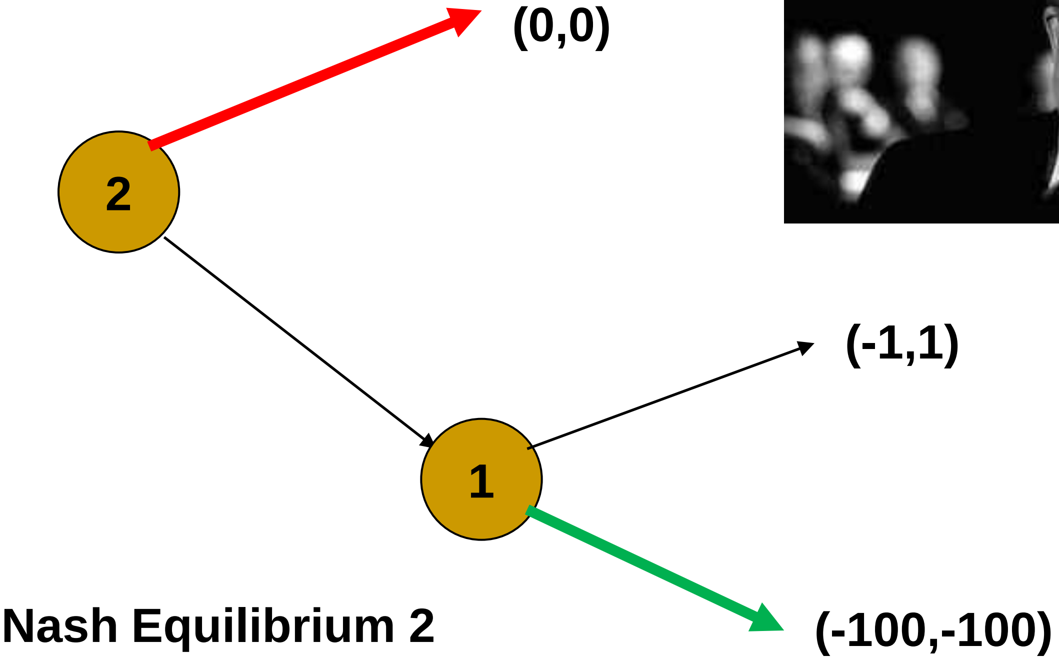
(0,0)

(-1,1)

(-100,-100)

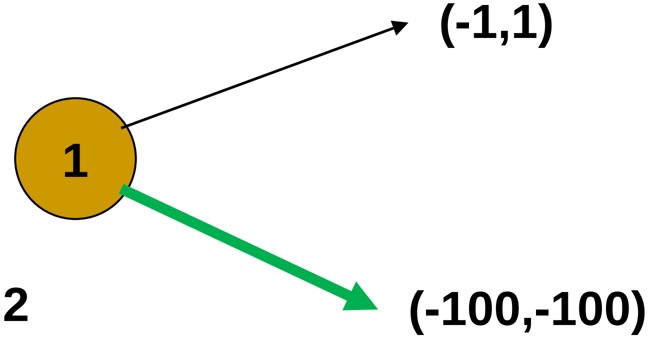
Nash Equilibrium 1

Doomsday Game



Nash Equilibrium 2

Doomsday Game

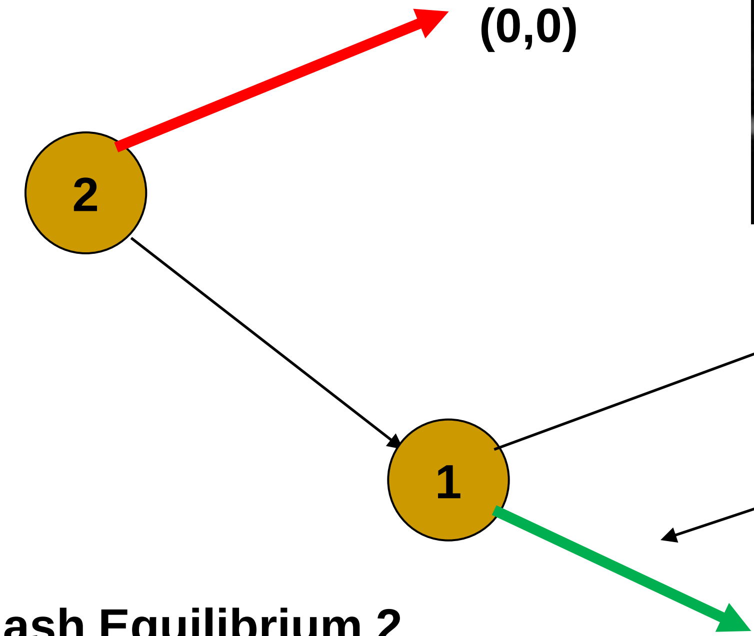

1

(-1,1)

(-100,-100)

Nash Equilibrium 2

Doomsday Game

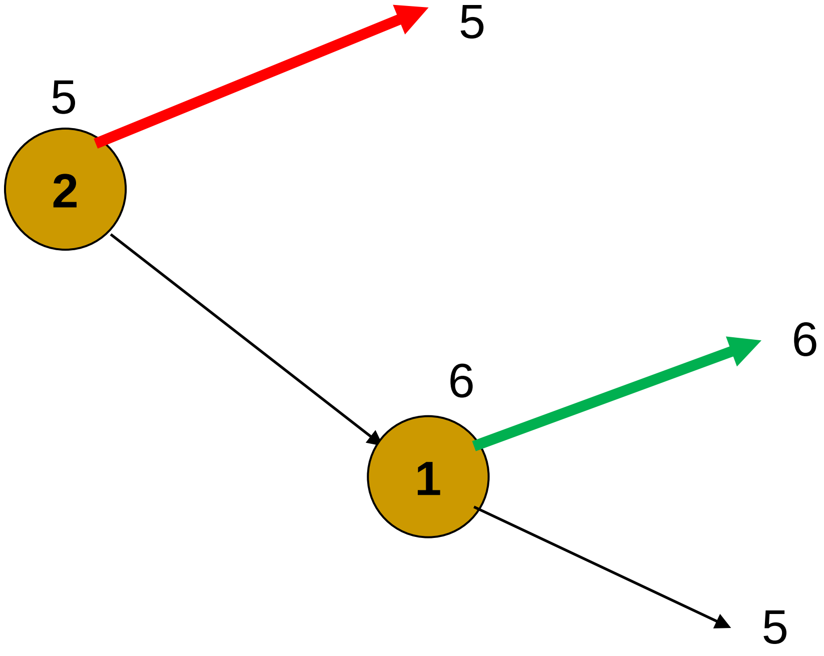


Non-credible threat

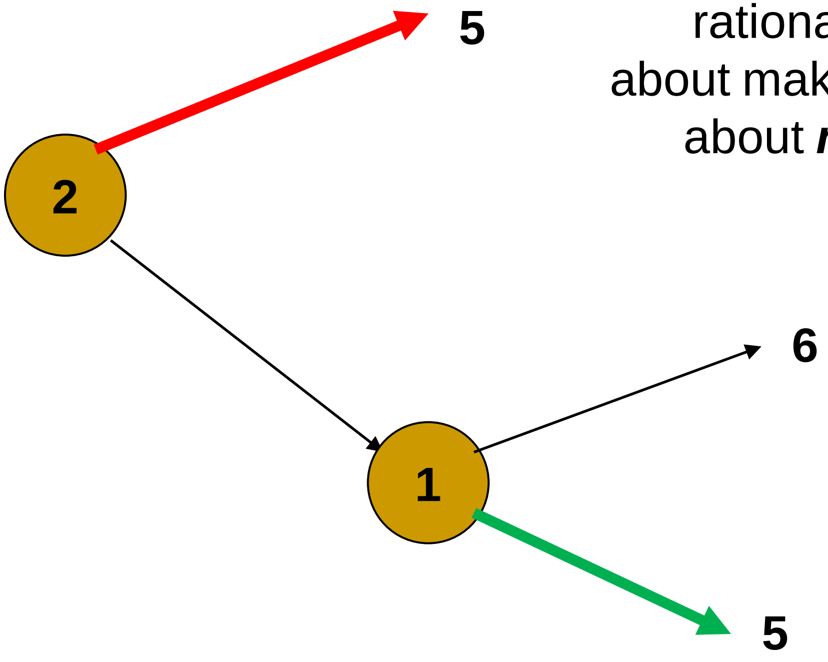
Nash Equilibrium 2

is not *subgame perfect.*

Nash eq. found by backwards induction



Another Nash equilibrium!



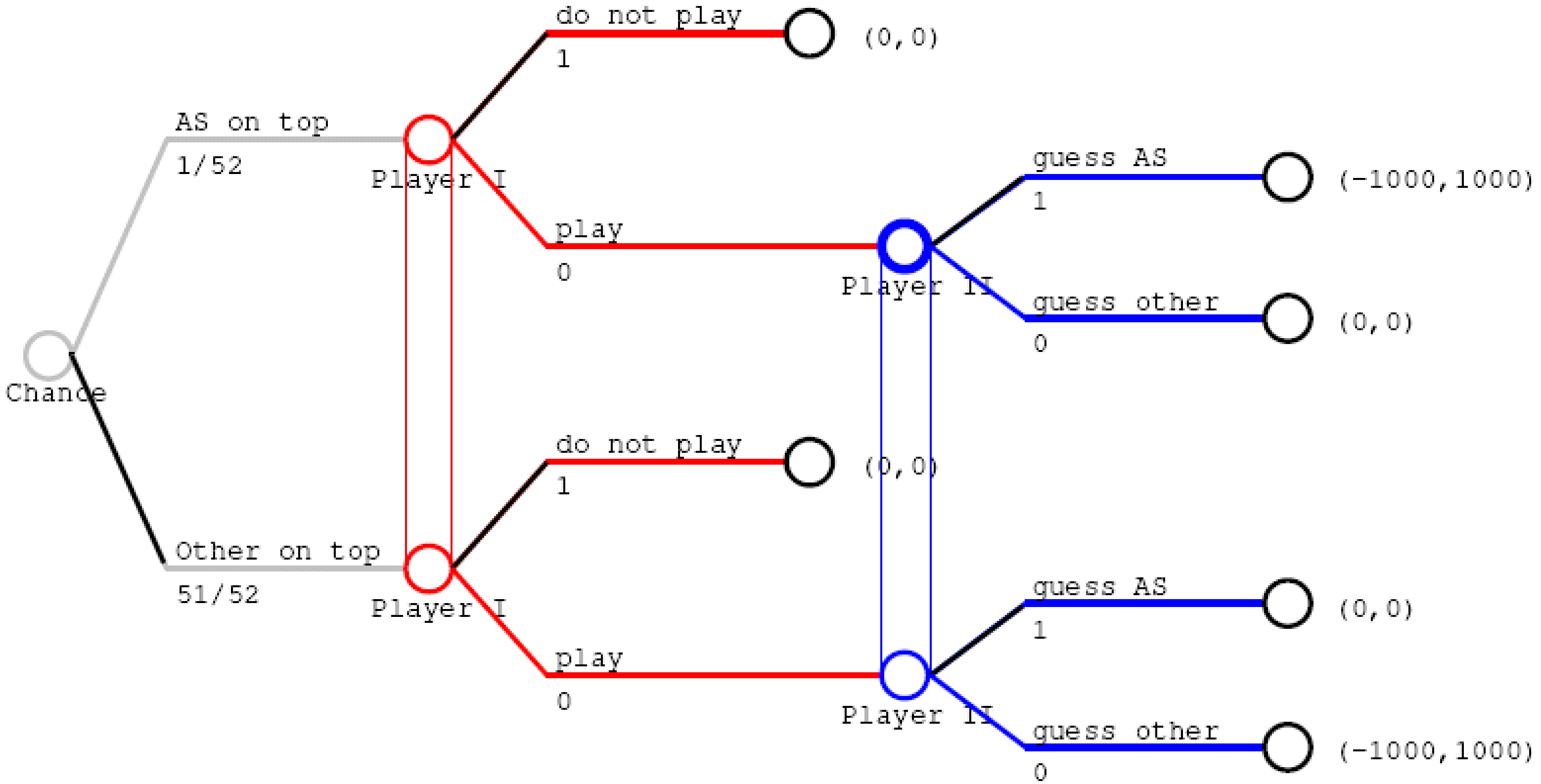
Not subgame perfect!

In zero-sum games, sequential rationality is not so much about making credible threats as about ***not returning gifts***

How to compute a subgame perfect equilibrium in a zero-sum game

- Solve each subgame separately.
- Replace the root of a subgame with a leaf with its computed value.

Guess-the-Ace, bad Nash equilibrium found by Gambit by KMvS algorithm

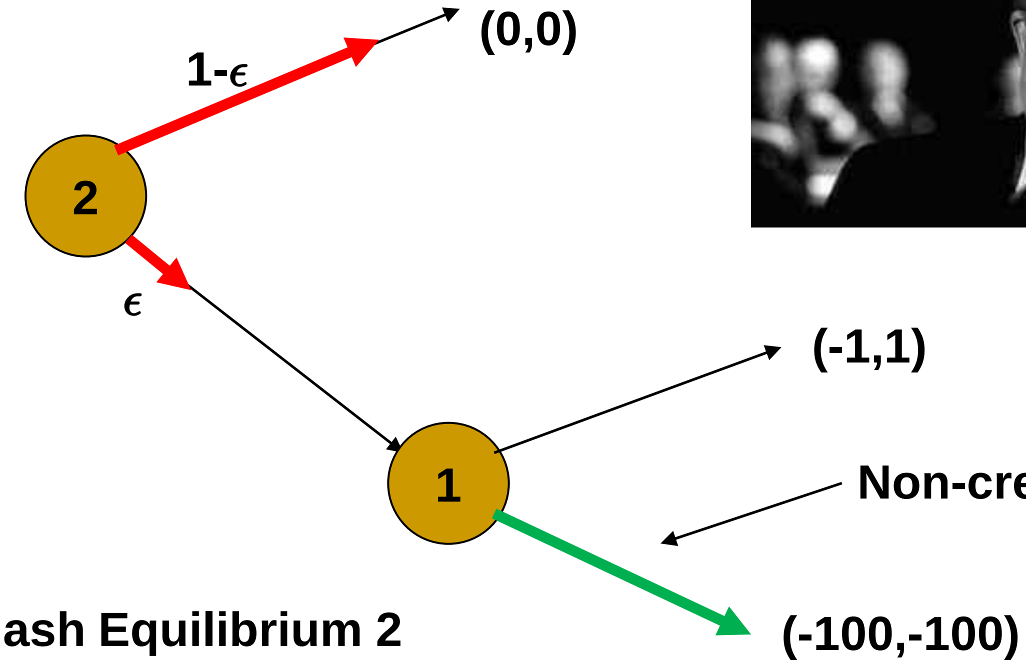


It's subgame perfect!

(Extensive form) trembling hand perfection (Selten'75)

- ***Perturbed game***: For each information set, associate a parameter $\epsilon > 0$ (a tremble). Disallow behavior probabilities smaller than this parameter.
- A limit point of equilibria of perturbed games as $\epsilon \to 0$ is an equilibrium of the original game and called ***trembling hand perfect.***
- **Intuition**: Think of ϵ as an infinitisimal (formalised in paper by Joe Halpern).

Doomsday Game

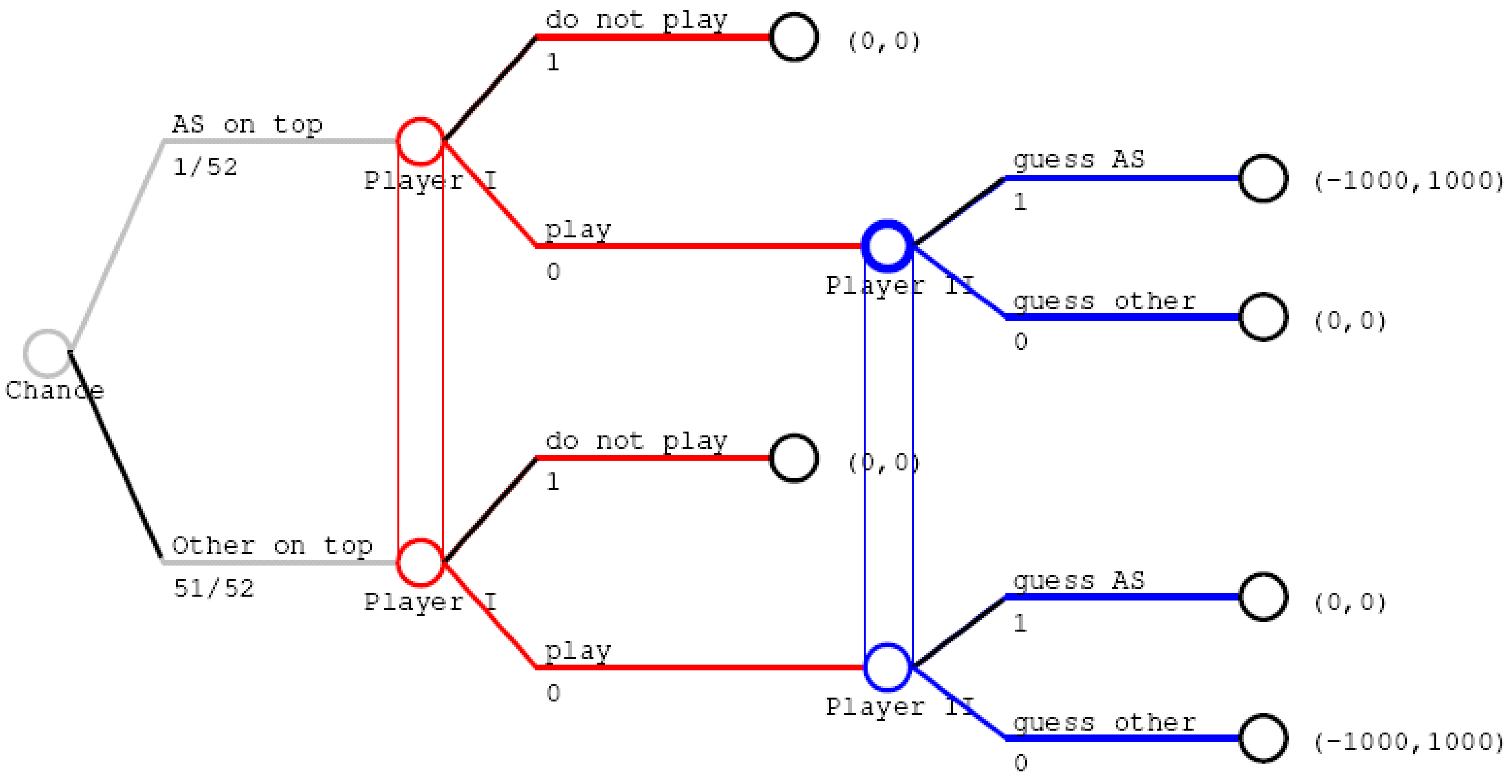


Non-credible threat

Nash Equilibrium 2

is not *trembling hand perfect:* If Player 1 worries just a little bit that Player 2 will attack, he will not commit himself to triggering the doomsday device

Guess-the-Ace, Nash equilibrium found by Gambit by KMvS algorithm

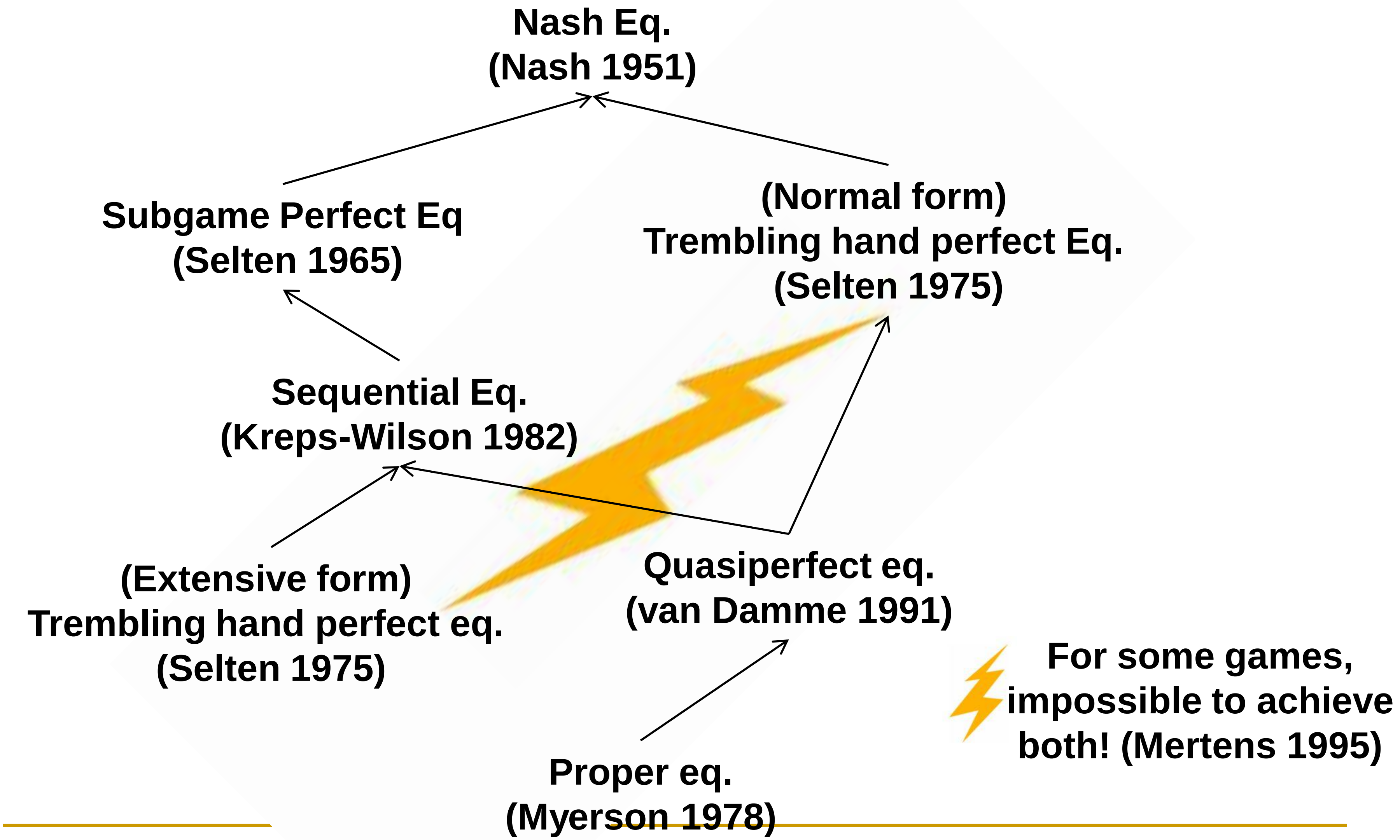


It’s not trembling hand perfect!

Computational aspects

- Can an extensive form trembling-hand perfect equilibrium be computed for a given zero-sum extensive form game (two player, perfect recall) in polynomial time?
- **Open problem**(!) (I think), but maybe not too interesting, as…

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

(Extensive form)
Trembling hand perfect eq.
(Selten 1975)

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

For some games, impossible to achieve both! (Mertens 1995)

(Normal form) trembling hand perfect equilibria

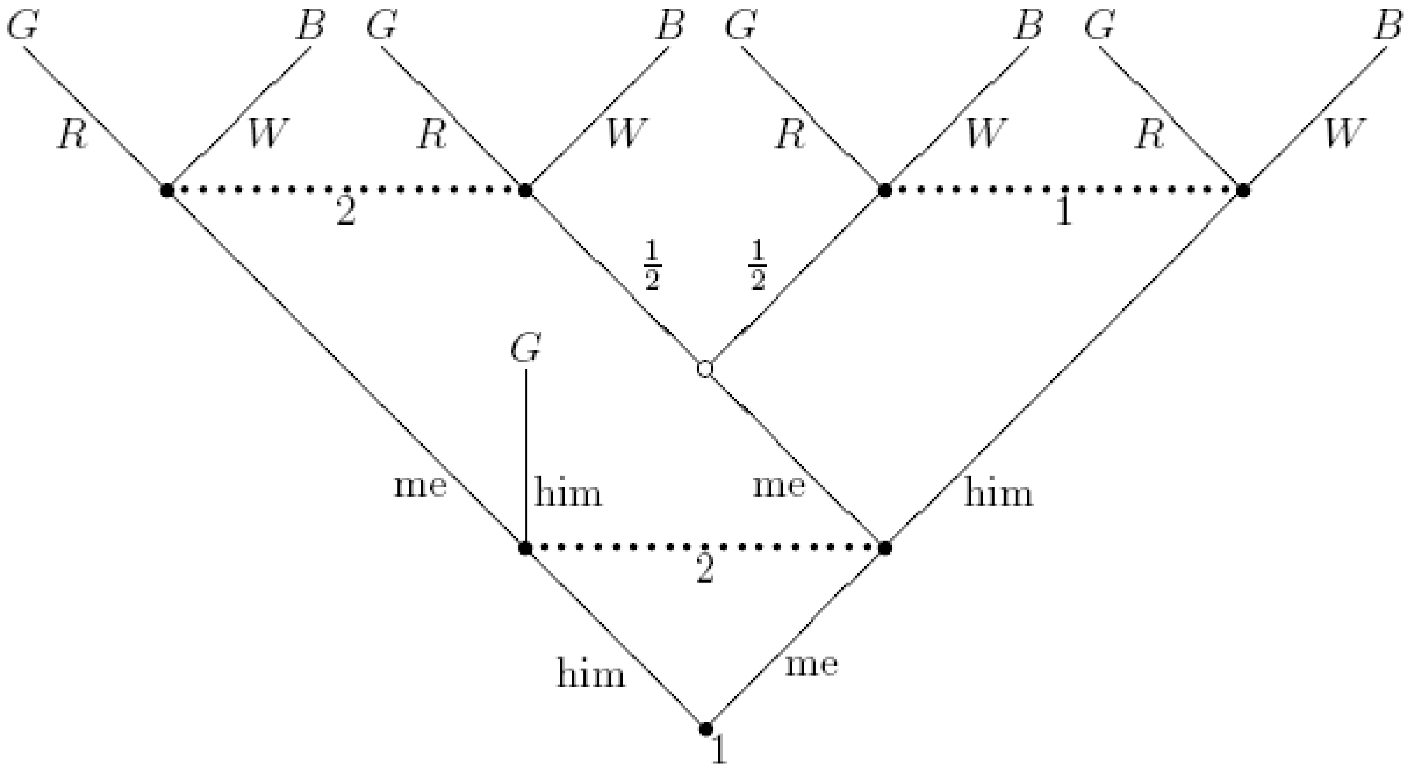
- Transform the game from extensive form to normal form.

- Transform the normal form back to an extensive form with just one information set for each player and apply the definition of extensive form trembling hand perfect equilibria.

- For a two-player game, a Nash equilibrium is normal form perfect if and only if it consists of two undominated strategies.

Mertens’ voting game

- Two players must elect one of them to perform an effortless task. The task may be performed either correctly or incorrectly.
- If it is performed correctly, both players receive a payoff of 1, otherwise both players receive a payoff of 0.
- The election is by a secret vote.
    - If both players vote for the same player, that player gets to perform the task.
    - If each player votes for himself, the player to perform the task is chosen at random but is not told that he was elected this way.
    - If each player votes for the other, the task is performed by somebody else, with no possibility of it being performed incorrectly.


G
B
G
B
G
B
G
B
R
W
R
W
R
W
R
W
2
1
½
½
G
me
him
me
him
2
him
me
1

Normal form vs. Extensive form trembling hand perfection

- The normal form and the extensive form trembling hand perfect equilibria of Mertens’ voting game are ***disjoint:*** Any extensive form perfect equilibrium has to use a dominated strategy.

- One of the two players has to vote for the other guy.

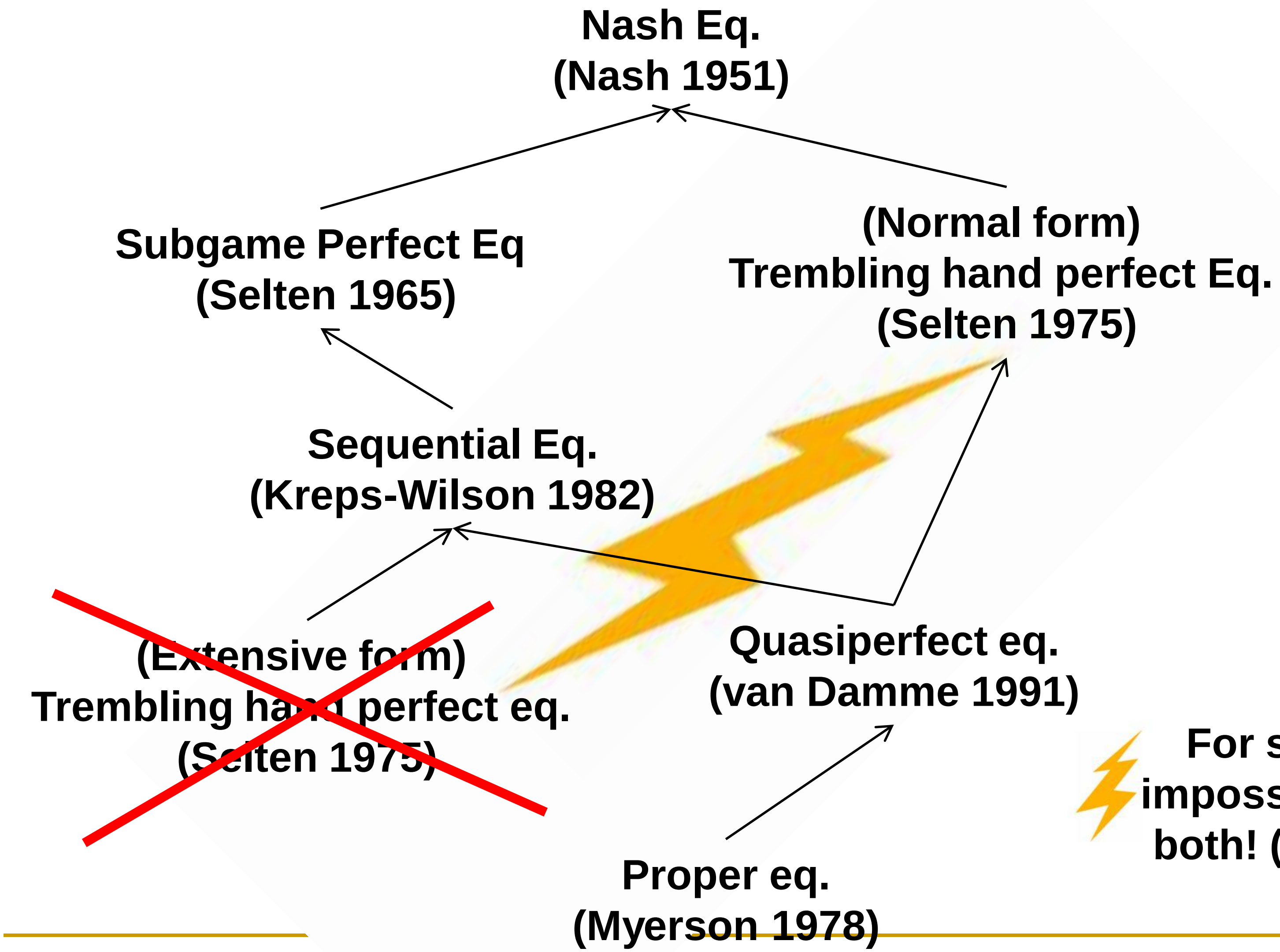
What’s wrong with the definition of trembling hand perfection?

- The extensive form trembling hand perfect equilibria are limit points of equilibria of perturbed games.
- In the perturbed game, the players ***agree*** on the relative magnitude of the trembles.
- This does not seem warranted!

Open problem

- Is there a zero-sum game for which the extensive form and the normal form trembling hand perfect equilibria are disjoint?

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

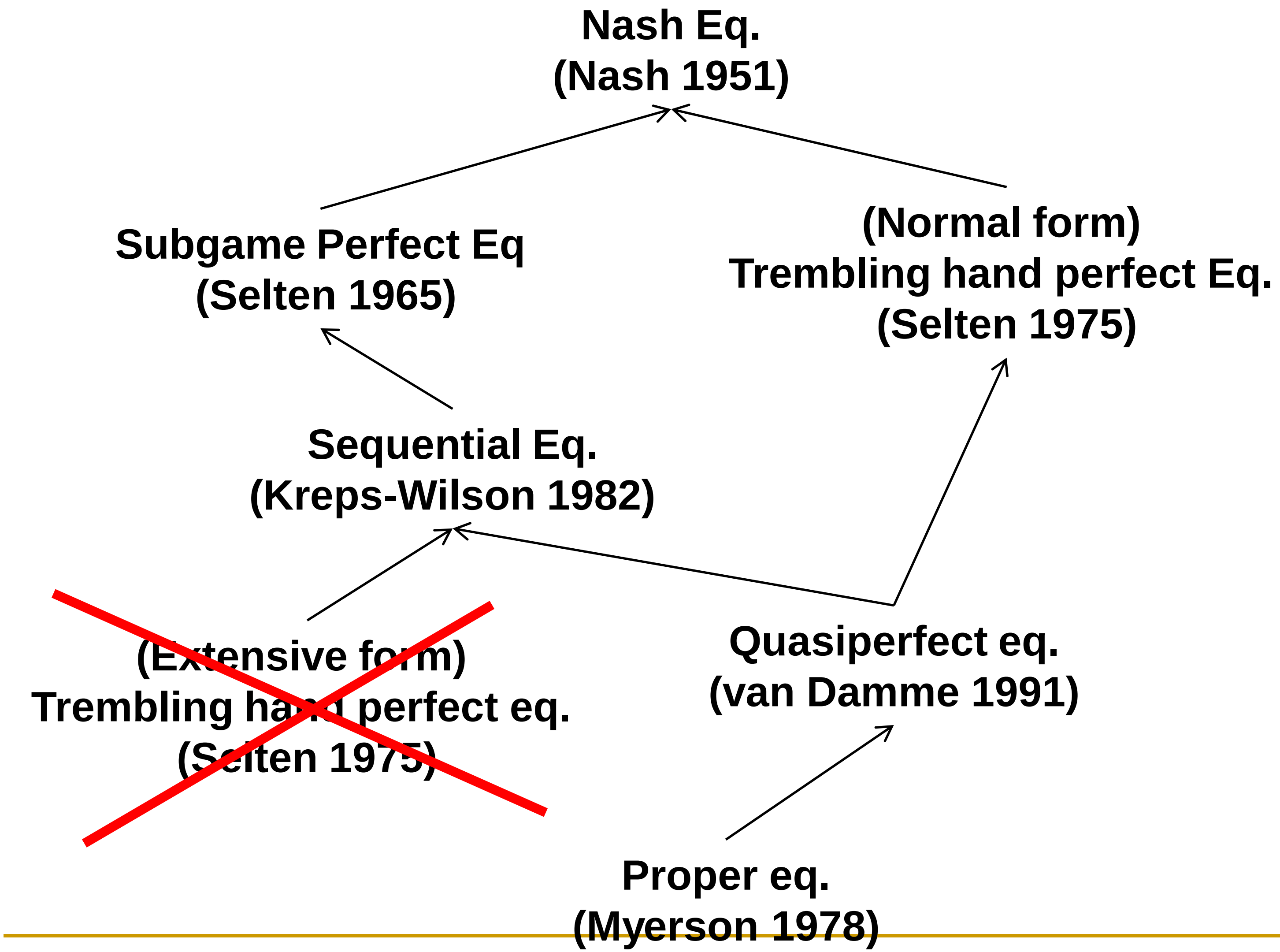
~~**(Extensive form)**~~
~~**Trembling hand perfect eq.**~~
~~**(Selten 1975)**~~

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

⚡ **For some games, impossible to achieve both! (Mertens 1995)**

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

~~(Extensive form)
Trembling hand perfect eq.
(Selten 1975)~~

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Computing a normal form perfect equilibrium of a zero-sum game, easy hack!

- Compute the value of the game using KMvS algorithm.

- Among all behavior plans achieving the value, find one that maximizes payoff against some fixed fully mixed strategy of the opponent.

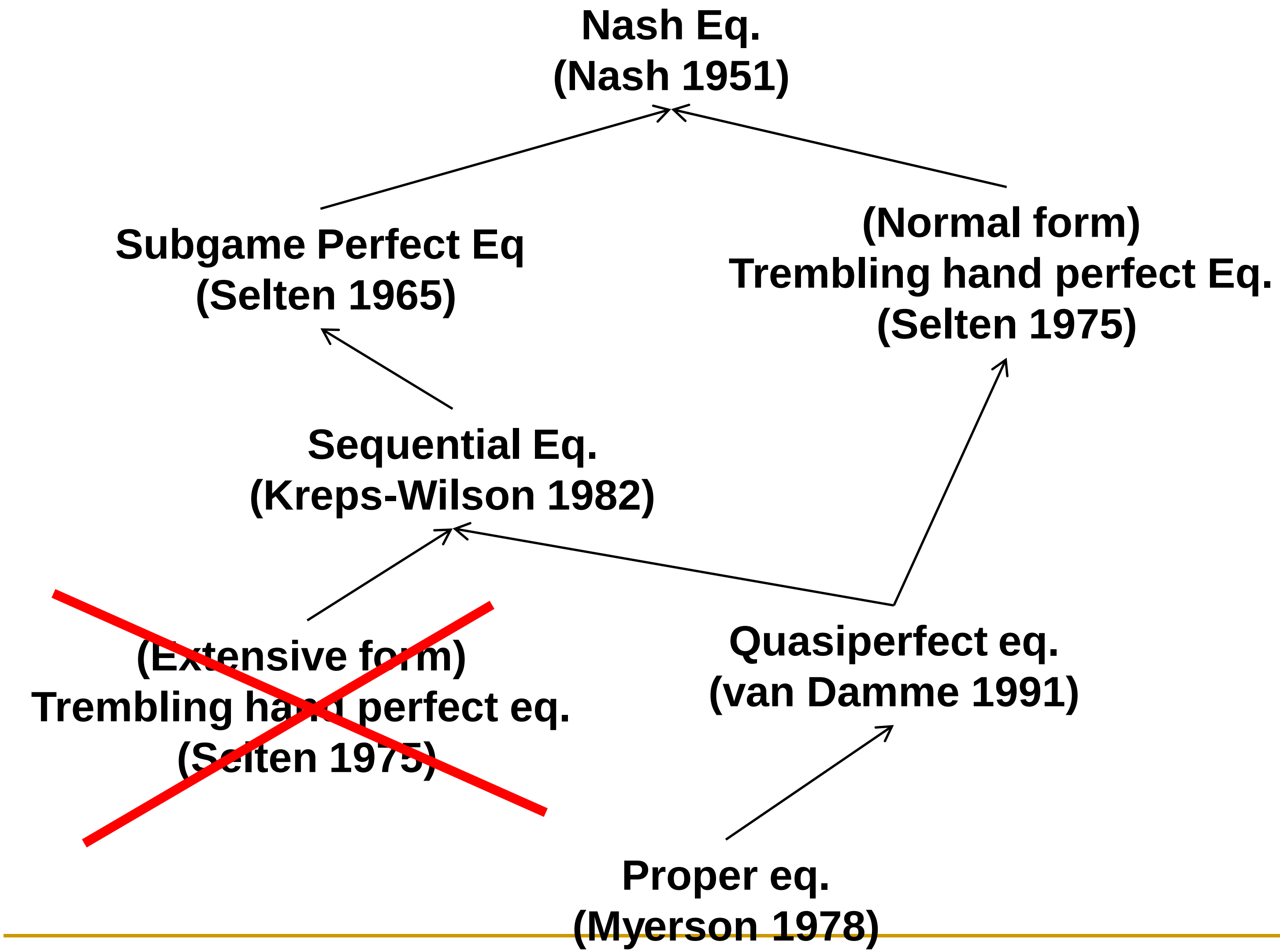
- ***But***: A normal form perfect equilibrium is not guaranteed to be sequentially rational (keep gifts).

Example of ”bad”(?) behavior in a normal form perfect equilibrium

- Rules of the game:
  - Player 2 can either stop the game or give Player 1 a dollar.
  - If Player 1 gets the dollar, he can either stop the game or give Player 2 the dollar back.
  - If Player 2 gets the dollar, he can either stop the game or give Player 1 two dollars.

It is part of a normal form perfect equilibrium for Player 1 to give the dollar back if he gets it.

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

~~**(Extensive form)**~~
~~**Trembling hand perfect eq.**~~
~~**(Selten 1975)**~~

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Sequential Equilibria

(Kreps and Wilson, 1982)

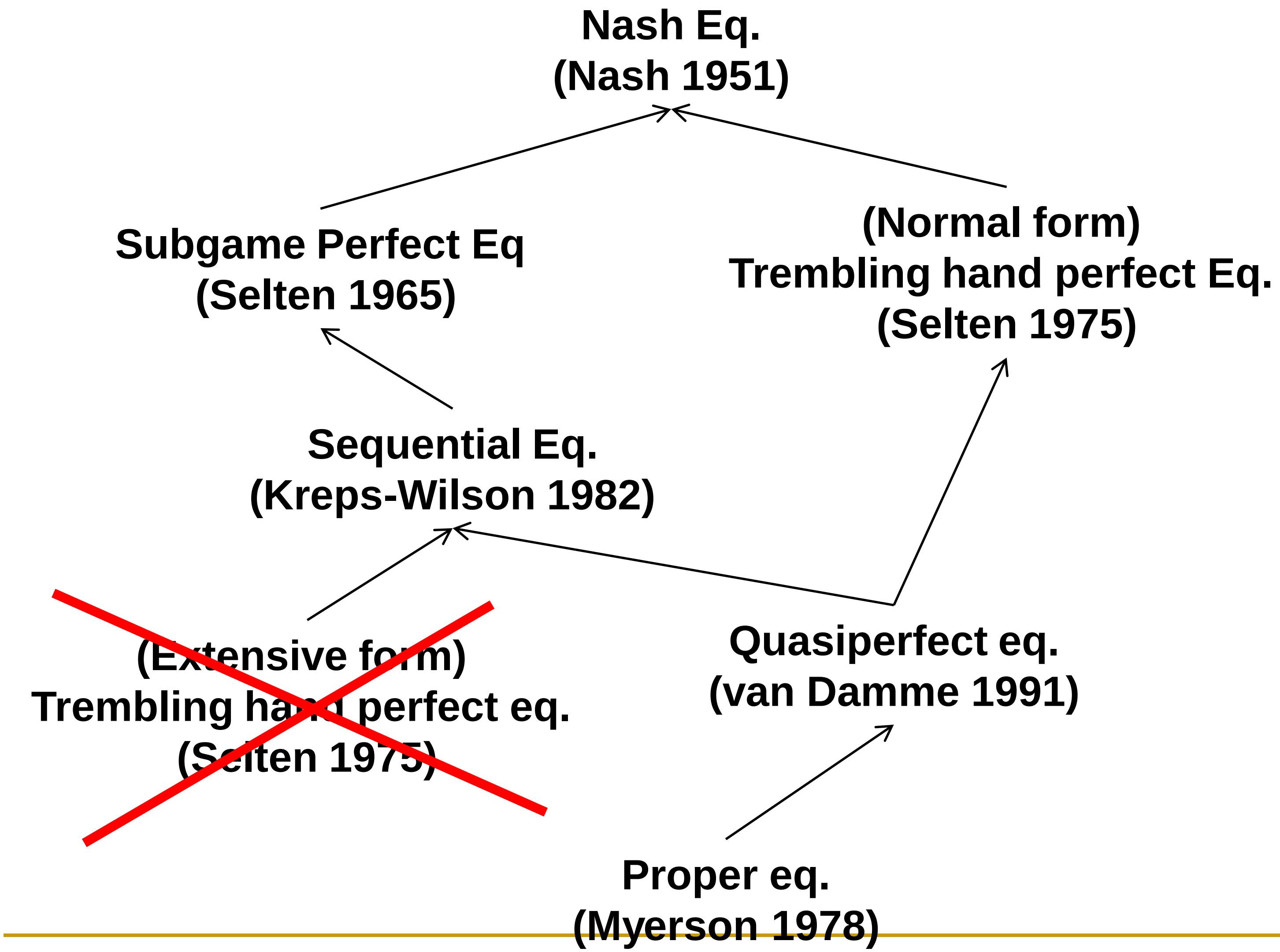
- In addition to prescribing two strategies, the equilibrium prescribes to every information set a ***belief:*** A probability distribution on nodes in the information set.
- At ***each*** information set, the strategies should be “sensible”, given the beliefs.
- At ***each*** information set, the beliefs should be “sensible”, given the strategies.

- Unfortunately, a sequential equilibrium may use dominated strategies.

Sequential equilibrium using a dominated strategy

- Rules of the game:
  - Player 1 either stops the game or asks Player 2 for a dollar.
  - Player 2 can either refuse or give Player 1 a dollar
- It is part of a sequential equilibrium for Player 1 to stop the game and ***not*** ask Player 2 for a dollar.
- Intuition: A sequential equilibrium reacts correctly to mistakes done in the past but does not anticipate mistakes that may be made in the future.

Equilibrium Refinements



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

~~**(Extensive form)**~~
~~**Trembling hand perfect eq.**~~
~~**(Selten 1975)**~~

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Quasi-perfect equilibrum (van Damme, 1991)

- A quasi-perfect equilibrium is a limit point of ***ϵ-quasiperfect behavior strategy profile*** as $\epsilon > 0$.
- An ϵ-quasi perfect strategy profile satisfies that if some action is not a ***local best response*,** it is taken with probability at most ϵ.
- An action a in information set h is a local best response if there is a plan π for completing play after taking a, so that best possible payoff is achieved among all strategies agreeing with π except possibly at h and afterwards.
- Intuition: A player trusts himself over his opponent to make the right decisions in the future – this avoids the anomaly pointed out by Mertens.
- *By some irony of terminology, the ”quasi”-concept seems in fact far superior to the original unqualified perfection.* Mertens, 1995.

Computing quasi-perfect equilibrium

M. and Sørensen, *SODA'06* and *Economic Theory*, 2010.

- Shows how to modify the linear programs of Koller, Megiddo and von Stengel using ***symbolic perturbations*** ensuring that a quasi-perfect equilibrium is computed.

- Generalizes to non-zero sum games using linear complementarity programs.

- Solves an open problem stated by the computational game theory community: How to compute a sequential equilibrium using realization plan representation (McKelvey and McLennan) and gives an alternative to an algorithm of von Stengel, van den Elzen and Talman for computing an nromal form perfect equilibrium.

Perturbed game G(ε)

- G(ε) is defined as G **except** that we put a constraint on the mixed strategies allowed:

A position that a player reaches after making d moves must have realization weight at least ε^d.

Facts

- $G(\varepsilon)$ has an equilibrium for sufficiently small $\varepsilon>0$.
- An expression for an equilibrium for $G(\varepsilon)$ can be found in practice using the simplex algorithm, keeping ε a symbolic parameter representing sufficiently small value.
- An expression can also be found in worst case polynomial time by the ellipsoid algorithm.

Theorem

- When we let $\varepsilon \to 0$ in the behavior strategy equilibrium found for G(ε), we get a behavior strategy profile for the original game G. This can be done symbolically
- This strategy profile is a quasi-perfect equilibrium for G.
- … note that this is perhaps surprising - one could have feared that an extensive form perfect equilibrium was computed.

Questions about quasi-perfect equilibria

- Is the set of quasi-perfect equilibria of a zero-sum game 2-player game a Cartesian product (as the sets of Nash and normal-form proper equilibria are)?
- Can the set of quasi-perfect equilibria be polyhedrally characterized/computed (as the sets of Nash and normal-form proper equilibria can)?

All complaints taken care of?

- ~~[..] the strategies are not guaranteed to take advantage of mistakes when they become apparent. This can lead to very counterintuitive behavior. For example, assume that player 1 is guaranteed to win $1 against an optimal player 2. But now, player 2 makes a mistake which allows player 1 to immediately win $10000. It is perfectly consistent for the ‘optimal’ (maximin) strategy to continue playing so as to win the $1 that was the original goal.~~

  Koller and Pfeffer, 1997.

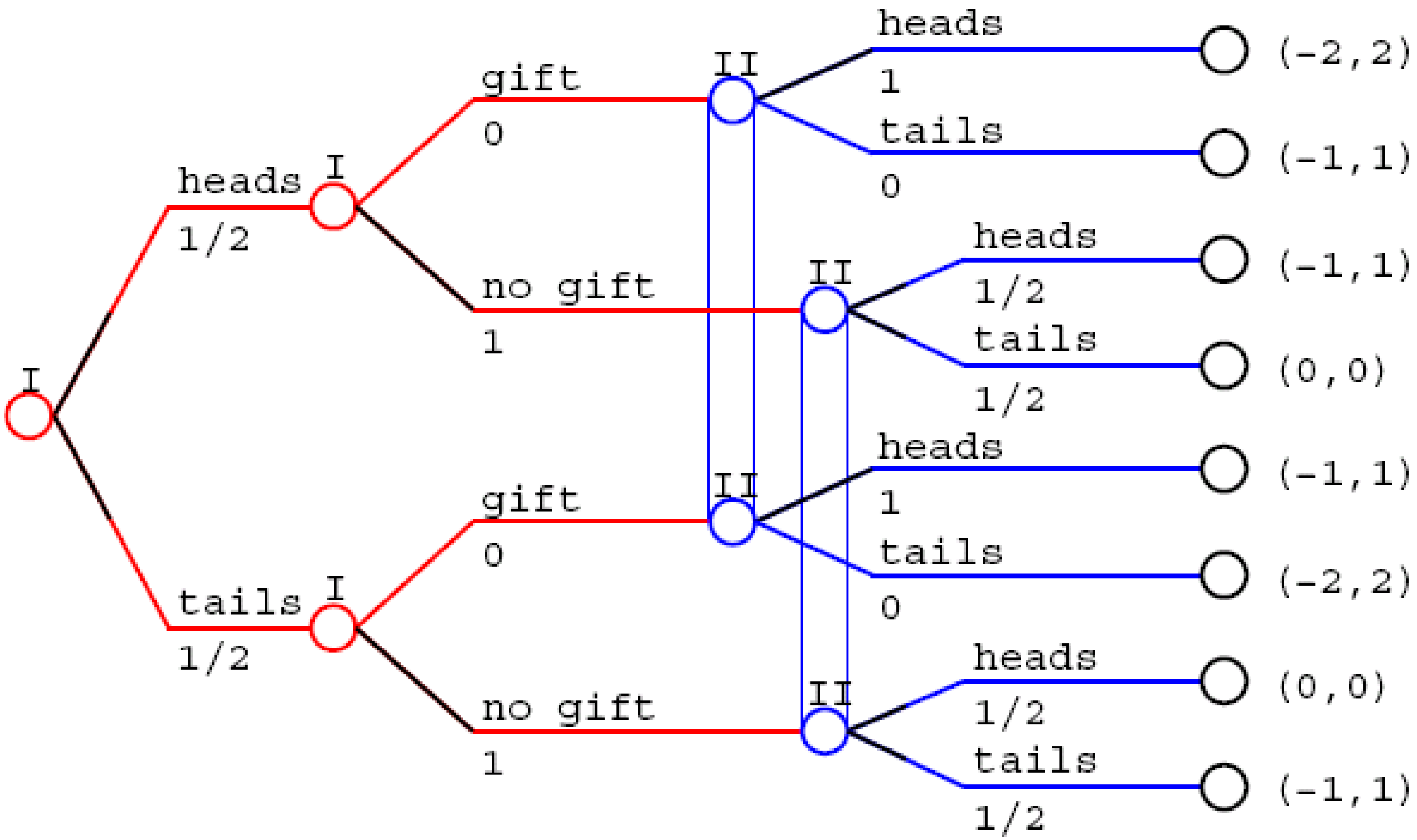
- ~~If you run an=1 bl=1 it tells you that you should fold some hands (e.g. 42s) when the small blind has only called, so the big blind could have checked it out for a free showdown but decides to muck his hand. Why is this not necessarily a bug? (This had me worried before I realized what was happening).~~

  Selby, 1999.

Matching Pennies on Christmas Morning

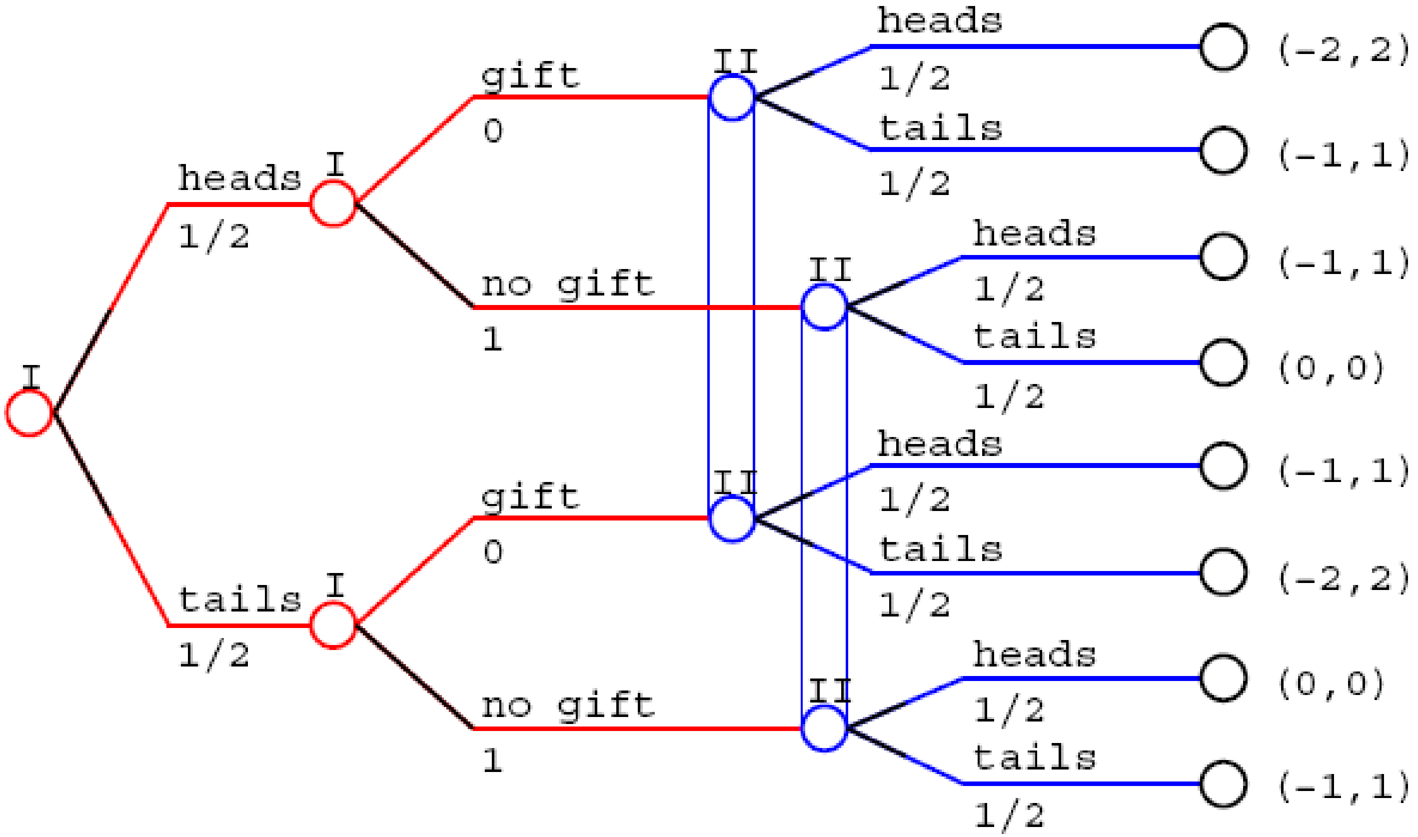
- Player 1 hides a penny.
- If Player 2 can guess if it is heads up or tails up, he gets the penny.
- How would you play this game (Matching Pennies) as Player 2?
- ***After*** Player I hides the penny but ***before*** Player 2 guesses, Player I has the option of giving Player 2 another penny, no strings attached (after all, it's Christmas).
- How would you play this game as Player 2?

Matching Pennies on Christmas Morning, bad Nash equilibrium



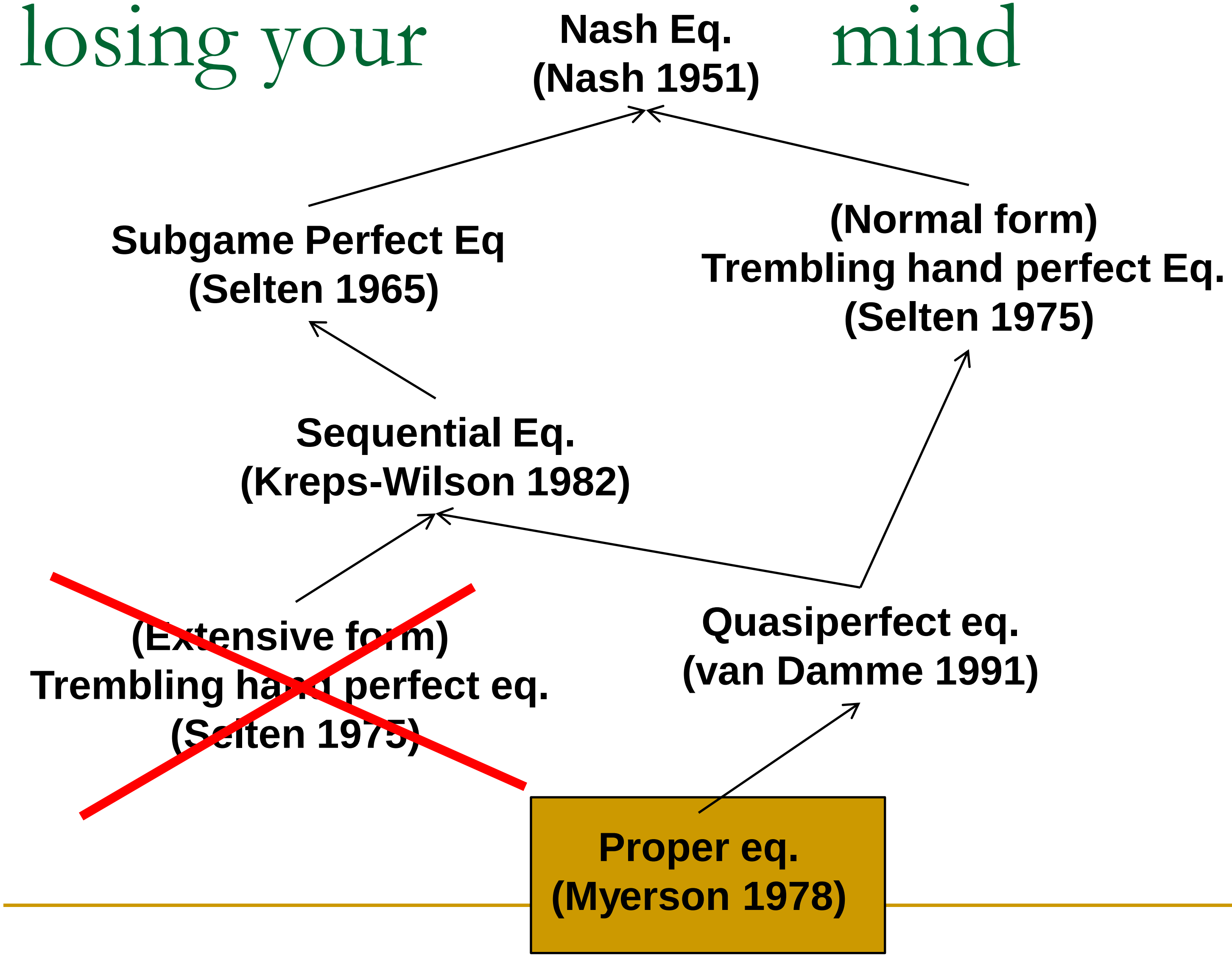
The bad equilibrium is quasi-perfect!

Matching Pennies on Christmas Morning, good equilibrium



The good equilibrium is not a basic solution to the KMvS LP!

How to celebrate Christmas without losing your mind



Nash Eq.
(Nash 1951)

Subgame Perfect Eq
(Selten 1965)

(Normal form)
Trembling hand perfect Eq.
(Selten 1975)

Sequential Eq.
(Kreps-Wilson 1982)

~~**(Extensive form)**~~
~~**Trembling hand perfect eq.**~~
~~**(Selten 1975)**~~

Quasiperfect eq.
(van Damme 1991)

Proper eq.
(Myerson 1978)

Normal form proper equilibrium (Myerson ’78)

- A limit point as $\varepsilon \to 0$ of ε-proper strategy profiles.

- An ε-proper strategy profile are two *fully mixed* strategies, so that for any two pure strategies i,j belonging to the same player, if j is a worse response than i to the mixed strategy of the other player, then $p(j) \leq \varepsilon\, p(i)$.

Normal form proper equilibrium (Myerson ’78)

Intuition:

- Players assume that the other player may make mistakes.
- Players assume that mistakes made by the other player are made in a ***rational manner***.

Normal-form properness

- The good equilibrium of Penny-Matching-on-Christmas-Morning is the unique normal-form proper one.

- Properness captures the assumption that mistakes are made in a “rational fashion”. In particular, after observing that the opponent gave a gift, we assume that ***apart from this*** he plays sensibly.

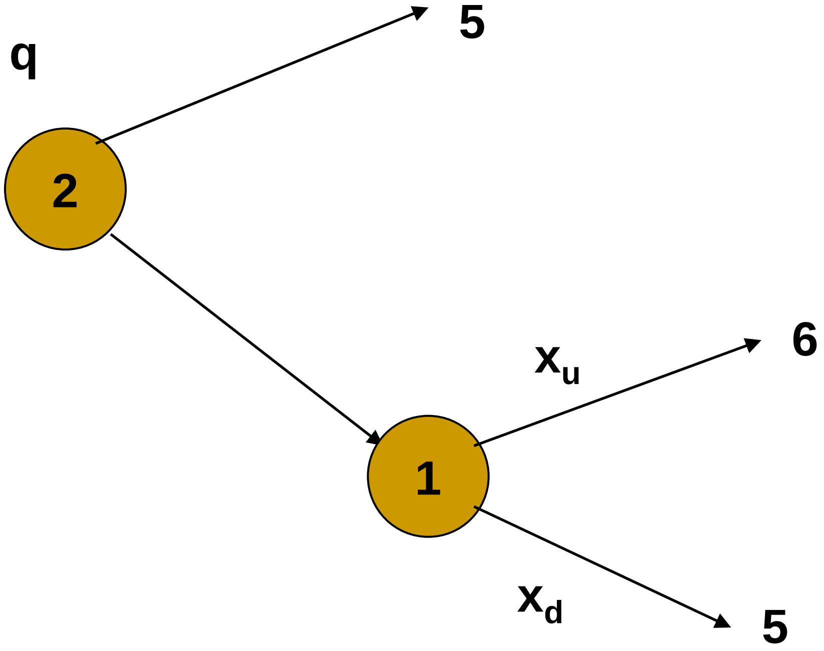
Properties of Proper equilibria of zero sum games (van Damme, 1991)

- The set of proper equilibria is a Cartesian product $D_1 \times D_2$ (as for Nash equlibria).
- Strategies of D_i are *payoff equivalent*: The choice between them is arbitrary against ***any*** strategy of the other player.

Miltersen and Sørensen, SODA 2008

- For imperfect information games, a normal form proper equilibrium can be found by solving a sequence of linear programs, based on the KMvS programs.
- The algorithm is based on finding solutions to the KMvS “balancing” the ***slack*** obtained in the inequalitites.

Up or down?



Max q

$q \leq 5$

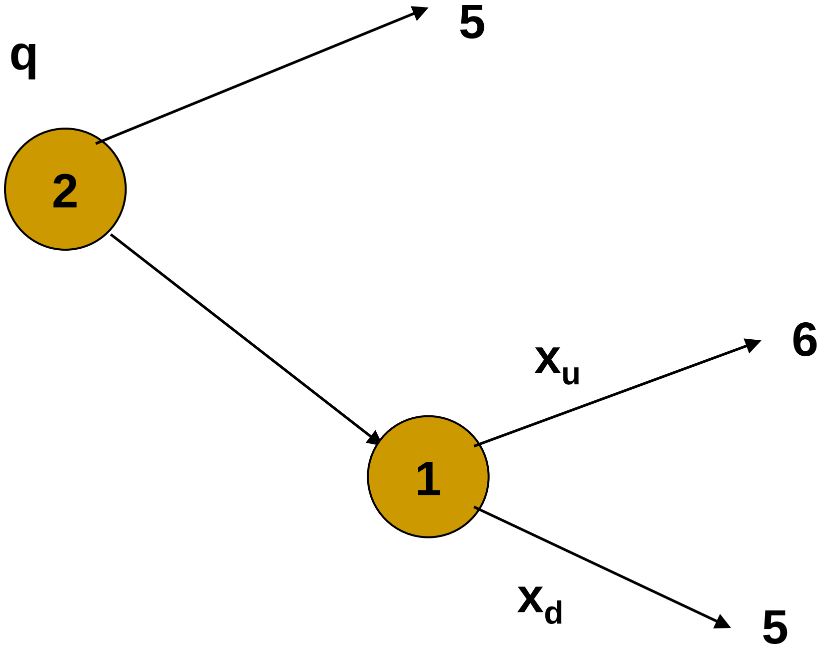
$q \leq 6\ x_u + 5\ x_d$

$x_u + x_d = 1$

$x_u, x_d \geq 0$

“Bad” optimal solution

$q=5, \quad x_u = 0, x_d = 1$



Max q

$q \leq 5$ ← No slack

$q \leq 6\, x_u + 5\, x_d$ ←

$x_u + x_d = 1$

$x_u, x_d \geq 0$

Good optimal solution

$q=5, \quad x_u = 1, x_d = 0$

$\textbf{Max } q$

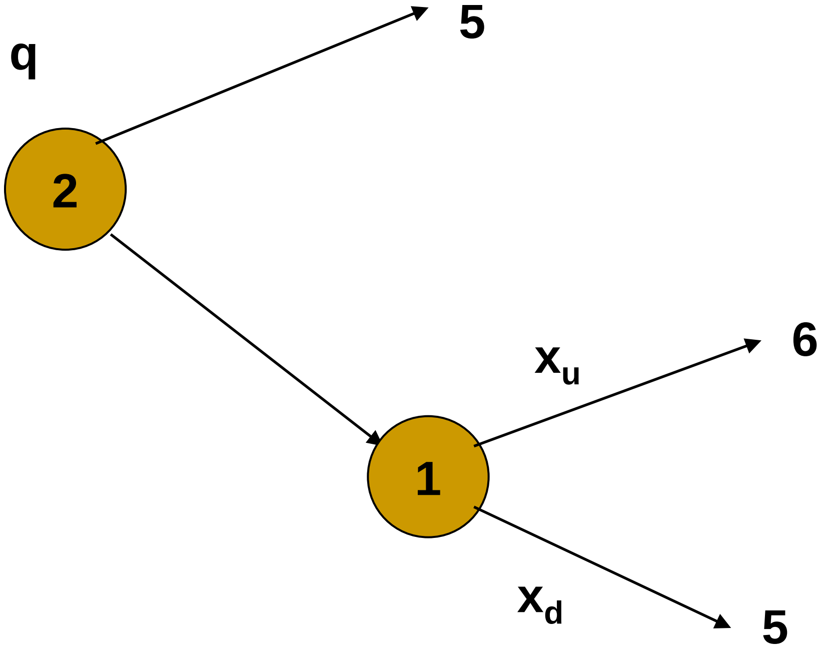
$q \leq 5$

$q \leq 6\, x_u + 5\, x_d$

Slack!

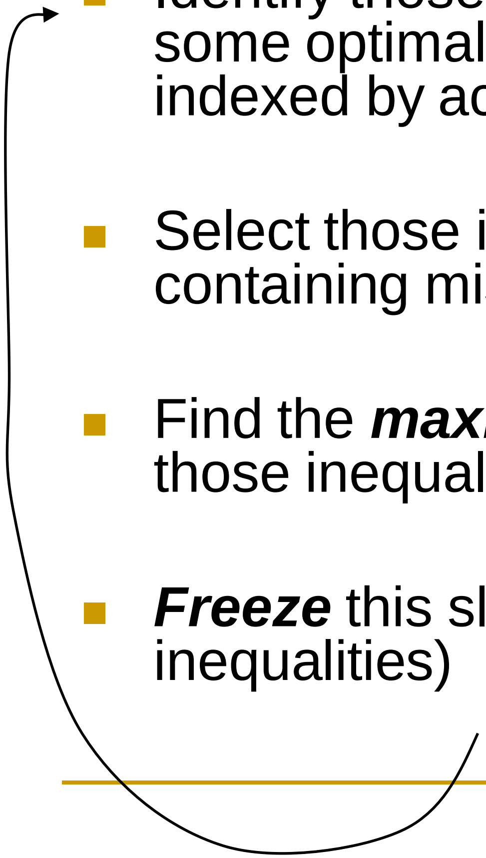
$x_u + x_d = 1$

$x_u, x_d \geq 0$



Intuition: Left hand side of inequality in solution is what Player 2 *could* achieve, right hand side is what he actually achieves by taking the action, so slack is good!

The algorithm



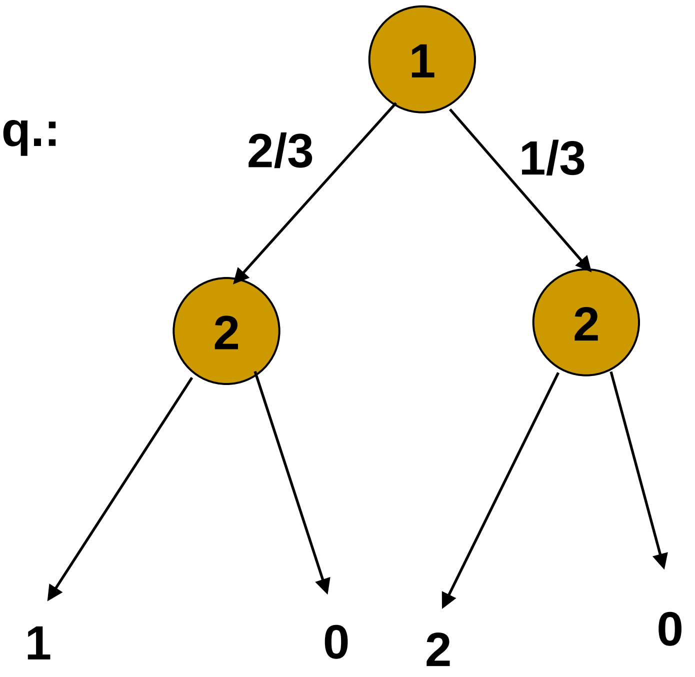
- Solve original KMvS program.
- Identify those inequalities that may be satisfied with slack in some optimal solution. **Intuition**: These are the inequalities indexed by action sequences containing ***mistakes***.
- Select those inequalities corresponding to action sequences containing mistakes but having no prefix containing mistakes.
- Find the ***maximin*** (min over the inequalities) possible slack in those inequalities.
- ***Freeze*** this slack in those inequalities (strengthening the inequalities)

Proof of correctness

- Similar to proof of correctness of “Dresher’s procedure” – characterizing the proper equilibria of a matrix game.
- Step 1: Show that any proper equilibrium “survives” the iteration.
- Step 2: Show that all strategies that survive are “payoff”-equivalent.

Left or right?

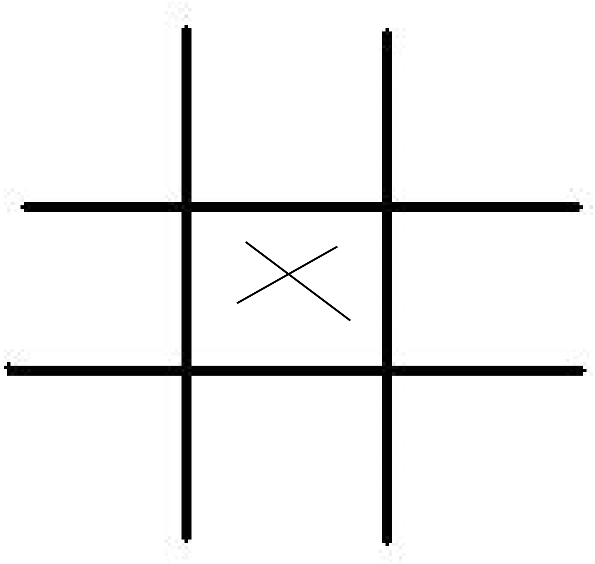
Unique proper eq.:



Interpretation

- If Player 2 never makes mistakes the choice is arbitrary.
- We should imagine that Player 2 makes mistakes with some small probability but can ***train*** to avoid mistakes in either the left or the right node.
- In equilibrium, Player 2 trains to avoid mistakes in the “expensive” node with probability 2/3.
- Similar to “meta-strategies” for selecting chess openings.
- The perfect information case is easier and can be solved in linear time by “a backward induction procedure” without linear programming.
- This procedure assigns three values to each node in the tree, the “real” value, an “optimistic” value and a “pessimistic” value.

The unique proper way to play tic-tac-toe



…. with probabiltiy 1/13

Questions about computing proper equilibria

- Can a proper equilibrium of a ***general-sum*** bimatrix game be found by a “pivoting algorithm”? Is it in the complexity class **PPAD**? Can one convincingly argue that this is ***not*** the case?

- Can an ϵ-proper strategy profile (as a system of polynomials in ϵ) for a matrix game be found in polynomial time). **Motivation:** This captures a “lexicographic belief structure” supporting the corresponding proper equilibrium.

Plan

- ***Representing*** finite-duration, imperfect information, two-player zero-sum games and ***computing*** minimax strategies. ✓
- Issues with minimax strategies. ✓
- Equilibrium refinements (a crash course) and how refinements resolve the issues, and how to modify the algorithms to compute refinements. ✓
- (If time) Beyond the two-player, zero-sum case.

Finding Nash equilibria of general sum games in normal form

- Daskalakis, Goldberg and Papadimitriou, 2005. Finding an approximate Nash equilibrium in a 4-player game is **PPAD**-complete.

- Chen and Deng, 2005. Finding an exact or approximate Nash equilibrium in a 2-player game is **PPAD**-complete.

- … this means that these tasks are polynomial time equivalent to each other and to finding an approximate Brower fixed point of a given continuous map.

- This is considered evidence that the tasks cannot be performed in worst case polynomial time.

- .. On the other hand, the tasks are not likely to be NP-hard.: If they are NP-hard, then **NP**=**coNP**.

Motivation and Interpretation

- ***The computational lens***
- ”If your laptop can’t find it neither can the market” (Kamal Jain)

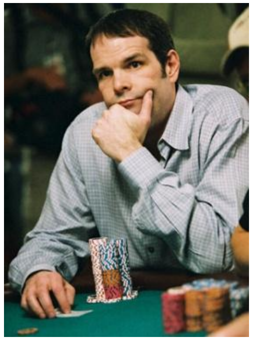
What is the situation for equilibrium refinements?

- ***Finding*** a refined equilibrium is at least as hard as finding a Nash equilibrium.

M., 2008: ***Verifying*** if a given equilibrium of a 3-player game in normal form is trembling hand perfect is NP-hard.

Two-player zero-sum games

Player 1:
Gus, the Maximizer



Player 2 :
Howard, the Minimizer

Maxmin value (lower value, security value):

$$\underline{v} = \max_{x \in \Delta(S_1)} \min_{y \in \Delta(S_2)} u_1(x, y)$$

Minmax value (upper value, threat value):

$$\overline{v} = \min_{y \in \Delta(S_2)} \max_{x \in \Delta(S_1)} u_1(x, y)$$

von Neuman's minmax theorem (LP duality): $\underline{v} = \overline{v}$

Three-player zero-sum games

Player 1: Gus, the Maximizer

Players 2 and 3: Alice and Bob, the Minimizers

"honest-but-married"

Maxmin value (lower value, security value):

$$\underline{v} = \max_{x \in \Delta(S_1)} \min_{(y,z) \in \Delta(S_2) \times \Delta(S_3)} u_1(x, y, z)$$

Minmax value (upper value, threat value):

$$\bar{v} = \min_{(y,z) \in \Delta(S_2) \times \Delta(S_3)} \max_{x \in \Delta(S_1)} u_1(x, y, z)$$

↑

***Uncorrelated* mixed strategies.**

Three-player zero-sum games

Player 1:
Gus, the
Maximizer

Players 2 and 3:
Alice and Bob, the
Minimizers

“honest-but-married”

Maxmin value (lower value, security value):

$$\underline{v} = \max_{x \in \Delta(S_1)} \min_{(y,z) \in \Delta(S_2) \times \Delta(S_3)} u_1(x, y, z)$$

Minmax value (upper value, threat value):

$$\bar{v} = \min_{(y,z) \in \Delta(S_2) \times \Delta(S_3)} \max_{x \in \Delta(S_1)} u_1(x, y, z)$$

Bad news:

- **Lower value $\leq$ upper value but in general not =**
- **Maxmin/Minmax not necessarily Nash**
- **Minmax value may be irrational**

equality?

Computable in P, given table of u_1

Maxmin value (lower value, security value):

$$
\begin{aligned}
\underline{v} &= \max_{x\in\Delta(S_1)} \min_{(y,z)\in\Delta(S_2)\times\Delta(S_3)} u_1(x,y,z) \\
&= \max_{x\in\Delta(S_1)} \min_{(y,z)\in S_2\times S_3} u_1(x,y,z) \\
&= \min_{(y,z)\in\Delta(S_2\times S_3)} \max_{x\in\Delta(S_1)} u_1(x,y,z)
\end{aligned}
$$

Correlated **mixed strategy (married-and-dishonest!)**

Minmax value (upper value, threat value):

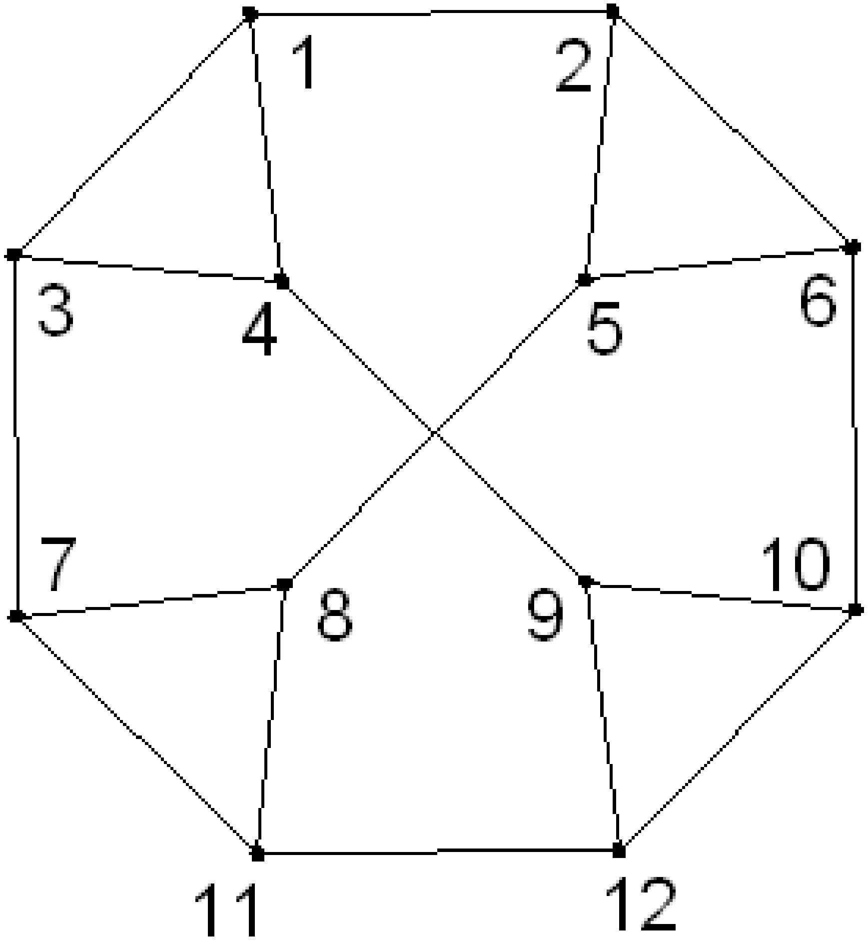
$$\bar{v} = \min_{(y,z)\in\Delta(S_2)\times\Delta(S_3)} \max_{x\in\Delta(S_1)} u_1(x,y,z)$$

Borgs *et al*., STOC 2008:
NP-hard to approximate, given table of u_1!

Borgs et al., STOC 2008

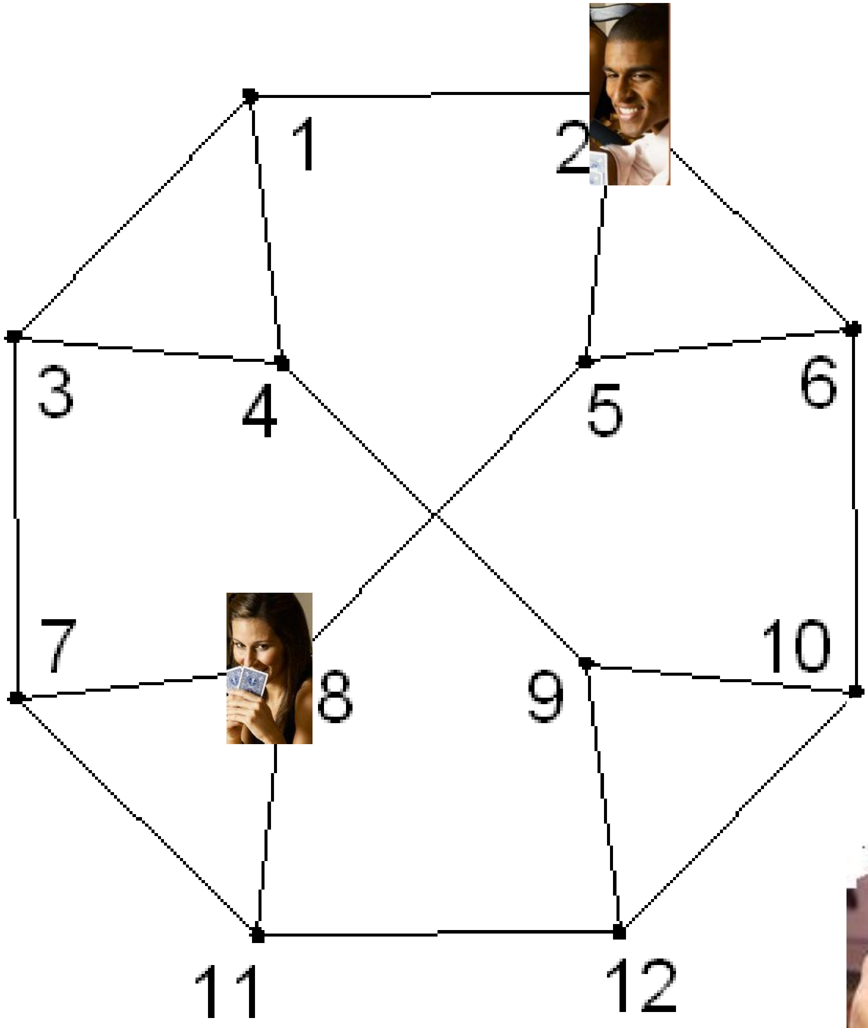
It is **NP**-hard to approximate the minmax-value of a 3-player n x n x n game with payoffs 0,1 (win,lose) within additive error $3/n^2$.

Proof – Hide and seek game



Alice and Bob hide in an undirected graph.

Proof – Hide and seek game



Alice and Bob hide in an undirected graph.

Gus, blindfolded, has to call the location of one of them.

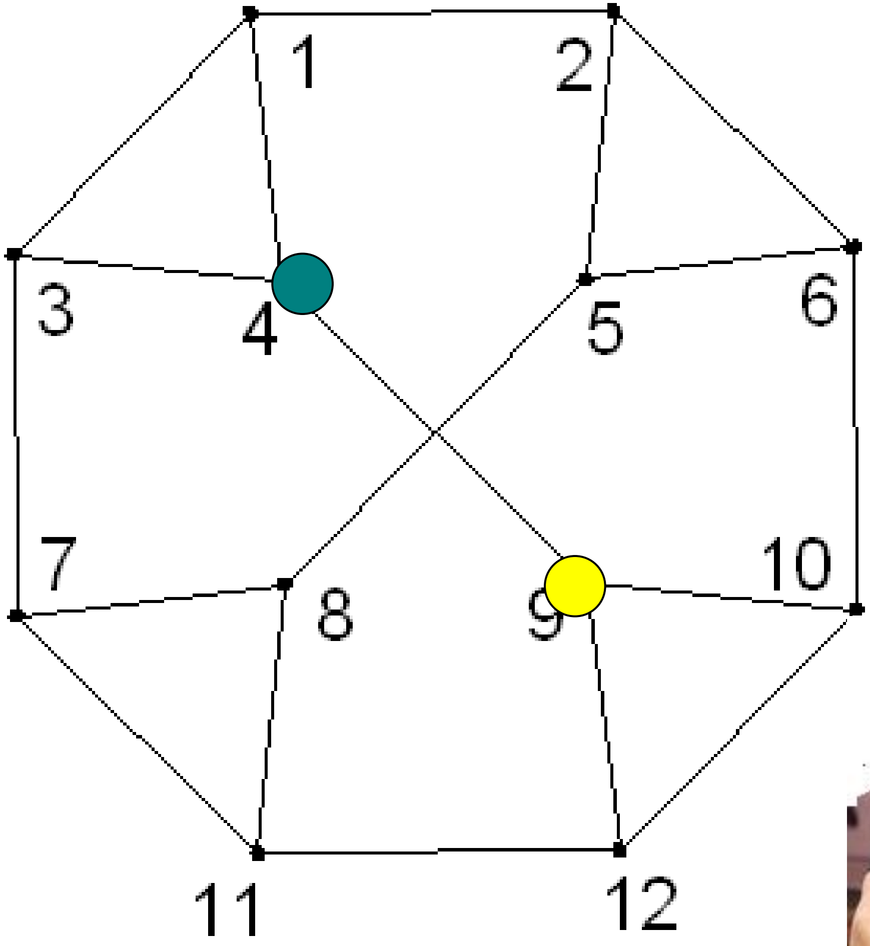
Analysis

- Optimal strategy for Gus
  - Call arbitrary player at random vertex.

- Optimal strategy for Alice and Bob
  - Hide at random vertex

- Lower value = upper value = 1/*n*.

Hide and seek game with colors

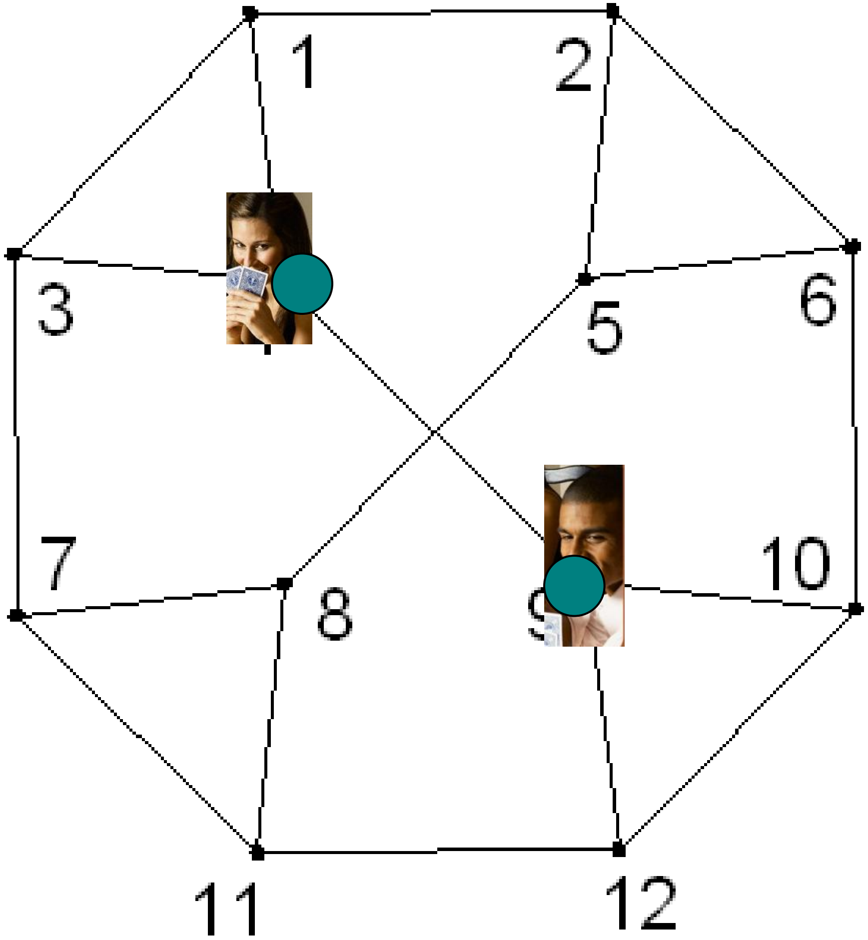


Alice and Bob hide in an undirected graph.

.. and declare a color in ● ● ●

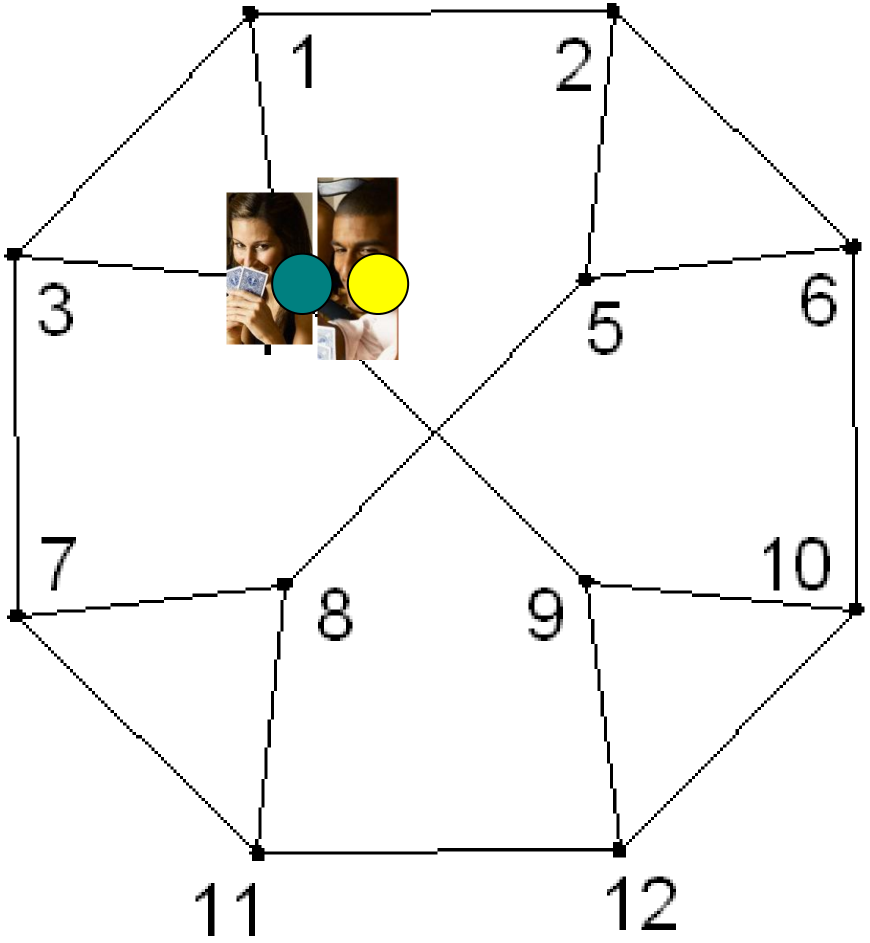
Gus, blindfolded, has to call the location of one of them.

Hide and seek game with colors



Additional way in which Gus may win: Alice and Bob makes a declaration inconsistent with 3-coloring.

Hide and seek game with colors



Additional way in which Gus may win: Alice and Bob makes a declaration inconsistent with 3-coloring.

Analysis

- If graph is 3-colorable, minmax value is $1/n$: Alice and Bob can play as before.
- If graph is not 3-colorable, minmax value is at least $1/n + 1/(3n^2)$.

Reduction to deciding trembling hand perfection

- Given a 3-player game G, consider the task of determining if the min-max of Player 1 value is bigger than α+ϵ or smaller than α-ϵ.

- Define G* by augmenting the strategy space of each player with a new strategy *.

- Payoffs: Players 2 and 3 get 0, no matter what is played.

- Player 1 gets α if at least one player plays *, otherwise he gets what he gets in G.

- **Claim:** (*,*,*) is trembling hand perfect in G* *if and only if* the minmax value of G is smaller than α - ϵ.

Intuition

- If the minmax value is less than α - ϵ, he may believe that in the equilibrium (*,*,*) Players 2 and 3 may tremble and play the exactly the minmax strategy. Hence the equilibrium is trembling hand perfect.

- If the minmax value is greater than α+ϵ, there is no ***single*** theory about how Players 2 and 3 may tremble that Player 1 could not react to and achieve something better than α by not playing *. This makes (*,*,*) imperfect,

- Still, it seems that it is a ”reasonable” equilibrium if Player 1 does not happen to have a fixed belief about what will happen if Players 2 and 3 tremble(?)…..

Questions about NP-hardness of the general-sum case

- Is deciding trembling hand perfection of a 3-player game *in* NP?
- Deciding if an equilibrium in a 3-player game is *proper* is NP-hard (same reduction). Can properness of an equilibrium of a 2-player game be decided in **P**? In **NP**?

- Thank you!