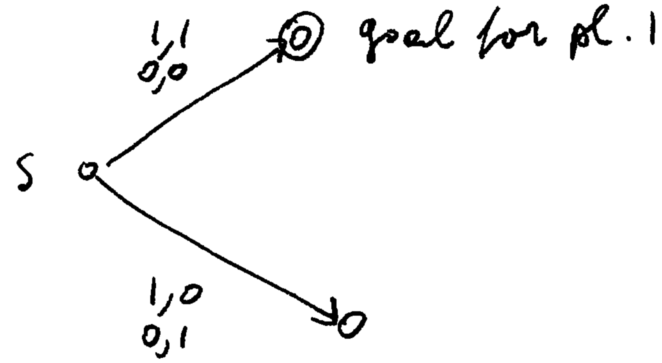
Concurrent Games

- Both players have moves at a state;
- The choice of moves is simultaneous, and independent.



How do we phrase " player 1 can win" ?

Attempt 1

$$\exists a \in \Gamma_1(s).\ \forall b \in \Gamma_2(s).\ \tau(s,a,b) \in T$$

No: too difficult for pl. 1. He can never win!

Attempt 2 (invert the quantifiers)

$$\forall b \in \Gamma_2(s).\ \exists a \in \Gamma_1(s).\ \tau(s,a,b) \in T$$

No: too easy for player 1: he can win all the time.

Essentially, the quantification that comes later can "see" the value of the one that comes first. This is similar to the reason for the following inequality, for $f : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$:

$$\inf_x \sup_y f(x,y) \geq \sup_y \inf_x f(x,y)$$

We can choose y knowing x, so we are better able to maximize f, compared to having to choose x without knowing y.

So for games, we need a "simultaneous" quantifier, something like:

$$\begin{pmatrix} \exists a \in \Gamma_1(s) \\ \forall b \in \Gamma_2(s) \end{pmatrix} . \ \tau(s,a,b) \in \Gamma.$$

How to do it? Randomization!

Mixed Moves

A mixed move is simply a probability distribution over normal ("pure") moves.

$\xi_1 \in \mathrm{Distr}(\Gamma_1(s))$ mixed move of pl. 1 at s.

Let: $D_1(s) = \mathrm{Distr}(\Gamma_1(s))$
$D_2(s) = \mathrm{Distr}(\Gamma_2(s))$.

Valuations

Given a state space S, a valuation is a mapping

$$f : S \mapsto [0,1]$$

that assigns a value to each state.

We let

$$\mathcal{V} = \{ S \mapsto [0,1] \}$$

be the set of all valuations.

Quantitative Predecessor

QPre is a valuation transformer: given a valuation, it computes a new valuation. For $X \in \mathcal{V}$, we let, for $s \in S$:

$$QPre(X)(s) = \sup_{\xi_1 \in D_1(s)} \inf_{\xi_2 \in D_2(s)} E_s^{\xi_1,\xi_2}(X),$$

where:

$$E_s^{\xi_1,\xi_2}(X) = \sum_{a_1 \in \Gamma_1(s)} \sum_{a_2 \in \Gamma_2(s)} \sum_{t \in S} \xi_1(a_1) \cdot \xi_2(a_2) \cdot \delta(s,a_1,a_2)(t) \cdot X(t).$$

Note that this is an instance of a classical matrix game.

Precisely, consider for $a_1 \in \Gamma_1(s)$, $a_2 \in \Gamma_2(s)$:

$$v_{a_1,a_2} = \sum_{t \in S} \delta(s,a_1,a_2)(t) \cdot X(t).$$

Then, we can write:

$$QPre(X)(s) = val\,[v_{a_1 a_2}]$$

where, for a matrix indexed by sets A, B (<u>finite</u>!) (that is, $[v_{ab}]_{a\in A,\, b\in B}$),

$$val\,[v] = \sup_{\xi_1 \in Distr(A)} \; \inf_{\xi_2 \in Distr(B)} E^{\xi_1,\xi_2}(v).$$

Now, matrix games are a classical notion: one player chooses the row, the other the column, and the matrix entry gives the payoff; player 1 is the maximizer, player 2 the minimizer. Von Neumann in 1929 proved the <u>minimax</u> theorem:

$$\sup_{\xi_1 \in Distr(A)} \; \inf_{\xi_2 \in Distr(B)} E^{\xi_1,\xi_2}(v) =$$

$$= \inf_{\xi_2 \in Distr(B)} \; \sup_{\xi_1 \in Distr(A)} E^{\xi_1,\xi_2}(v).$$

Note that, as usual,

$$E^{\xi_1,\xi_2}(v) = \sum_{a\in A}\sum_{b\in B} \xi_1(a)\cdot\xi_2(b)\cdot v_{a,b}$$

It is also not hard to show that, for finite sets A, B, (omitting the ranges of ξ_1, ξ_2, as above):

$$\sup_{\xi_1}\inf_{\xi_2} E^{\xi_1,\xi_2}(v) = \max_{\xi_1}\min_{b\in B} E^{\xi_1,b}(v).$$

Intuitively, once the minimizer knows the distribution ξ_1 chosen by the maximizer, randomization is no longer needed. So in summary:

$$\begin{array}{ccc} \sup\limits_{\xi_1}\inf\limits_{\xi_2} E^{\xi_1,\xi_2}(v) & \overset{\text{von Neumann}}{=} & \inf\limits_{\xi_2}\sup\limits_{\xi_1} E^{\xi_1,\xi_2}(v) \\ \| & & \nparallel \\ \max\limits_{\xi_1}\inf\limits_{b_2\in B} E^{\xi_1,\xi_2}(v) & \neq & \min\limits_{\substack{\cancel{\xi_2} \\ b_2\in B}}\max\limits_{\xi_1} E^{\xi_1,\xi_2}(v). \end{array}$$

Reachability Games (concurrent)

Aside from QPre, we define also these valuation transformers: for $f, g \in \mathcal{V}$:

$\sqcup$ (pointwise max) $$(f \sqcup g)(s) = f(s) \sqcup g(s) = \max\{f(s), g(s)\}.$$

$\sqcap$ $$(f \sqcap g)(s) = f(s) \sqcap g(s) = \min\{f(s), g(s)\}.$$

Note also that, given a set $R \subseteq S$, we define its characteristic function $[R]$ by, for all $s \in S$:

$$[R](s) = \begin{cases} 1 & s \in R \\ 0 & s \notin R. \end{cases}$$ so, $[R] \in \mathcal{V}$.

Given $f, g \in \mathcal{V}$, we also let $f \leq g$ iff $f(s) \leq g(s)$ for all $s \in S$.

Then, it is easy to see that $(\mathcal{V}, \leq)$ is a complete lattice.

<u>Reachability algorithm:</u>

$$\langle 1 \rangle \Diamond R = \mu X. \left([R] \sqcup QPre(X) \right)$$

Computing the fixpoint iteratively:

$$X_0 = \bar{0} = \lambda s.0$$ (the identically 0 valuation)

$$X_1 = [R] \sqcup \underbrace{QPre(\bar{0})}_{= \bar{0}}$$

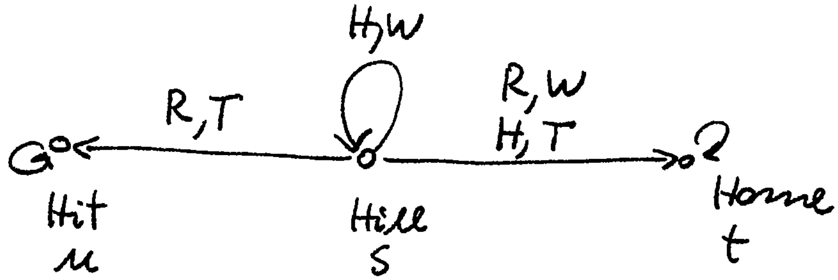
$$X_2 = [R] \sqcup QPre(X_1)$$

$$\dots$$

$$X_* = \mu X. \left([R] \sqcup QPre(X) \right)$$

Idea: prove by induction that $X_k(s)$ is a lower bound for the probability of reaching R in at most $k-1$ steps.

An advanced example: the hill-snowball game.



Player 1 does not have a strategy π_1 to get Home with probability 1, or:

$$\forall \pi_1 \; \exists \pi_2 . \; Pr_s^{\pi_1 \pi_2} (\Diamond \{t\}) < 1 .$$

In fact, there are two cases:

- If Pl. 1 uses π_1 that never runs before seeing a snowball, pl. 2 can just wait forever. Formally, if for all k, $\pi_1(s^k)(R) = 0$,

  then take $\pi_2(s^k)(W) = 1$ for all k.

- If $\overset{\text{smallest}}{\exists} k . \; \pi_1(s^k)(R) > 0$, then choose

  π_2 such that: $\forall j < k : \pi_2(s^j)(R) = 0$

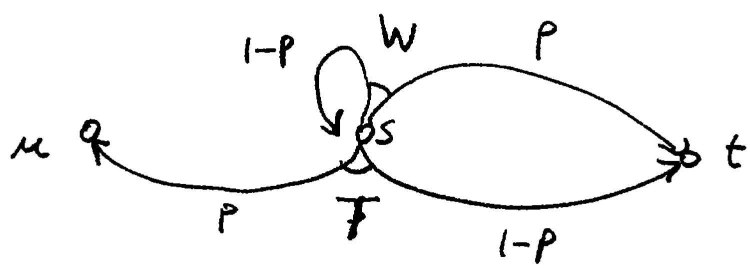
  $\pi_2(s^k)(R) = 1$

where " $\overset{\text{smallest}}{\exists} k$ " means to take the smallest such k.

However, if Pl. 1 plays:

$$\xi_1: \quad \begin{array}{lll} R & \text{prob.} & p \\ H & \text{prob} & 1-p \end{array}$$

at every round, then the situation for player 2 is as follows:



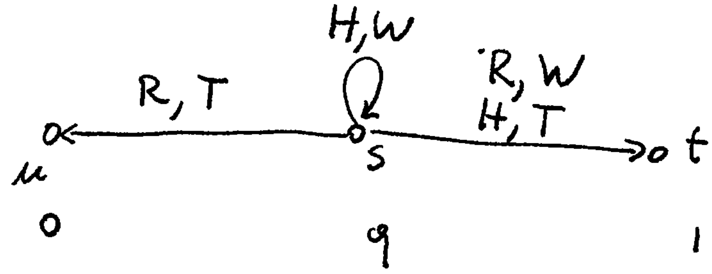
So the best pl. 2 can do is throw immediately (there is no gain in waiting), and pl. 1 reaches t (Home) with probability at least $1-p$.

As this is true for all p,

$$(\langle 1\rangle \diamond \{t\})(s) = \sup_{\pi_1} \inf_{\pi_2} \Pr_s^{\pi_1 \pi_2} (\diamond \{t\}) = 1.$$

However, there is no optimal strategy for player 1: the value 1 is never attained.

Now, let us run the reachability algorithm on the hill-and-snowball game:



The valuation $X_k(t)$ is 1 for all $k > 0$, and $X_k(u)$ for all k.

We need to compute QPre (X_k) given X_k.

Call $q = X_k(s)$, to simplify the notation.

Consider a mixed move ξ_1 of player 1, that plays:

R with prob. α

H " " $1-\alpha$.

In response, player 2 can play any combination of W, T, but we know that a pure answer W, T will be just as good.

Assuming pl. 1 plays $\alpha R + (1-\alpha) H$ for some α, what is a ~~goo~~ good (minimizer) response for player 2?

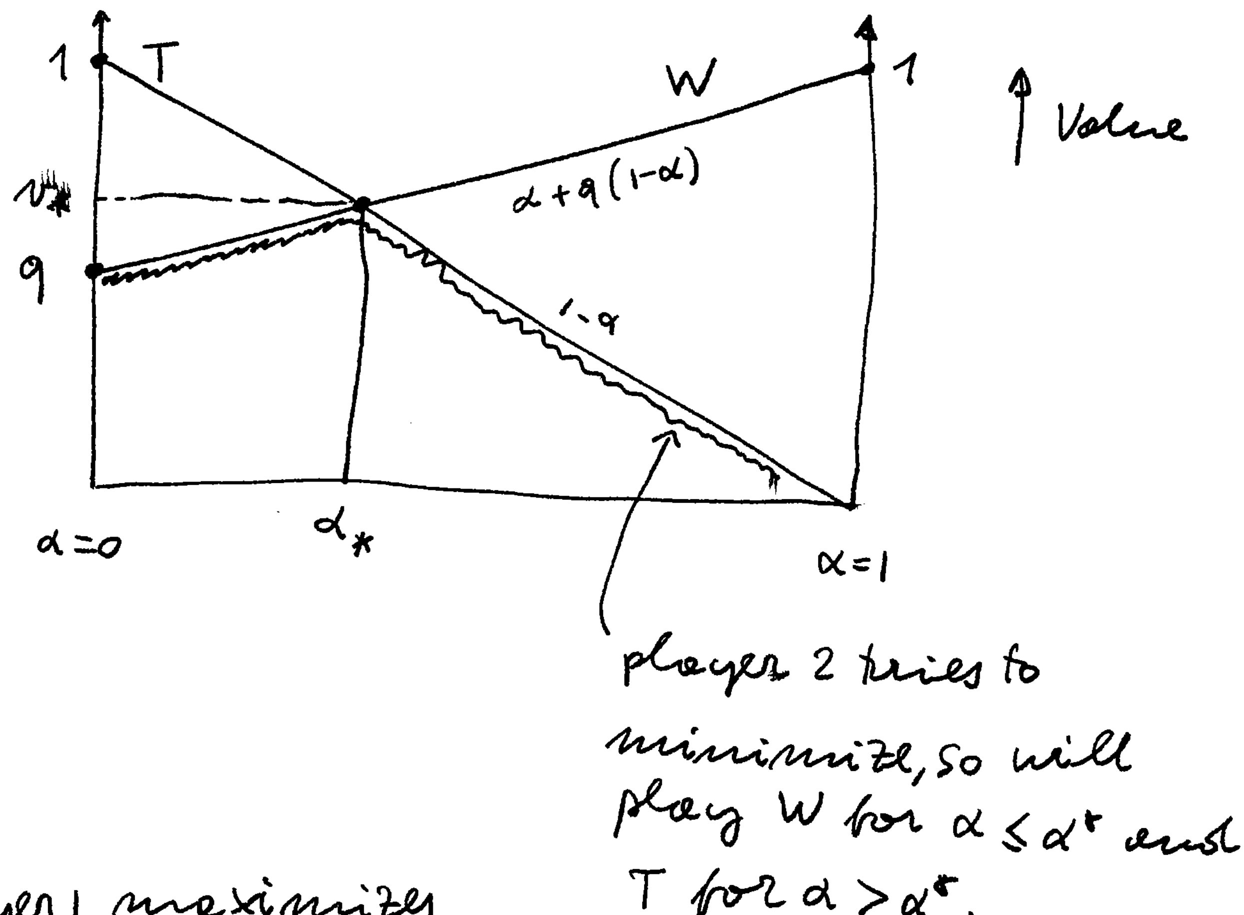
If Pl. 2 plays:

			Value		Prob		
T	R	$\rightsquigarrow$	0	$\cdot$	α	$=$	0
	H	$\rightsquigarrow$	1	$\cdot$	$1-\alpha$	$=$	$1-\alpha$
						Total:	$1-\alpha$
W	R	$\rightsquigarrow$	1	$\cdot$	α	$=$	α
	H	$\rightsquigarrow$	q	$\cdot$	$(1-\alpha)$	$=$	$q(1-\alpha)$
						Total:	$\alpha + q(1-\alpha)$

So:

			Value
Pl. 2 plays	T	$\rightsquigarrow$	$1-\alpha$
	W	$\rightsquigarrow$	$\alpha + q(1-\alpha)$

The situation can be summarized as follows:



Player 1 maximizes, so he will play α_*.

What is α_*? Solve:

$$1-\alpha = \alpha + q(1-\alpha)$$

$$\alpha_* = \frac{1-q}{2-q}$$

So, $v_* = 1-\alpha_* = \alpha_* + q(1-\alpha_*) = \dfrac{1}{2-q}$.

(simple calculations).

This tells us that, if $X(s) = q$, then

$$QPre(X)(s) = \frac{1}{2-q}.$$

Let's look at how the algorithm proceeds:

$X_0 =$	0	0	0
$X_1 =$	0	0	1
$X_2 =$	0	$\frac{1}{2-0} = \frac{1}{2}$	1
$X_3 =$	0	$\frac{1}{2-\frac{1}{2}} = \frac{2}{3}$	1
$X_4 =$	0	$\frac{1}{2-2/3} = \frac{3}{4}$	1

In general, $X_k(s) = \frac{k-1}{k}$.

$$\lim_{k \to \infty} X_k(s) = 1.$$

How to construct an optimal strategy from the proof:

at round k, we have (for $k > 1$)

$$X_{k-1}(s) = \frac{k-2}{k-1} \quad . \qquad (*)$$

Pl. 1 plays the α_* obtained by setting $q = (*)$, or:

$$\alpha_k = \frac{1 - \frac{k-2}{k-1}}{2 - \frac{k-2}{k-1}} \quad .$$

This gives a strategy that depends on the round k. We know that there is also a memoryless strategy (play R with prob. p at all rounds, for sufficiently small p).

a memoryful strategy that wins with probability as close to optimal as we wish can be built as follows.

Given $\varepsilon > 0$, choose $n > 0$ so that $|X_* - X_n| < \varepsilon$. Then, I play:

~~[illegible]~~

$$\xi_1^{n} \quad \xi_1^{n-1} \quad \xi_1^{n-2} \quad \cdots\cdots \quad \xi_1^{0}$$

where ξ_1^{k} is the ~~selector~~ (mixed move) for pl. 1 at s that chooses $\alpha_k R + (1-\alpha_k) H$.

This memoryful strategy is optimal if there is a bound on the n. of rounds in the game, of $n-1$ rounds.