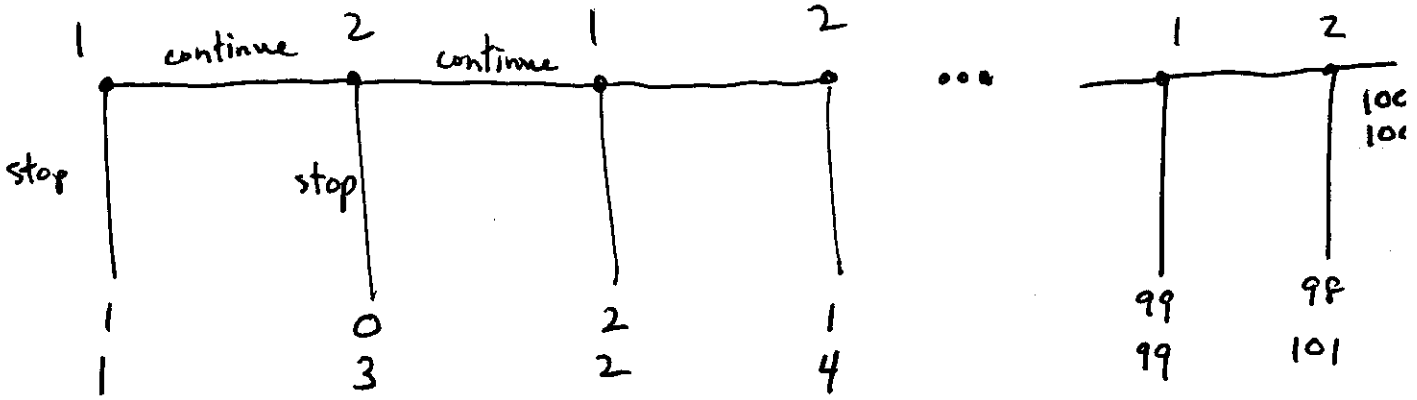
Critique of SPNE and backwards induction.
It requires common knowledge of rationality.



the Centipede game

- Would your opponent be so likely to use the "backwards induction" logic that it does not pay to continue?

- Choosing "continue" is inconsistent with the theory underlying backwards induction. Will your opponent remain convinced that you will stop in the future?

- A small probability of being a crazy guy who continues changes the way rational people would play.

Multi-stage Games (with observed actions)

actions chosen in stage k:

$$a^k = (a_1^k, \ldots, a_i^k, \ldots a_I^k)$$

history at the end of stage k:

$$h^{k+1} = (a^0, \ldots, a^k)$$

set of feasible actions after history h^{k+1}:

$$A_i(h^{k+1})$$

A **pure strategy** is a contingent plan:

$$s_i : \underset{\text{set of possible histories}}{H^{k+1}} \to \underset{\text{set of feasible actions across all histories}}{A_i(H^{k+1})}$$

$$h^{k+1} \xrightarrow{s_i} A_i(h^{k+1})$$

payoffs depend on the entire history (usually, average payoff or discounted sum)

SPNE — the game after history h^K is a subgame. Payoffs and "continuation strategies" are defined in the obvious way.

example: repeated games

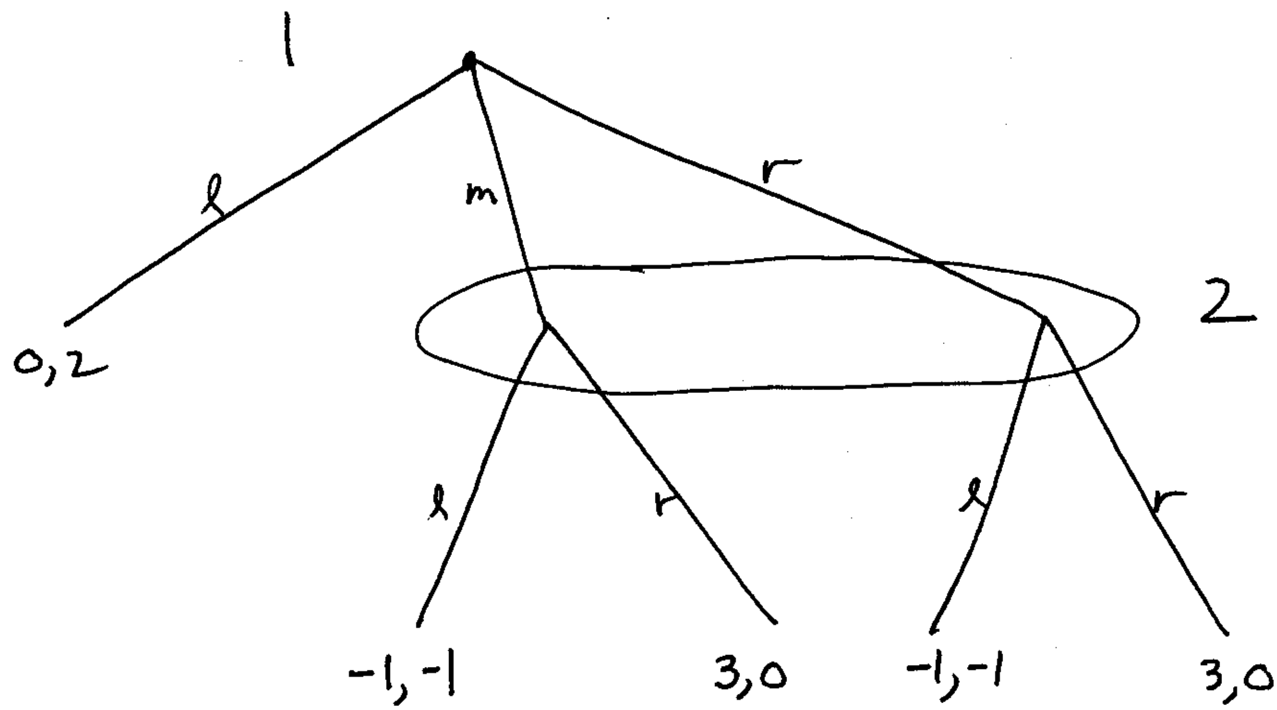
	c	d
c	3,3	0,4
d	4,0	1,1

- grim trigger strategy
- other payoffs are possible

Folk Thm Any payoff that is feasible and individually rational in the stage game is a NE payoff of the repeated game.

Refinements of SPNE

Subgame perfection is not enough to rule out all threats that are not credible.



This game has no smaller subgames, so (l,l) is a SPNE. However, player 2's threat to play l is not credible.

- m and r are essentially the same strategy. (The distinction could be based on which pencil 1 uses to write his strategy choice.) But combining m and r changes the set of SPNE.

Sequential rationality (starting at a particular information set) depends on <u>beliefs</u> about which one is the correct node.

<u>Def.</u> A system of beliefs, μ, is a specification of a (nonnegative) probability for each decision node, x, such that

$$\sum_{x \in H} \mu(x) = 1$$

holds for all information sets, H.

<u>Def.</u> A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_I)$ is <u>sequentially</u> <u>rational</u> at information set H given a system of beliefs μ if for the player, i, who moves at H

$$E[u_i \mid H, \mu, \sigma_i, \sigma_{-i}] \geq E[u_i \mid H, \mu, \tilde{\sigma}_i, \sigma_{-i}]$$

for all continuation strategies $\tilde{\sigma}_i$. If σ satisfies this condition at all information sets, then σ is sequentially rational given belief system μ.

The problem with "bad" equilibria is that there may be no sensible beliefs a player can have to justify the action taken.

Just as sequential rationality depends on beliefs, beliefs should depend on the (rational) strategies.

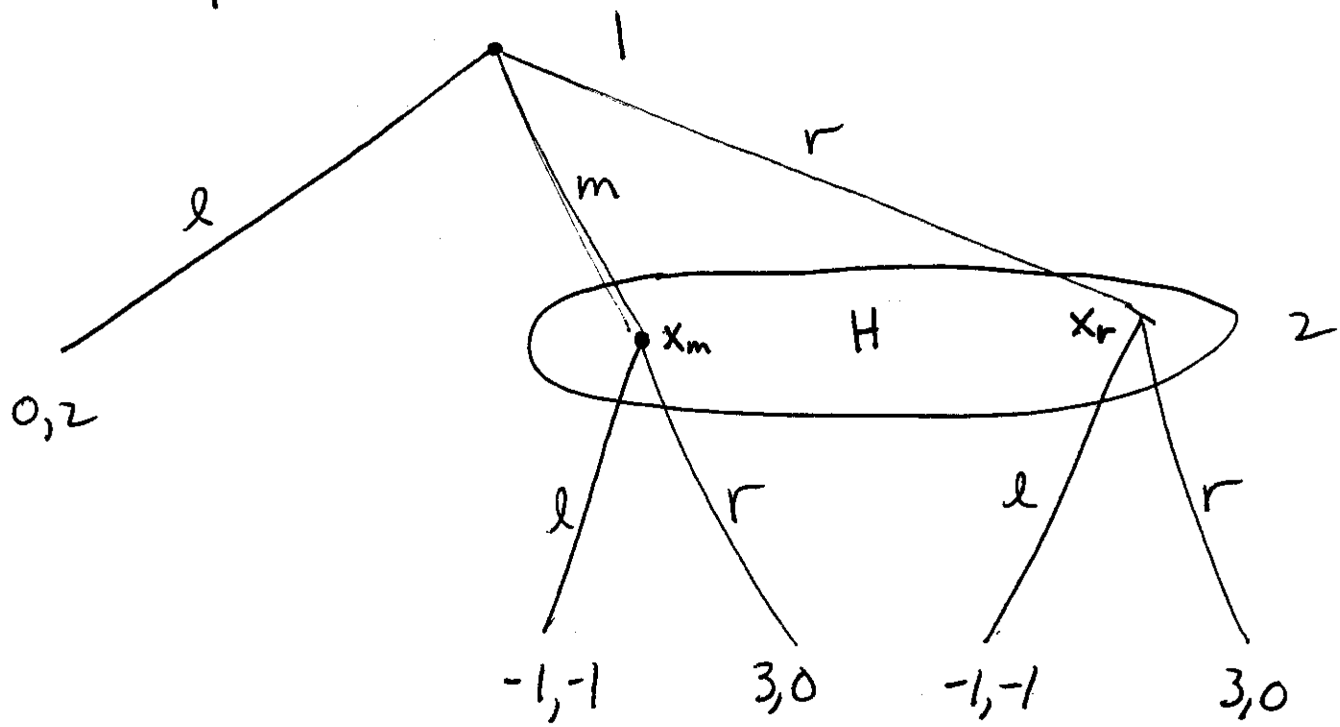
<u>Def.</u> A profile of strategies and a system of beliefs (σ, μ) is a <u>weak</u> <u>Perfect</u> <u>Bayesian</u> <u>equilibrium</u> if

(i) σ is sequentially rational, given belief system μ.

(ii) μ is derived from σ according to Bayes' rule whenever possible.
For any information set "on the equilibrium path" [where $\text{prob}(H|\sigma) > 0$]

$$\mu(x) = \frac{\text{Prob.}(x|\sigma)}{\text{Prob}(H|\sigma)} \qquad \forall x \in H.$$

example



- (l, l) is a SPNE but not a WPBE.
- No matter what beliefs are at player 2's information set, utility is higher by playing r
- $\{(m,r), \ \mu(x_m)=1, \ \mu(x_r)=0\}$ is a WPBE

If player 1 plays m, $\mu(x_m) = \frac{prob(x_m|\sigma)}{prob(H|\sigma)} = 1$

Example: A game of imperfect information.

This game could have been derived from a game of perfect but incomplete information.