Wikipedia: Prisoner's Dilemma

generalized version:

A \ B	coop	defect
coop	R, R	S, T
defect	T, S	P, P

relationships: $T > R > P > S$.

for iterated games, we require $2R > T + S$ so that cooperation is better than alternating.

iterated prisoner's dilemma also known as: Peace-War game

so yes, OK, my point about how scoring depends on the other submitted algorithms makes sense.

what if we only allow finite automaton players?

the tournament format also matters a great deal

— if single elim, obv a "don't lose" strat like always-defect is dominant

maybe something can be said about what happens if we're given a distribution.

why does the evolutionary approach make sense?

↳ because it captures how meta game usually develops going back and forth also, if no progress is made in the evolutionary game then things would be perfect

→ I DO EXPECT, however, that no such situation actually exists.

POSSIBLY INTERESTING QUESTION:

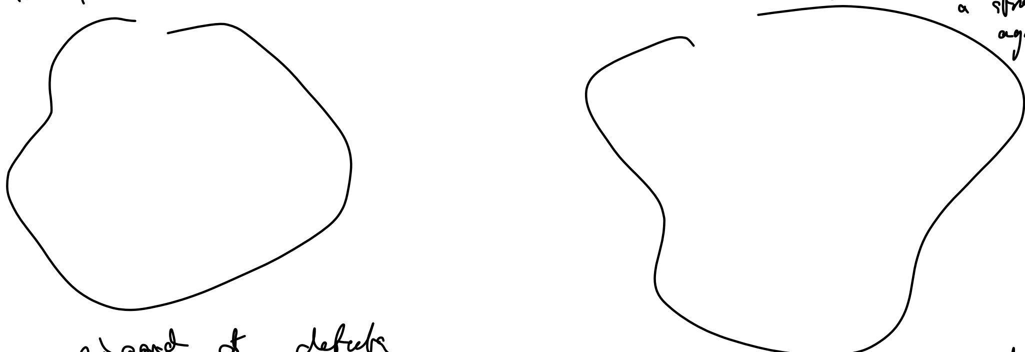
Does there exist a collection of strategies for which no outside strategy exists that would outperform them?

N rounds (maximize)

(hmm this does not really capture what I want. because a strategy could identify itself and tit for tat and do well against it

hmmm. I kinda like the way this is set up.

wait. can be solved by saying that we just play a strategy against itself.



cloud of defects

cloud of cooperates

$N \cdot P$

$N \cdot T$

$N \cdot S$

$N \cdot R$

$P + T \;\; ? \;\; S + R$

win against tit for tat? only over 1 defect-cooperation point.

Axelrod: paper 1

"what is effective is likely to depend not only upon the characteristics of a particular strategy but also upon the nature of other strategies with which it must interact."

what seems like a reasonable evolutionary way to look at this?

→ start with a bunch of strategies

→ in every step, add a strategy that would win and remove the loser

questions: 1: would this procedure produce an equilibrium?

2: would this equilibrium always be the same?

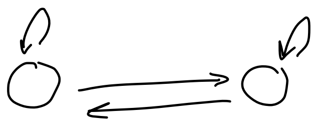
hmm

OK so Sipser's idea is to restrict to finite automata.

---

hmm.

what is my idea?



play against each other + against itself.

why? if only against each other there is no reason to cooperate!

---

let's say we only can remember last move.

so cc cd dc dd

move:

2^4 different strategies.

let's just say you only can base it on opponent (start c)

opponent's last move:	c	d
your move:		

cc
cd ← titfortat
dc
dd

cd:

cd ↔ cc tie

cd ↔ dc

yeah so tit-for-tat very clearly optimal!

R S T

P

(finite case)

cc	R
cd	S
dc	T
dd	P

so $\overset{5}{T} > \overset{3}{R} > \overset{1}{P} > \overset{0}{S}$

and $R > \frac{S+T}{2}$.

<u>claim</u>: TFT loses w/ at most T−S pts against any other strategy.

<u>proof</u>:

TFT:	c	d	c	
other:	d	c	d	

cancel out

<u>claim</u>: for deterministic strategies, if P_1 is played against its clone P_2, then P_1 and P_2 get the same result.

<u>proof</u>: they'll just do the same thing

<u>corollary</u>: self-playing games are just cc's and dd's.

<u>corollary</u>: in order not to loose against TFT, a strategy needs to play only cc's against itself (and possibly a few dd's)

———— d

4 rounds.
let's do ccc d.

then: against TFT: +(T − S)
against itself: − R + P
so net: T + P − R − S > 0

so that strategy is better than TFT,
in this model.

oh but it is worse than other strategies.

so TFT > DD > LASTD > TFT

could say: we have finite automata.
then eliminate LASTD
could also say: don't know when we
will end

det A wins against B if ∃ N s.t.
for all n ≥ N A wins over B in an
n-length game.

hmm what about a strategy that:

d first round. if opponent also d's then always c. if opponent c's always d.

simple finite automaton

so I guess this approach doesn't work either.

and in all scenarios I think there will exist degenerate strats like these that win against TFT.

maybe then the best option is to play everyone against everyone?

why is O a worse strategy than TFT?

because it loses to DD?

hmm ok let's try everyone-against-everyone.

what does that mean?

I'm computing the average score against all possible strategies — a term I will have to define.

yea ok, this actually makes some sense.

$$\begin{array}{c|cccc} & cc & cd & dc & dd \\ cc & & & & \\ cd & & & & \\ dc & & & & \\ dd & & & & \end{array} \quad \begin{pmatrix} p_1 & p_2 & p_3 & p_4 \\ \vdots & & & \vdots \\ \vdots & & & \vdots \\ \vdots & & & \vdots \end{pmatrix} := A$$

$\underbrace{V \cdot A^n X}$ ← answer, kinda.

AX so becomes:

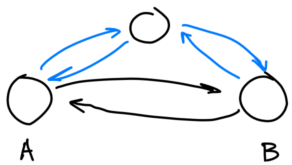
$$\boxed{V \cdot (A + A^2 + \dots + A^7) X}$$

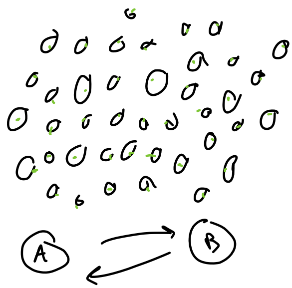
OK, play against every possible strat.

every possible strategy: all-d always best.
because the problem is that we include stupid strategies that smart strategies will not do so well against.

hmm so I guess that means, ...
that we're back to square 1.

OK new goal: construct ANY context in which TFT would win.
(no matter how specific and stupid it is)





okay maybe this: A wins against B if A wins the major

$$
\begin{pmatrix} \\ 0 & 1 & \cdots & N \end{pmatrix}
\underbrace{\begin{pmatrix} 1 & 0 & & \\ \cdot & \cdot & & \cdot \\ & \ddots & \ddots & \ddots \\ & & 0 & 1 \end{pmatrix}}_{N+1} N+1
$$

$x_1 =$

oh sure it seems determined.

$$x_{N-1} = \frac{P_{N-1,N} + P_{N-1,N-2} \cdot x_{N-2}}{1 - P_{N-1,N-1}}$$

oki not obvious. but I guess provable by induction.

the only stationary vector should be

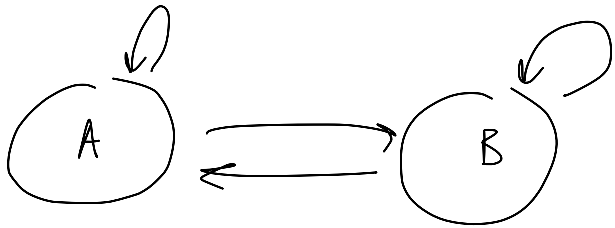
$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix}$ right?

not obvious

oh ok it is obvious: at every step we're leaking into the end states, so can't have stationary distr that is not zero on middle states ü

this idea is very appealing.

so this actually models



we want to prove something about this

so, it will be hard to show that one strategy is "optimal".

can we formulate as a normal DFA

e.g. "accepts iff s[-1] is the action taken against s[:-1]"

yea. this would enable us to use NFAs wow.

maybe even extend this to PDAs and CFGs?

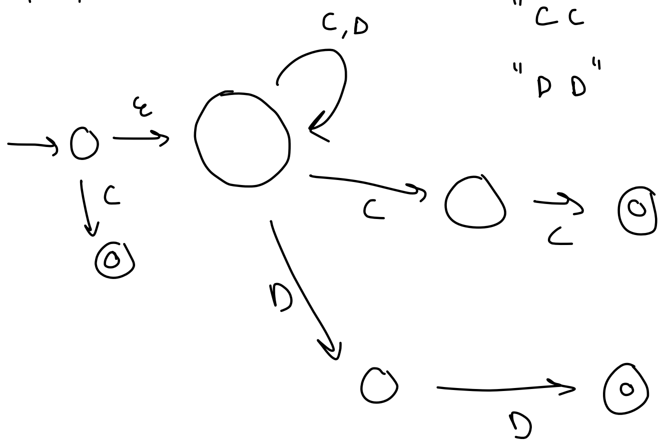
"C"

"CC"

"DD"

so anything that ends with CC or DD

TFT:



what are the relevant definitions at play?

- evolutionary stability

↳ x^* becomes relevant

OK so that paper captures essentially everything:

(1) to be ES, you need to cooperate w/ yourself

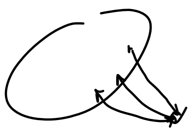
(2) there exists an ES strategy

this definition deals with infinitesimally small addition.

possible follow-up: is that strategy unique?

ALLD vs Pavlov not obvious to me.

hmm ok so problem is that infinite - infinite seems degenerate



(1) either invader is better against pervader than pervader is against

or

(2) invader is equally good against pervader but better against itself than pervader is against itself

wait so in the infinite game it is obvious that TFT is evolutionarily stable because it only loses by a little bit.

↳ there's still, however, a possibility of other strategies existing (i.e. other strategies are not killed off) ← which would be impossible to ensure, however

OK, how to prove that Pavlov is ES?

Pavlov depends on it.

$$Q: \ \frac{P+T}{2} > R \ ?$$

in this particular example, no, otherwise unclear.

C
C
D
D } C

C
D } D

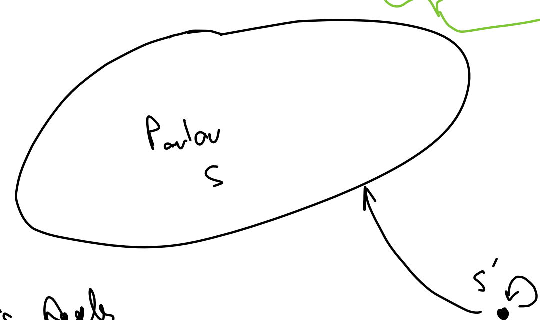
DC

Pavlov	All D
C	D
D	D
C	D
D	D
C	D
D	D

$\frac{P+T}{2}$

$\frac{P+S}{2}$

does not matter much.



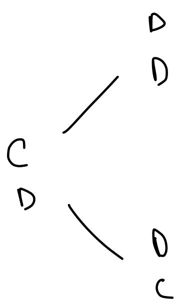
~~P>T~~

$T > R > P > S$

~~dafcg~~ this needs wrong broken

$$v^m(s,s) = u^* \quad \forall m.$$

$$v^{m^*}(s',s') > v^{m^*}(s,s') \ ?$$



so best strategy will ~~aaaye~~ always C against Pavlov.

?

s' needs to get $\geq$ C-payoffs against Pavlov.

oh ok, clean that this but works.

oh so against mostly-D: do mostly-D until mistake, then tit-for-tat

hmm oh so double-Pavlov is also ES.

yeah ok, really any strategy for which

$$V^m(S,S) = u^* \quad \forall m$$

is going to be ES under this definition

and that is really super easy because you can keep track .... wait no, you can't keep track of all m mistakes

so really, Pavlov can be generalized to:
generally always $\frac{C}{C}$,
if opponent defects then do something

maybe do finite game with infinite population

↪ not nice for finite automaton etc.

---

all pumping lemma $\Rightarrow$ $\exists M$ s.t. defects $< M$ times against itself

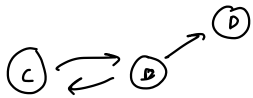
else $\Rightarrow$ $\exists M$ s.t. after turn M it can defect against itself at most once.

$\Rightarrow$ $\exists M$ s.t. after turn M it will never defect against itself.

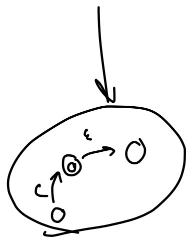
ummm ok this argument is trivialized by "eventually repeats itself"

1. my FA idea (15 min)
2. understanding the proof in the given paper (30 min)
3. applying that proof to finite games. (30 min)
4. figure out what to say to sipser (30 min)

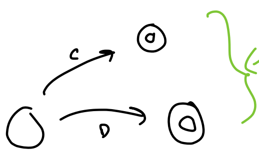
1



etc is equiv to DFA.



✓



hard to enforce cannot accept more than once

OK. So we certainly can convert the natural FA to a normal DFA/NFA and then use pumping lemma.

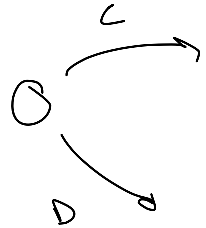
Ask sipser: is there another way of converting that natural FA to a normal FA that maybe has equivalence?

2

note again: finite automata are the way to go.
↪ equivalent to finite-memory Turing machines

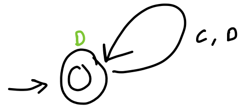
so how do we model noise?

<u>option 1:</u> w/ probability, the finite auto



ok. so we basically have probability $1-p$ of following right transition, and p of following the wrong one.

ex allD:



oh wait. that's bad.

second option: probability p of saying the wrong thing.

all D:

↖ should be equivalent. almost.

OK so how to formulate ES.

probabilistic!

could make them same probably

$$\text{Let } v(s_1, s_2) := \lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{T} E\left[p\left(s_1(h_{i,2}), s_2(h_{i,1})\right)\right]$$

move E[] out of the sum?

where $h_{i,2} = h_{i-1,2} \circ s_2(h_{i-1,1})$

concat

$h_{i,1} = h_{i-1,1} \circ s_1(h_{i-1,2})$

$$\text{and } p(a,b) = \begin{array}{cc} & b: \\ & \begin{array}{cc} C & D \end{array} \\ a: \begin{array}{c} C \\ D \end{array} & \begin{pmatrix} R_3 & S_0 \\ T_5 & P_1 \end{pmatrix} \end{array}$$

where $T > R > P > S$ and $(2R > T+P$

s_2

s_1

def the fitness $F(s)$ of strategy s in a population P is

$$\sum_{p \in P} f(p) \cdot v(s, p) \quad \text{where } f(p) \text{ is fraction of } p \text{ in } P.$$

def strategy s_2 can invade strategy s_1 if regardless of the distribution in P, $F(s_2) > F(s_1)$.

(→ thought: "drift" – when $F(s_2) = F(s_1)$ e.g. All C)

def strategy s_1 is evolutionary stable if it cannot be invaded by any other strategy s_2.

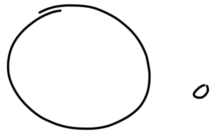
corollary: strategy s_1 is evolutionary stable iff

$$v(s_1, s_1) > v(s_2, s_1)$$

or

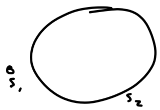
$$v(s_1, s_1) = v(s_2, s_1) \text{ and } v(s_1, s_2) \geq v(s_2, s_2)$$

$v(s_1, s_1) > v(s_2, s_1)$



wait.... what happens with which something is weird

$v(s_1, s_2) > v(s_2, s_2)$



oh wait. definitely a problem with definitions here

↑ tiny: can grow

→ is that invasion?

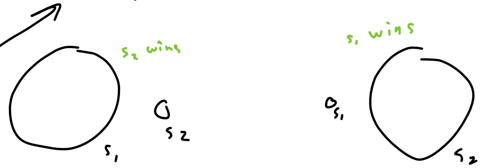
or: can take over completely?

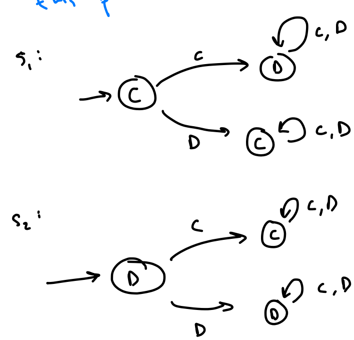
wait so:

$v(s_1, s_1) < v(s_2, s_1)$

$v(s_1, s_2) > v(s_2, s_2)$

← is this possible? yes





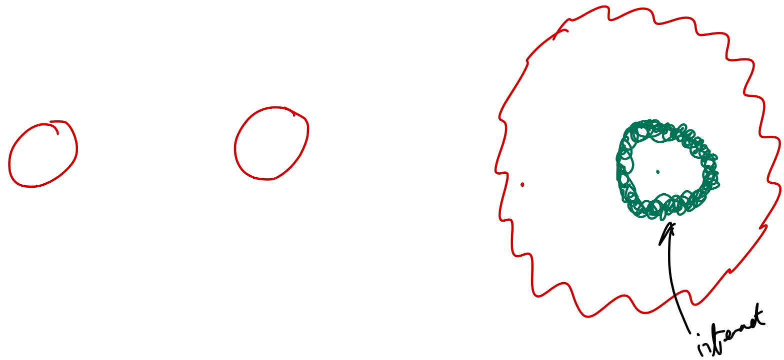
if c is true, then there exists an equilibrium position

oh yeah. is that reasonable to then say they invade each other?

very interesting!!

point 1: if s_2 can reach 70% of population, then yes definitely invade. → I do still think that maybe their definition makes more sense

scenario 2: if s_2 can reach 1% of population?



hmmm how does it affect probability.

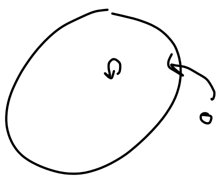
so the problem here is coexistence of multiple strategies in an equilibrium

→ and that that can happen organically.

for now I think I'll ignore this.

how define invasion then?

$v(s_1, s_1) > v(s_2, s_1)$



can find arbitrarily small size to get contradiction

2nd condition obvious.

ok ES ⇒ then

¬then ⇒ ¬ES.

$v(s_1, s_1) \leq v(s_2, s_1)$

$v(s_1, s_1) < v(s_2, s_1)$
can invade

other: also can invade.

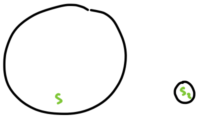
OK good so I have the definitions down.

what do I want to prove?

check just prove this

ok, start with condition 1:

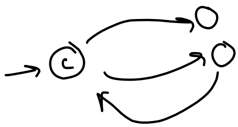
$\forall s_2: v(s,s) > v(s_2,s)$



$\iff \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[p(s(h_{t-1}), s(h_{t-1}))]$

$> \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[p($

uhhhh this definition makes no sense

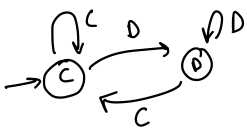


$p(s_1, s_2 | h_{t-1})$

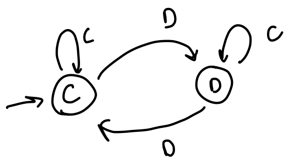
kinda.
o. no no yea ok.

oh wait. linearity of expectation means that we actually can do this, and treat them as separate things actually. Interesting.

TFT:



Pavlov:

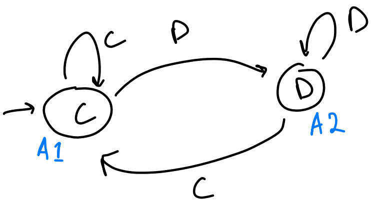


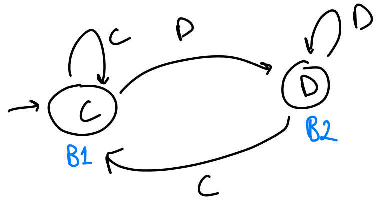
no nonono makes no sense does it.

interesting, they're very similar.

suppose tft does 1 mistake. then really bad.

ok start to compute $v(s_{TFT}, s_{TFT})$ because it is not at all obvious how one would do so.





$v(s,s)$? oh wait I can't compute it without a definition lol

but no I have the intuitive definition clear:

$$\lim_{T\to\infty} E[\text{payoff given } T \text{ rounds}]$$

what model do we have? follow transition.

aperiodic?

ok what if we take cross-product states and define transitions.

$\{A1B1, A1B2, A2B1, A2B2\}$

← regular Markov chain

$$M = \begin{array}{c} \\ A1B1 \\ A1B2 \\ A2B1 \\ A2B2 \end{array} \begin{array}{cccc} A1B1 & A1B2 & A2B1 & A2B2 \\ \hline \end{array}$$

$$M = \begin{array}{c|cccc} & A1B1 & A1B2 & A2B1 & A2B2 \\ \hline A1B1 & (1-p)^2 & (1-p)p & p(1-p) & p^2 \\ A1B2 & p(1-p) & p^2 & (1-p)^2 & (1-p)p \\ A2B1 & (1-p)p & (1-p)^2 & p^2 & p(1-p) \\ A2B2 & p^2 & p(1-p) & (1-p)p & (1-p)^2 \end{array}$$

so markov matrix in $\pi^* = \pi^* M$ form

$\boxed{M: \text{go from row} \to \text{column}}$

hmm this feels like a matrix product kinda oh wow. we're interested in the stationary state we arrive at starting from

$$(1, 0, 0, 0).$$

oh nah. does it necessarily have a stationary state?

no eg.

$$\begin{pmatrix} & & & 1 \\ 1 & & & \\ & 1 & & \\ & & 1 & \end{pmatrix}$$

which is periodic.

but if it had a stationary state, does that mean that it converges to, does that the time spent in that that is the each state ???? ahhh.

oh maybe because lazy (you can

oh no.... if more states there's no guaranteed unique convergence

which makes sense because you're not guaranteed to converge if you start in disconnected state.

wait really the only problem is periodicity.

$$\pi M = \pi$$

$$\pi (M - I) = 0$$

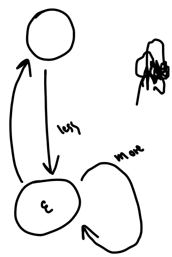
ok but now solve TFT

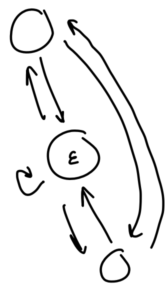
* has unique stationary dist
*

in TFT: will converge to a unique distribution

→ then, can compute the value by multiplying with value vector. eller?

yea ok if converges then doesn't matter how fast it converges, but we can just take stationary state and multiply with payoffs.





$\varepsilon_1 \to \varepsilon_2 \to \varepsilon_1$

ahh. always less than 1 coming out.

so self loop

en nettoinkomst never escapes!

ok so self loop state has constant value ε. ahh but needs to be strongly connected typ så den inte samlar

ah ok makes sure

so in TFT case we can just compute steady state vector

beräkna exp. expectation om det är alltid

hamnar i steady state

$\longrightarrow$ $\bar{j}^a$, för steady state är en prob dist.

regular if some ~~power~~ power has positive elements

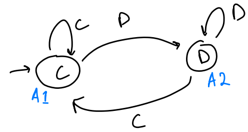
hur ofta finns alltid risk att man stannar kvar

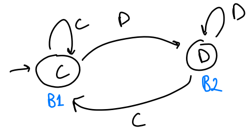
not ~~necessarily~~ true!

bäst av TFT: (has unique stationary state)

$\pi M = \pi$

$$\begin{array}{c|cccc} & A1B1 & A1B2 & A2B1 & A2B2 \\ \hline A1B1 & (1-p)^2 & (1-p)p & p(1-p) & p^2 \\ A1B2 & p(1-p) & p^2 & (1-p)^2 & (1-p)p \\ A2B1 & (1-p)p & (1-p)^2 & p^2 & p(1-p) \\ A2B2 & p^2 & p(1-p) & (1-p)p & (1-p)^2 \end{array}$$





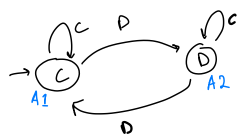
payoff (for A): $\left(\overset{A1B1}{R} \quad S \quad T \quad P \right)$ $\quad (3 \quad 0 \quad 5 \quad 1)$

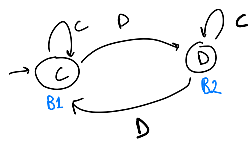
stationary vector: $(.25 \quad .25 \quad .25 \quad .25)$

wtf.

Pavlov then:

$$\begin{array}{c|cccc} & A1B1 & A1B2 & A2B1 & A2B2 \\ \hline A1B1 & (1-p)^2 & (1-p)p & p(1-p) & p^2 \\ A1B2 & p^2 & p(1-p) & (1-p)p & (1-p)^2 \\ A2B1 & p^2 & p(1-p) & (1-p)p & (1-p)^2 \\ A2B2 & (1-p)^2 & p(1-p) & (1-p)p & p^2 \end{array}$$





interesting:

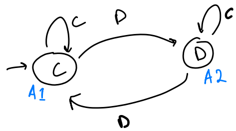
stationary: $(0.448,\ p(1-p),\ p(1-p),\ 0.232)$

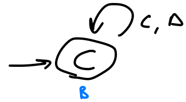
$2 \qquad 2p - 2p^2 + k = 1$

$1 - 2p + 2p^2$

Pavlov vs All C

$s_1 = $ Pavlov, $s_2 = $ All C.





$$\begin{array}{ccc} & A1B & A2B \\ A1B & \multicolumn{2}{c}{\multirow{2}{*}{$\begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}$}} \\ A2B & & \end{array}$$

(eig: $(\frac{1}{2}, \frac{1}{2})$)

$(-4p^3+7p^2, \quad -p^2, \quad -p^2, \quad 4p^3-5p^2)$

$(1-4p, \quad .9p, \quad .9p, \quad 22p)$

$$v(s_1, s_1) = (-4p^3+7p^2-4p+1,\ p(1-p),\ p(1-p),\ 4p^3-5p^2+2p) \cdot (R,S,T,P)$$

$$v(s_1, s_2) = \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right) \cdot (R,S,T,P)$$

$$v(s_2, s_1) = \left(\frac{1}{2}, \frac{1}{2}, 0, 0\right) \cdot (R,S,T,P)$$

$$v(s_2, s_2) = (1, 0, 0, 0) \cdot (R,S,T,P)$$

$v(s_1, s_1) > v(s_2, s_1)$ as $p \to 0$ so not invaded.

makes a lot of sense.

yeah noise makes a lot of sense
this model overall makes a lot of sense

Pavlov vs all D? does not affect evolutionary stability.

aha. so we first compare as $p \to 0$
then $\frac{1}{p}(\ )$ as $p \to 0$
then $\frac{1}{p^2}$ . . . .
and so on

nothing special about being perfect there.

so ..... what's my conclusion?

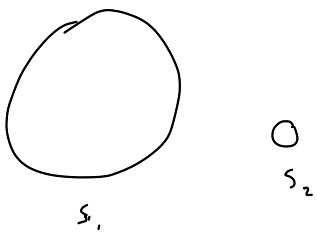
→ I need to prove the damn thing.

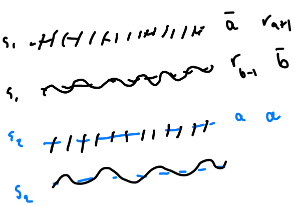
Suppose $v(s_1, s_1) < R$

then create $s_2 \sim s_1$ s.t. $v(s_2, s_1) = v(s_1, s_2) = v(s_1, s_1)$

and $v(s_2, s_2) = R$

necessarily end up in equilibrium?





$v(s_1, s_1) = v(s_2, s_1)$

$v(s_1, s_2) < v(s_2, s_2)$

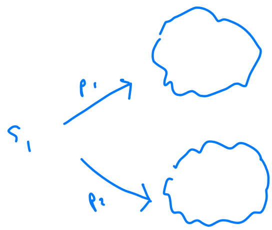
plan.

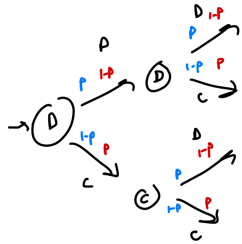
consider strategy that does allD + $\frac{1}{n}$ C
and allD if deviates.
oh → allD num.

idea is: flip two adjacent states
eg:

yeah I think we can just do that always without changing final outcome right.

nonono





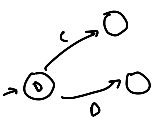
suppose it tries to do smth differently.

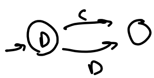
ok. so if we invert both then we good.

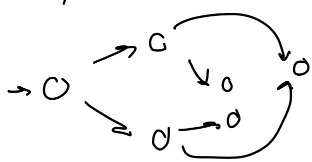
ahh lots of wrongs here.

eg: know that s_1-s_1 does D-D → perform C there.

← always better to perform C here! nono not necessarily







then it works.

ahhhh two models very different.

① perception error:

② action error: we know that we'll always & have a nonzero possibility being anywhere

hmm maybe my $p \to 0$ thing needs to happen first

nonono, not different me stupid.

uhhhhhh

ugh

what does $p \to 0$ mean?

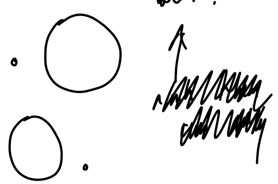
no can't think like that.

but what does it mean that $p \to 0$.

I think size of p probably doesn't matter for invasion.

why not define ~~invade~~ instead of have extinguish?

i.e. always ~~invade~~ [illegible].



$poly(p) > 0$?

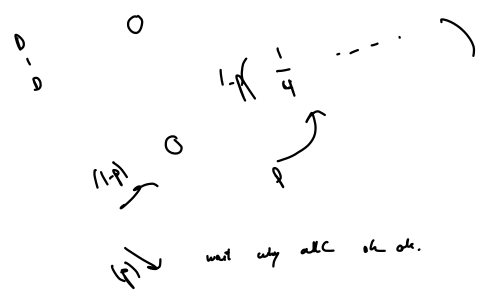
must have root at some $p \in (0,1)$ then.

probably has though.

$p \to 1-p$ completely reverses all behavior (?)

hmm ok.

if we know constant ~~[illegible]~~ probability of D-D for example why would that change things?



wait why allC oh ok.

so maybe their idea isn't too stupid ....

given worst history, self identify.

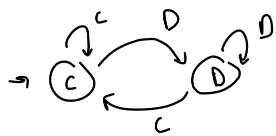
so what is their strategy again?

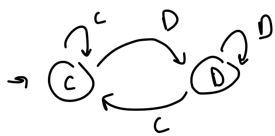
given history

I actually think it will work. ~~the~~ like our result is honestly very very similar come to think about it. the $p \to 0$ thing of $v(s,s) = R$ is precisely just the top level. wow ok. really try, next time.

OK let's try to invade TFT.

clearly h^* doesn't matter. because whatever the distribution is, will end up in same state





ok so take $p \to 0$

doesn't help I think

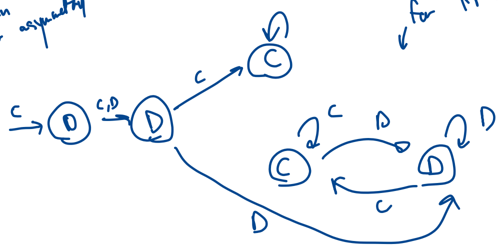
so ok. given history that minimizes future gain.

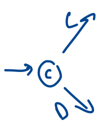
then doing will actu

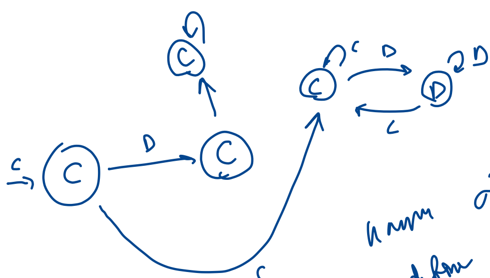
maybe not

ok sure can be after asymmetry

for TFT this works.







hmm ok if we define it another way.

w.r.t given proportion I think we can design strategy

C	C
D	C

ensure this only happens suppose that then.

oh interesting:

WE CAN DESIGN A STRAT GIVEN THE FRACTION OF INVASION, SO THE STRAT WILL CHANGE WITH THE PROPORTION BUT I THINK THAT IS FINE

so, define invasion in terms of parameter α.

hmm this is weird but this definition makes a lot more sense, and also makes this possible to prove.

also: strengthens evolutionary stability, so it seems reasonable

oh, I guess the good thing with the infinitesimal is that it is actually (attaining same level)

OK so we probably could change our definitions like that.

but can we do something actually interesting?

OK I like the idea of infinite repetition. like you need to prove yourself infinitely many times

problem: $(1-p)^\infty = p^\infty$

hmm. can we construct

like maybe $f(1-p) = 1$ $\quad \lim_{n\to\infty} f(p) = 0$?

~~(as / p → 1 ∞)~~

ok so step function is great but hard to construct.

$1-p + (1-p)^2 + \dots =$

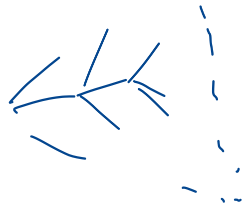
$= \frac{x^{n+1}-1}{x-1} = \frac{1}{1-(1-p)} = \frac{1}{p}$

- $p^\infty = 0 \qquad (p+\frac{1}{2})^\infty = 0$
- ok so construct $\frac{1}{2}$? ← probably possible

hmm kanske kan fungera

strategi:

dela upp i branches:



→ skapa $\frac{1}{2}$ i en av dem, p i en av dem

→ kombinera dem

→ oh shit.

ok säg att vi kan skapa $p(p+\frac{1}{2})$

hur kan vi använda det?

→ kan inte $p^\infty (p+\frac{1}{2})^\infty$

ok I think what I'm doing makes no sense.

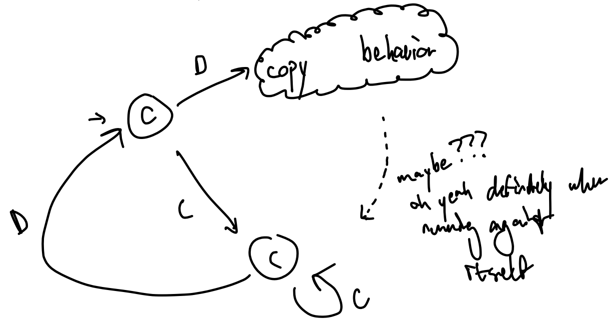
ok so the problem is that a strategy could possibly [illegible] the other

we just need to make sure that doesn't happen to us.

sure, 1 way to do that is to just copy. but there has to be other ways.

(assuming here we only have 1 Markov CC but can easily extend)

oh ok very simple I think:



when bad: D D

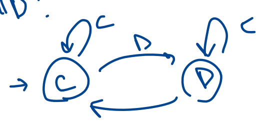
oh ok so when change D → C potentially bad

hmmm

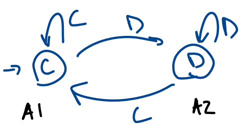
oh ok so this self-identifies continuously. maybe ??? that just works...

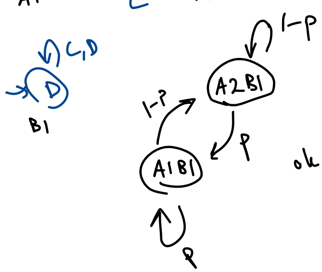
oh wow very interesting

wait..... can Pavlov actually invade allD?

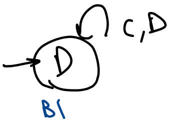


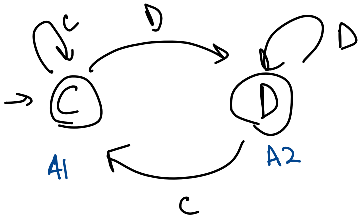
ah no. but TFT?





ok good





ok så allt som är steady
state: D.

i ~~okej allD can not~~
~~be invaded~~

ok ok

	A1 B1	A2 B1
A1 B1	p	1 - p
A2 B1	p	1-p

fuck.

vector is (p, 1-p), so we get that TFT loses

darn.

damn ok same thing again: for sufficiently small p, can invade.

oh ok so we have a bigger problem at hand:

AllD can now not be invaded.

Possible ways to fix this:

① make $p \to 0$ earlier

i.e. instead of fixing p and then letting $\alpha \to 0$,

fix α and let $p \to 0$.

② define invasion differently:

a strategy is ES iff given a populat

Try ①

invade if:

$F(s_2) > F(s_1)$. Let $f(s_2) = \alpha$ typically $\alpha << 1$.

$s_1 = \text{All D}$

that is:

$$f(s_1)\, v(s_2, s_1) + f(s_2)\, v(s_2, s_2) > f(s_1)\, v(s_1, s_1) + f(s_2)\, v(s_1, s_2)$$

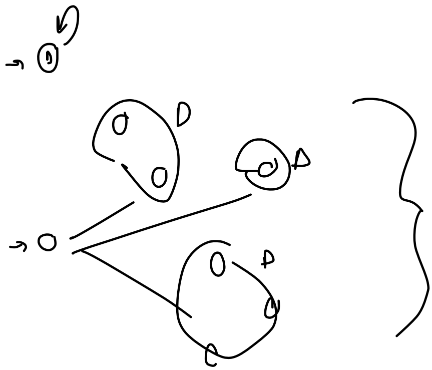
then:

$$(1-\alpha)\underbrace{v(s_2, s_1)}_{D - p^k \cdot \sqcup} + \alpha \cdot \underbrace{v(s_2 s_2)}_{\text{big}} > (1-\alpha)\underbrace{v(s_1, s_1)}_{D} + \alpha \cdot \underbrace{v(s_1, s_2)}_{D + p^k \cdot \sqcup}$$

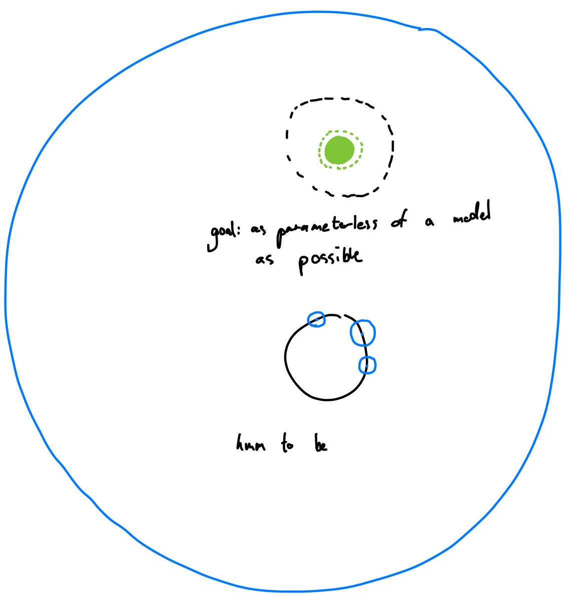
just approach this naturally.

then evolutionary stable would mean something like:

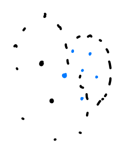
even in infinite populations, can't be invaded

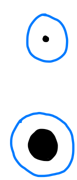


then it cannot collude with anyone.



again this seems hard without discreteness.





oh ok so this model suggests that it is reasonable to define invasion as starting from some fraction ε and then that fraction will only increase invade f:

$F(s_2) > F(s_1)$. Let $f(s_2) = \alpha$

$s_1 = \text{AllD}$

that is:

$$f(s_1)\,v(s_2,s_1) + f(s_2)\,v(s_2,s_2) > f(s_1)\,v(s_1,s_1) + f(s_2)\,v(s_1,s_2)$$

then:

$$(1-\alpha)\,v(s_2,s_1) + \alpha \cdot v(s_2,s_2) > (1-\alpha)\,v(s_1,s_1) + \alpha \cdot v(s_1,s_2)$$

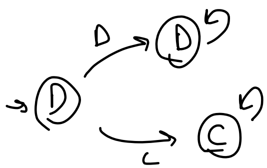
this is true, and we know that

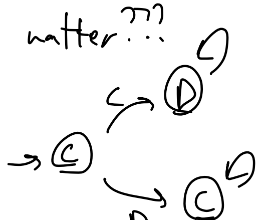
$\alpha + \varepsilon$ then:

$$-\varepsilon \cdot v(s_2,s_1) + \varepsilon \cdot v(s_2,s_2) \overset{?}{\geq} -\varepsilon \cdot v(s_1,s_1) + \varepsilon\, v(s_1,s_2)$$

$$v(s_1,s_1) + v(s_2,s_2) \overset{?}{\geq} v(s_2,s_1) + v(s_1,s_2)$$

weird; size of epsilon doesn't matter???





ok so this is not always true!

oh! but we get to choose s_2!!!!

so can we make sure that this doesn't happen?

oh only get to pick s_2 for disproving ESS, not proving it :(

$$(1-\alpha)\, v(s_2, s_1) + \alpha \cdot v(s_2, s_2) > (1-\alpha)\, v(s_1, s_1) + \alpha \cdot v(s_1, s_2)$$

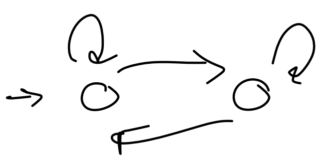
given α, find s_2. can easily prove that that is possible!

does this mean...... that we're done?

ahhhhhhhhhhhhh I think Pavlov is invadable now

yes $\ddot{n}$. So... we need to strike a balance......

wait ok I'm not convinced anymore.

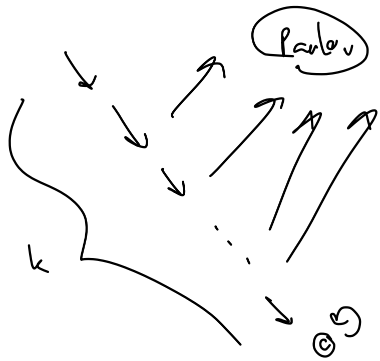


Pavlov

$$v(s_2, s_1) = (1-p^k)\, v(s_1, s_1) + p^k \cdot S$$

$$v(s_2, s_2) = p^{2k} \cdot R + (2p^k - 2p^{2k}) \cdot \left(\frac{S+T}{2}\right)$$

$$+ (1-p^{2k})^2 \cdot v(s_1, s_1)$$



$$v(s_1, s_2) = (1-p^k)\, v(s_1, s_1) + p^k \cdot T$$

$$(1-\alpha)\, v(s_2,s_1) + \alpha \cdot v(s_2,s_2) > (1-\alpha)\, v(s_1,s_1) + \alpha \cdot v(s_1,s_2)$$

FAKE NEWS.

$$v(s_2,s_1) = (1-p^k)\, v(s_1,s_1) + p^k \cdot S$$

$$v(s_2,s_2) = p^{2k} \cdot R + (2p^k - 2p^{2k}) \cdot \frac{S+T}{2} + (1-p^k)^2 \cdot v(s_1,s_1)$$

$$\gamma := v(s_1,s_1)$$

wow ok NOT obvious at ALL

$$v(s_1,s_2) = (1-p^k)\, v(s_1,s_1) + p^k \cdot T$$

$$-(1-\alpha)\cancel{p^k}\gamma + (1-\alpha)\cancel{\,}p^k S + \alpha p^{2k} R + \alpha(p^k - p^{2k})(\cancel{S}+\cancel{T}) + \alpha(1-\cancel{p^k})^2 \gamma - \alpha(1-\cancel{p^k})\gamma - \alpha \cancel{p^k} T \overset{?}{>} 0$$

$$\gamma\left(-p^k + \cancel{\alpha} p^k + \cancel{\alpha} - 2\cancel{\alpha} p^k + \alpha p^{2k} - \cancel{\alpha} + \alpha \cancel{p^k}\right)$$

$$\gamma(\alpha p^{2k} - p^k) + S(p^k - \cancel{\alpha p^k} + \cancel{\alpha p^k} - \alpha p^{2k})$$

$$S(p^k - \alpha p^{2k})$$

$$+\, T(\alpha p^k - \alpha p^{2k} - \alpha p^k)$$

$$T(-\alpha p^{2k})$$

$$+\, \alpha p^{2k} R \overset{?}{>} 0$$

$$\gamma(\alpha p^{2k} - p^k) + S(p^k - \alpha p^{2k}) - T\alpha p^{2k} + \alpha p^{2k} R \overset{?}{>} 0$$

$$\gamma(\alpha p^k - 1) + S(1 - \alpha p^k) - T\alpha p^k + \alpha p^k R \overset{?}{>} 0$$

$\gamma(\alpha p^k - 1) + s(1 - \alpha p^k) - T\alpha p^k + \alpha p^k R \overset{?}{\geq} 0$

$\gamma \alpha p^k + S + \alpha p^k R \overset{?}{\geq} \gamma + S \alpha p^k + T \alpha p^k$

omg wow. as $p \to 0$ this is NEVER true.

lolol I played myself

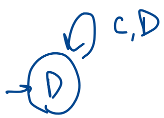
i've been fake news this entire time lololol

$T > R > P > S$

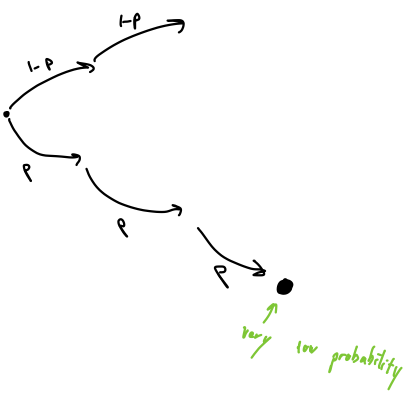
how invade AllD here?

$(1-\alpha) v(s_2, s_1) + \alpha \cdot v(s_2, s_2) > (1-\alpha) v(s_1, s_1) + \alpha \cdot v(s_1, s_2)$

$\forall \alpha$ (small), find s_2 such that ↑ is true



$v(s_1, s_1) = P$



$(1-\alpha) v(s_2,s_1) + \alpha \cdot v(s_2,s_2) > (1-\alpha) v(s_1,s_1) + \alpha \cdot v(s_1,s_2)$

p^2, $2p(1-p)$, $(1-p)^2$

oh so I don't quite like this.

$v(s_1,s_1) = \gamma.$

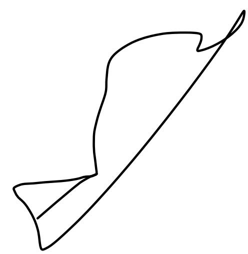
$v(s_2,s_1) = (1-p^k)\gamma + p^k \cdot S$

$v(s_2,s_2) = (1-p)^{2k} \cdot R + 2 \cdot (1-p)^k \cdot (1-(1-p)^k)\left(\frac{S+T}{2}\right) + (1-(1-p)^k)^2 \cdot \gamma$

$v(s_1,s_2) = (1-p^k)\gamma + p^k \cdot T$

$$(1-\alpha)(-p^k)\gamma + (1-\alpha)p^k S + \alpha(1-p)^{2k} R + \alpha(1-p)^k \cdot (1-(1-p)^k)(S+T) + \alpha(1-(1-p)^k)^2\gamma$$

$$-\alpha(1-p^k)\gamma - \alpha p^k \cdot T \overset{?}{>} 0$$



$$\gamma\left(-p^k + \alpha p^k + \cancel{\alpha} - 2\alpha(1-p)^k + \alpha(1-p)^{2k} - \cancel{\alpha} + \alpha p^k\right)$$

$\gamma = R - \beta$

$$+ S\left(p^k - \alpha p^k + \alpha(1-p)^k - \alpha(1-p)^{2k}\right)$$

$(R+\gamma)(1-p)^{2k} \overset{?}{>} (\gamma+\gamma)(1-p)^k$

$(R-2\beta)\left((1-p)^{2k} - (1-p)^k\right) + \beta(1-p)^{2k} \overset{?}{>} 0$

$$+ R\left(\alpha(1-p)^{2k}\right)$$

$a(q^2-q) + bq^2 \overset{?}{>} 0$

$$+ T\left(\alpha(1-p)^k - \alpha(1-p)^{2k} - \alpha p^k\right)$$

oh so seems like increasing k is just.... bad?

safely set α sufficiently small so l

I feel like this is fake news!

ohhhh the $\frac{S+T}{2}$ bound always holds which makes sense ($\frac{S+T}{2}$ is symmetric for large k & small p). for small p we get the R bound.

$$\gamma\left(-p^k + \alpha p^k + \cancel{\alpha} - 2\alpha(1-p)^k + \alpha(1-p)^{2k} - \cancel{\alpha} + \alpha p^k\right) + R\left(\alpha(1-p)^{2k}\right)$$

$$\gamma\left(-p^k + \alpha p^k - 2\alpha(1-p)^k + \alpha(1-p)^{2k} + \alpha p^k\right) + R\alpha(1-p)^{2k} \overset{\circ}{>} 0.$$

$$\gamma\left(2\alpha p^k + \alpha(1-p)^{2k}\right) + R\alpha(1-p)^{2k} \overset{\circ}{>} \gamma\left(p^k + 2\alpha(1-p)^k\right)$$

$$\gamma 2\alpha p^k + \alpha(1-p)^{2k}(\gamma + R) \overset{\circ}{>} \gamma p^k + \alpha(1-p)^k(\gamma+\gamma)$$

$$\gamma p^k(2\alpha - 1) + \alpha(1-p)^k\left((1-p)^k(\gamma+R) - 2\gamma\right) \overset{\circ}{>} 0$$

$R = \gamma + f(p)$ then.

$$\gamma p^k(2\alpha - 1) + \alpha(1-p)^k\left(2\gamma\left((1-p)^k - 1\right) + f(p)(1-p)^k\right) \overset{\circ}{>} 0$$

$1-p = q$

$$2\gamma\left(q^k - 1\right) + f(q)\, q^k \overset{\circ}{>} 0$$

$$q^k\left(2\gamma + f(q)\right) \overset{\circ}{>} 2\gamma$$

$$f(q) \overset{\circ}{>} \frac{2\gamma}{q^k} - 2\gamma$$

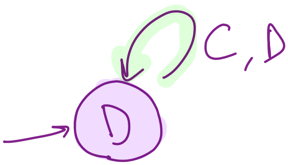
suppose $f(p) = c \cdot p$
then $f(q) = c(1-q) =$
$= c - cq$

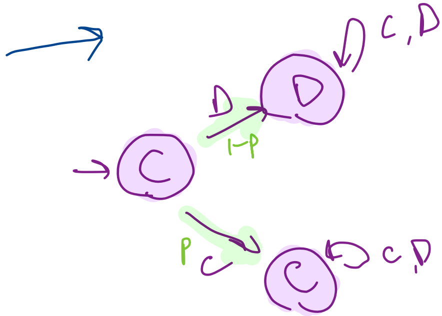
$2\gamma q^k + c q^k \overset{\circ}{>}$
$\overset{\circ}{>} 2\gamma + c q^{k+1}$

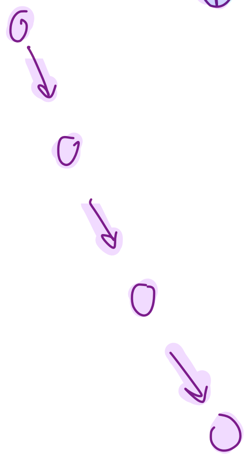
limits k by above

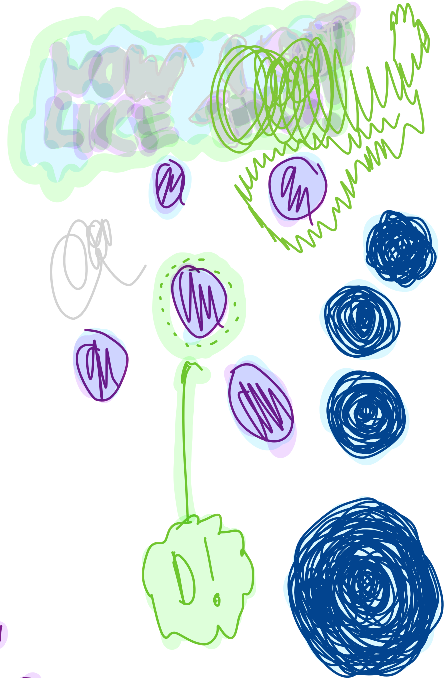
hmmmm definitely not doing

hmmm this is not really satisfying.
I would want it to hold for any $p < 0{,}5$ really.









$s_1 - s_1 : P$

$S_1 - S_2 : (1-p) \cdot P + p \cdot T$

$s_2 - s_1 : (1-p) \cdot P + p \cdot S$

$s_2 - s_2 : (1-p)^2 \cdot R + p(1-p)S + p(1-p)T + p^2 \cdot P$

more abstractly:

we have polynomial of

$\forall \alpha$

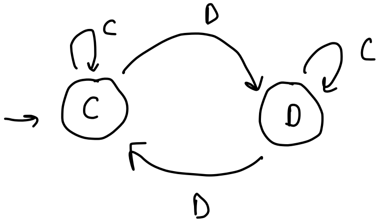
fix

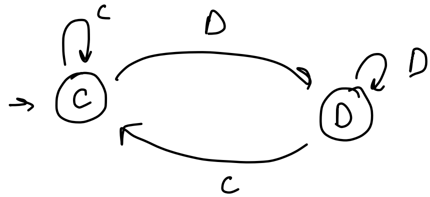
given

$p, \alpha, k, \delta.$

→ what properties of

only real parameter.





Leaves out:

* might want to have probabilistic players.

fantasy level: (really interesting)

* more states can kill all small state machine.

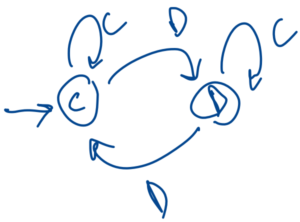
evolutionary: advantage to having more horsepower more states can always kill fewer states?

→ go back to deterministic model

→ look at language?

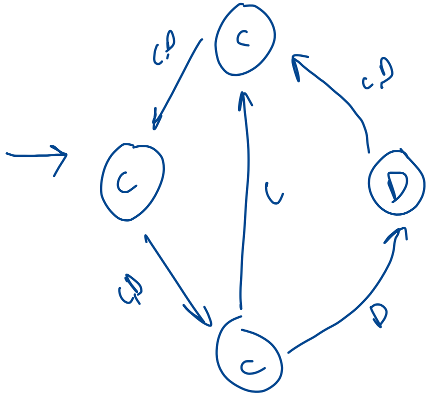
→ still look at infinite case?
(will have to look at loops)

so can create machine that splits into branches depending on which of n strategies there are!



look at machines with $\Theta(n)$, $\Theta(n^2)$ etc states?

→ then parameter n is game length!



uncommon occurrence

actually ahhh, now I know what my problem is.....
Pavlov needs to be ESS,
and for fix P that ∀ st
doesn't make sense.

VERDICT

the definition I use makes sense!

because for a given P might be better than Pavlov (eg longg loop)

so won't be able to state the result at all really!

oh, could say that

S_1 is ES $\Rightarrow$

$\Rightarrow V_{S_1}(S_1) = R - \sigma(p)$

random thought?

can we say that we

need $\delta \geq R - P$

or something like that?

$\gamma = R - f(p)$ then.

$$\gamma p^k (2\alpha - 1) + \alpha (1-p)^k \left(2\gamma((1-p)^k - 1) + f(p)(1-p)^k \right) \overset{?}{>} 0$$

$$2R((1-p)^k - 1) + 2f(p) - f(p)(1-p)^k \overset{?}{>} 0$$

$$f(p) > \frac{2R(1-(1-p)^k)}{2-(1-p)^k}$$

$k=1 \Rightarrow$

$$f(p) > \frac{2R \cdot p}{1+p}$$

So: must have $\geq R - \dfrac{2Rp}{1+p} =$

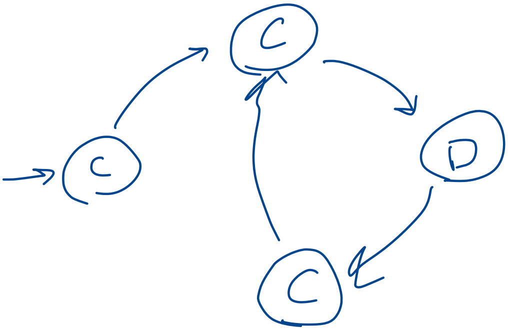
$$= \frac{R - Rp}{1+p} =$$

$$= R\left(\frac{1-p}{1+p}\right)$$

$p = 0.25$

then: $\geq 0.6R$

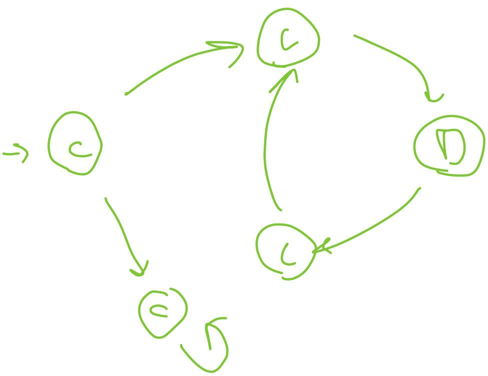
so $\cancel{\tfrac{3}{4}}\ \frac{2}{3}c$ is good



here, behavior is deterministic hmm

R+P vs S+T?

not obvious



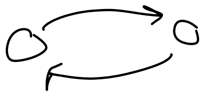
last n moves only

ergodic Markov chain

non-0 prob of occurring

OK, do we need to restrict to ergodic Markov chains here??

oh ok. if eigenvalue 1 then has stationary distribution.



wait oh. what can happen, really.

we take vP^n as $n \to \infty$

obviously it goes to some value right. because it is bounded

oh might also just alternate. but mehhhhhh that's quite obviously fine

~~in which cases not e~~

we do not need ergodicity

— we only need non-periodicity

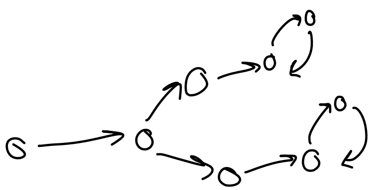
OK so this model doesn't really make sense.

what is periodicity?

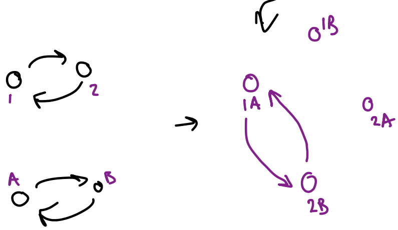
↳ no self-loops.

→ what if always small probability of transforming into another state

eg model "failure of mind"



~~can still define time average payoffs~~



problem: noise does nothing to transitions that take C & D to same place.

ok so it complicates the analysis but it should be fine.

ok Markov chain stationary state. and we're taking the limit of it. wow.

---

suppose you're in a chain

$p_1 \to p_2 \to \dots \to p_n \to p_1$

claim: $\frac{p_1+\dots+p_n}{n}$ is stationary.

proof: obvious

only now need to prove that we reach a cycle.

proof: at some point will get arbitrarily close to a previous point. ■

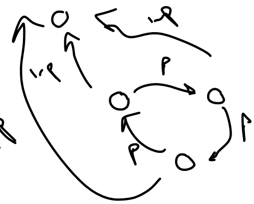
(can make ε-style boxes, use pigeonhole principle)

I want to say that it's a polynomial of p.

certainly a polynomial at every step

but.... $(p^n + \dots + 1) \to \frac{1}{1-p}$

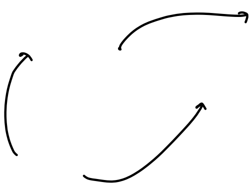
not polynomial :(



$V_{s_1}(s_1) \approx$ coefficient, stationary distribution.

hmmm not obvious to me.

altho note that $\Sigma \pi$ is always 1. so



like how can i know that the limit makes sense at all

evaluating the stationary dist

oh ok. is solution to:

$\pi P = \pi$.

all such solutions necessarily polynomials?

hmm so we never not divides. dividing by things.

$\frac{1}{p} - \frac{1}{p}$ for example!

$$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{matrix} x_1 \\ x_2 \end{matrix} = \begin{matrix} x_1 \\ x_2 \end{matrix}$$

$Av = v$

eg:

$$\begin{pmatrix} p^{10} & \\ & p^{10} \end{pmatrix} \begin{matrix} x_1 \\ x_2 \end{matrix} \quad \begin{matrix} x_1 \\ x_2 \end{matrix}$$

oh hmm

what is eigenvector?

$Ax = x$. also need $\sum x = 1$.

$(A - I)x = 0$

x is null vector. polynomial?

ok will be some polynomial in p

$$\begin{pmatrix} p & p^2 \\ 1 & p \end{pmatrix} x = 0$$

~~not obvious~~

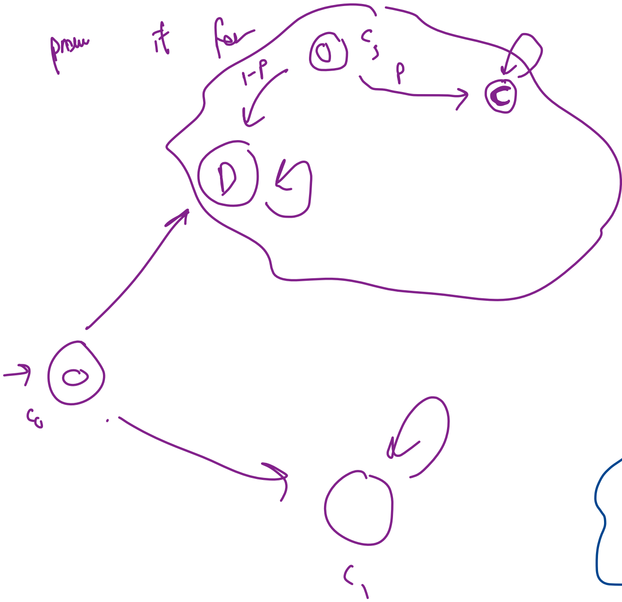
$x = \begin{pmatrix} p \\ -1 \end{pmatrix}$ $x = \begin{pmatrix} 1 \\ -\frac{1}{p} \end{pmatrix}$

is it polynomial? ahhh.

$1 + p + p^2 + \cdots + p^n$???

$$\boxed{v_s(s) \le R \quad \forall s}$$

ok can prove it for



Markov distributions are linear.

so we get $\frac{\gamma - p \cdot \text{smth}}{1-p}$

$\gamma - p \cdot \text{smth}$

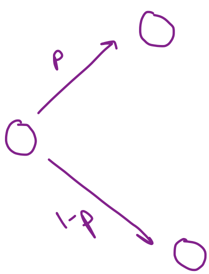
ah fuck.
not obvious at all

false, even

$F(s_2) - F(s_1) =$

ok. so we don't really need Markov chain at all here, again

$=$



pick $\varepsilon = \frac{\alpha}{2}$.

$R - 2Rp + p^2R \quad , \qquad (1-p)\,p(S+T) =$

$\geq$

$$(1-\varepsilon)\big((1-p)\gamma + pS\big) + \varepsilon \cdot \left((1-p)^2R + 2(1-p)p \cdot \left(\frac{S+T}{2}\right) + p^2\gamma\right) - (1-\varepsilon)\cdot\gamma - \varepsilon\big((1-p)\gamma + pT\big) =$$

$$= \overset{1}{\cancel{\gamma}} - \gamma p + pS - \overset{2}{\cancel{\varepsilon\gamma}} + \varepsilon p\gamma - \overset{3}{\cancel{\varepsilon pS}} + \varepsilon R - 2p\varepsilon R + \varepsilon p^2R + \overset{3}{\cancel{\varepsilon pS}} + \overset{4}{\cancel{\varepsilon pT}} - \varepsilon p^2S - \varepsilon p^2T + \varepsilon p^2\gamma \;\; - \overset{1}{\cancel{\gamma}} + \overset{2}{\cancel{\varepsilon\gamma}} \;\; - \varepsilon\gamma + \varepsilon p\gamma - \overset{4}{\cancel{\varepsilon pT}}$$

$$= -\gamma p + pS + \varepsilon p\gamma + \varepsilon R - 2p\varepsilon R + \varepsilon p^2R - \varepsilon p^2T + \varepsilon p^2\gamma - \varepsilon\gamma + \varepsilon p\gamma$$

$$= \left(pS + \varepsilon p\gamma + \varepsilon R + \varepsilon p^2R + \varepsilon p^2\gamma + \varepsilon p\gamma\right) - \left(\gamma p + 2p\varepsilon R + \varepsilon p^2T + \varepsilon\gamma\right)$$

$$= \varepsilon(R-\gamma) + p\left(S + \varepsilon\gamma + \varepsilon pR + \varepsilon p\gamma + \varepsilon\gamma - \gamma + 2\varepsilon R + \varepsilon pT\right)$$

ok, how else define $V_{s_1}(s_2)$?

$$\sum \pi_{c_1,c_2} \cdot r(c_1, c_2)$$

define as

$v_0 M^\infty$?

or $r(v) = \sum v_{c_1,c_2} \cdot r(G_{s_1}(c_1), G_{s_2}(c_2))$

then look at

expected value of time average

uhhh.

could also define as

$$\lim_{T \to \infty} \frac{1}{T} \sum r(a_i, b_i)$$

$$E\left[\lim_{T\to\infty} \frac{1}{T} \sum_{i=1}^{T} r(v_i)\right]$$

what do I need for $V_{s_1}(s_2)$?

if continuous? [redacted]

1. $\lim_{p \to 0} V_{s_1}(s_2)$ exists
2. $V_s(s) \leq R$ ← ez.
3. easy to calculate splits. ← ez.

could invent notion $E[R \mid X_i = \text{smth}]$

to get properties like

$$E[R] = E[R \mid X_1 = 2] + E[R \mid X_1 = 1]$$

ohh nice nice.

nice. oh wait 'tis not a series

$$X_n = \frac{1}{n} \sum_{i=1}^{n} v(X_i)$$

$$E\left[\lim_{n \to \infty} X_n\right] = \lim_{n \to \infty} E[X_n]$$

The limit definition makes the most intuitive sense, clearly.

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[r(\dots)]$$

always

$$\lim_{T\to\infty} E\left[\frac{1}{T}\sum_{t=1}^{T} r(\dots)\right]$$

always ~~only if converges to something!~~

So. Let X_n = r.v. for which state at time step n.

then, reward is

$$V_{s_1}(s_2) = E\left[\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} r((X_t)_1, (X_t)_2) \,\middle|\, X_0\right]$$

easy to interpret. the expected reward when playing infinitely long.

same as looking at stationary state? if can exchange limits then yes.

ah yes, if goes to stationary!!!

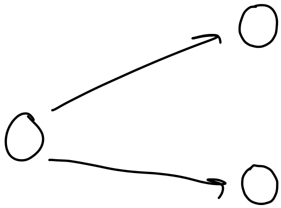
but not using ~~Fatou's lemma~~

dominated convergence theorem

can rewrite as weighted average of value of every state

the value for time $= E\left[\text{fraction of time in state } i \mid X_0 = \dots\right]$

$v_{s_1}(s_2)$



$$v_{s_1}(s_2) = \sim\sim \mid X_0 = (c_{start}(s_1), c_{start}(s_2))$$

reward

$$v_{s_1}(s_2) = E\left[\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} r(X_t) \;\middle|\; X_0 = (c_{start}(s_1), c_{start}(s_2))\right]$$

ordinary limit !!!!!

ah. the theorem becomes very easy to prove now

just condition on different events and stuff.

this definition doesn't make sense to me

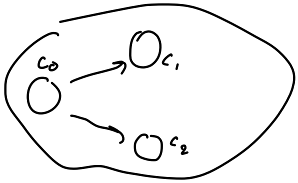
or kinda generate infinite sequence, apply this function

ohhhhh. I think that makes sense now.

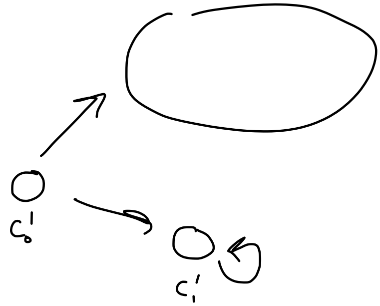
$$\lim_{T\to\infty} E\left[\frac{1}{T} \sum_{t=0}^{T} r(X_t) \;\middle|\; X_0 = \dots\right]$$

← very not obviously the same

→ much harder to argue existence.



s_1



$$v_{s_1}(s_1) = E[\dots \mid X_1 = (c_1, c_1)] \cdot P(X_1 = (c_1, c_1))$$

$$+ \dots$$

okok this works; nice.

$\lim_{p \to 0} v_{s_1}(s_2)$ exists?

$$\lim_{p \to 0} E\Big[\underbrace{\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} r(X_t)} \Big| X_0 = (c_{start}(s_1), c_{start}(s_2)) \Big]$$

does this limit exist?? uhhhh not really

$$\forall \varepsilon > 0 \ \exists N : |X_n - \bar{x}| < \varepsilon \ \forall n \geq N$$

$$\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[r(X_t)]$$

hmmmm seems like I don't wanna think about convergence

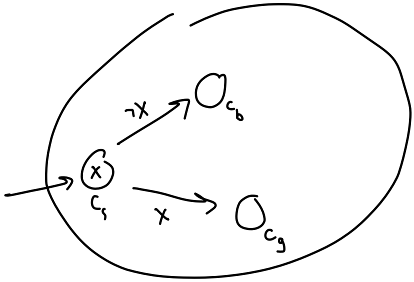
$$\lim_{p \to 0} V_{s_1}(s_2)$$

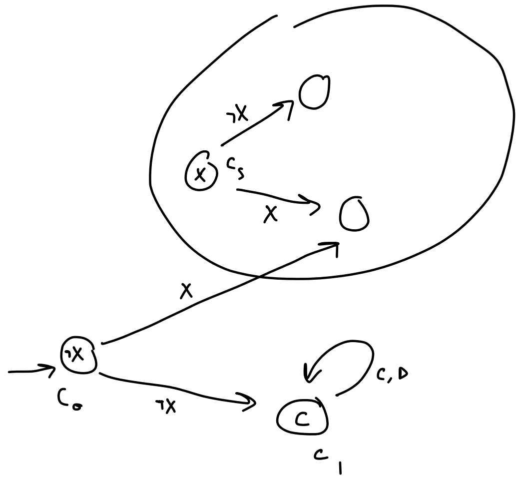
$$V_{s_1}(s_2) = E\Big[\underbrace{\lim_{T\to\infty}}_{\text{ordinary limit!!!}} \frac{1}{T} \sum_{t=0}^{T} \overbrace{r(X_t)}^{\text{reward}} \,\Big|\, X_0 = (c_{start}(s_1), c_{start}(s_2)) \Big]$$

$$\underbrace{E[\text{fraction of time in state } i \mid X_0 = \dots]}_{\text{defines time-average distribution}}$$

definitely an eigenvector

$$\pi_{c_1, c_2}$$





$$\pi^{(c_s, c_s)} = (1-p)^2 \, \pi^{(c_g, c_g)} \quad + \quad \cdots$$

structure?

→ intro ← why this specific.

→ setup

→ results

→ proofs

→ discussion of model

→ non-central but still interesting results?

→ appendix?

first state, then prove, or intersperse definitions with description?

email him

definition of payoffs?

→ which is easier to understand

how to ~~avo~~. avoid notation problems?

→ key idea is simple.

if no superscripts, assume start ...

before definition 4, outcome$_i$ = pair i correspond

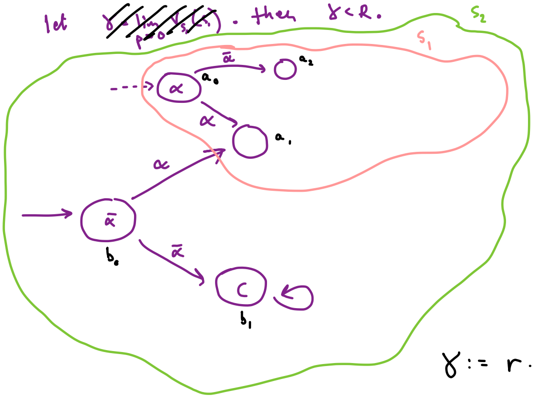
give names

pull out parts and explain them first.

use greedy

suppose that NOT $\lim_{p\to 0} v_{s_1}(s_1) = R$!

let ~~$\gamma = \lim_{p\to 0} v_{s_1}(s_1)$~~. then $\gamma < R$.



~~$\pi(a_1,a_1)$~~

for now, always pick an appropriate r. then

$r\pi^{(a_0,a_1)}$ is important

claim: $r\pi^{(c_1,c_2)} \le R$ $\forall (c_1,c_2)$.

not obvious!!!

why?

okok I need lemma 3 to be stronger

$\gamma := r \cdot \pi^{(a_1,a_1)}$

$$V_{s_1}(s_1) = r\pi^{(a_0,a_0)} = (1-p)^2 r\pi^{(a_1,a_1)} + \underbrace{2p(1-p)\, r\pi^{(a_2,a_1)}}_{+p(1-p)r\pi^{(a_2,a_1)} + p(1-p)r\pi^{(a_1,a_2)}} + p^2 r\pi^{(a_2,a_2)}$$

I think I just need much better notation...

$\not\to R$

$\Rightarrow \gamma < R$

$$\le (1-p)^2 r\pi^{(a_1,a_1)} +$$

$$\le (1-p)^2 \gamma + 2p(1-p)R + p^2 R.$$

ok cool.

$\frac{1}{O(\frac{1}{p^k})}$ maybe?

non oh ah nice $\frac{1}{\Omega(\frac{1}{p})}$

$$v_{s_1}(s_2) = r\pi^{(a_0,b_0)} = (1-p)^2 r\pi^{(a_2,a_1)} + p(1-p) r\pi^{(a_2,b_1)} + p(1-p) r\pi^{(a_1,a_1)}$$

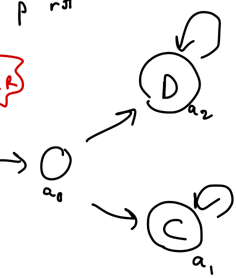
so basically, want to find 1 state that we get to with prob $1-O(p^k)$ that transitions into a_1 & a_2 s.t. $\pi^{a_1,a_2} = \pi^{a_1,a_1} = \pi^{a_2,a_1} = \gamma < R$

$+ p^2 r\pi^{(a_1,b_1)}$

$$(1-\varepsilon)\left(\pi(a_1,a_2) - \pi(a_1,a_1)\right) + \varepsilon\left(R - \pi(a_2,a_1)\right)$$

ah. can make ≠! cos 1 state..!

$\pi(a_1,a_2)$? $\pi(a_1,a_1)$ not necessarily ! eg here $\pi(a_1,a_2)$ bad, $\pi(a_1,a_1)$ good



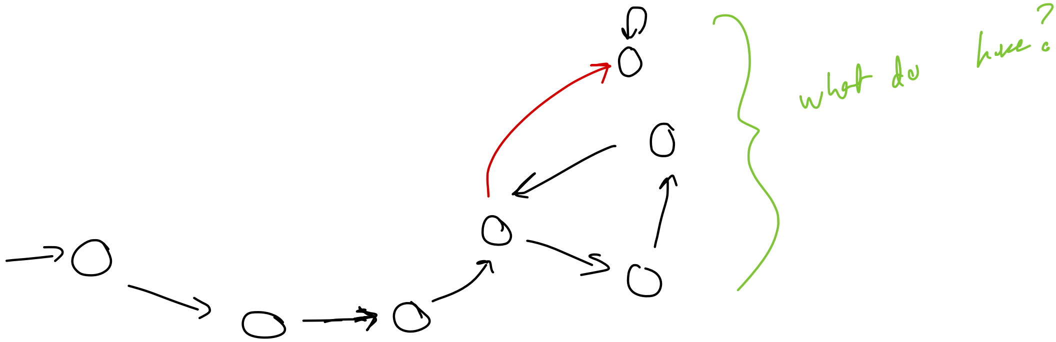
only thing that matters is $1-p$ transitions ! the rest can just be made arbitrarily small.

ahhh ok. first need to replicate things exactly until we get to an ergodic subgraph?

that's almost exactly the proof that I hated so much on.

PLAN

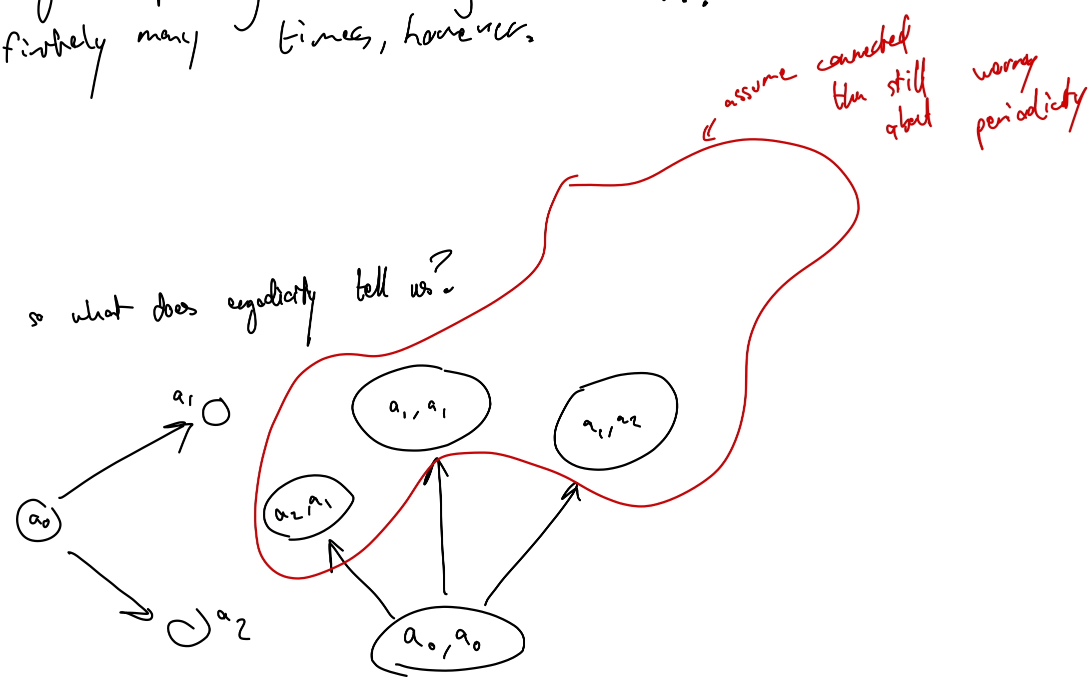
1. prove it using that weak assumption
2. try to fix the p thing



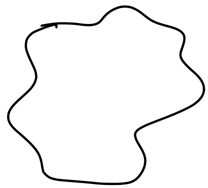
okok yeah the condition $\lim_{p \to 0} v_s(s) = R$ condition tells exactly and only about the ergodic state?

oh sure yes, but remember that is not the same as just following the big transitions!!! → can do that finitely many times, however.

ok so what does ergodicity tell us?

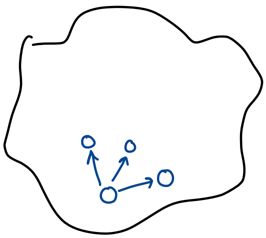


Suppose we get to an ergodic subgraph. (only in-edges)



then happy!!

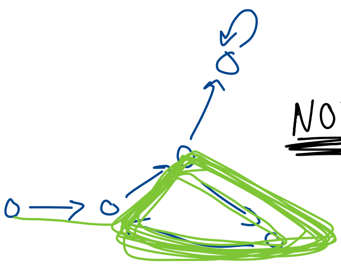
suppose we get to an isolated component (only in-edges)



then will have the same time-average distribution because there is a unique stationary distribution?

yeah...

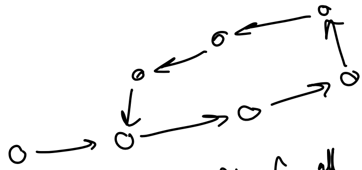
ok so do we always end up in SCC?



NOT obvious.

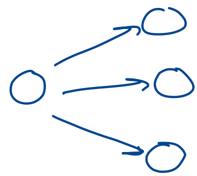
might even be false.

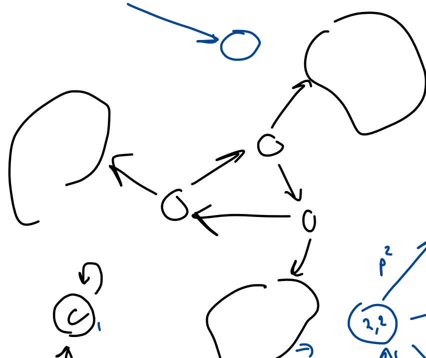
can we get this graph?
remember: playing against itself!

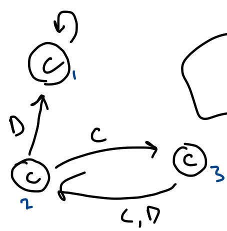


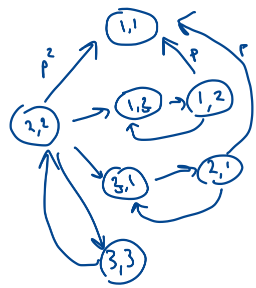
suppose is false for all nodes.

Then this is obviously NOT a stable state at all.









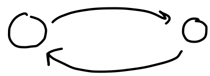
hmm so this is very problematic...

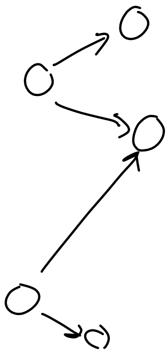
even worse:

oh ok I'm convinced this works again,
but it is very annoying.

OK, so core issue?

s_1 gets information from s_2.

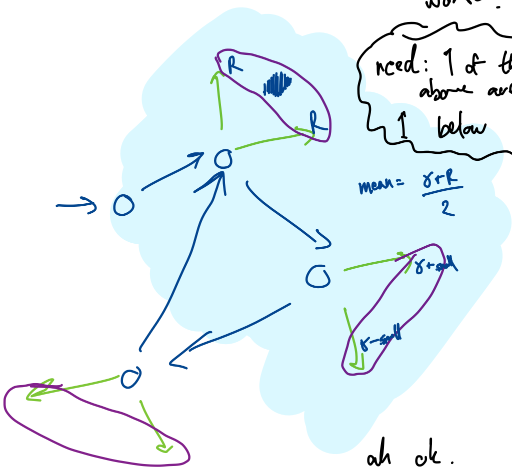




ahhh.
if loop on top level, then loop again.
so will get some for both

combine best w worst.

need: 1 of them above average, 1 below R.

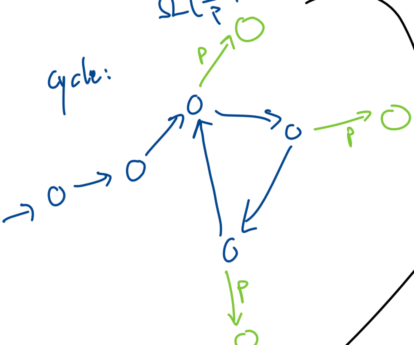


ah ok. so just find one that is $\geq$ mean.

OH ok so this:

follow $1 - \frac{1}{\Omega(\frac{1}{p})}$ until get ~~to~~

cycle:

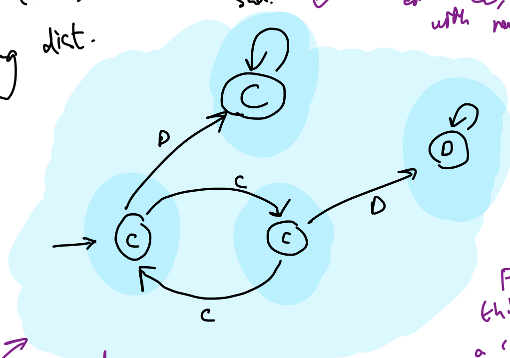


finite number, one of them has the ~~worst~~ worst limiting dist.

fuck this is so ugly

what happens here? obvious that will get to either CC, CD, DC, DD with roughly equal probability.

sad.

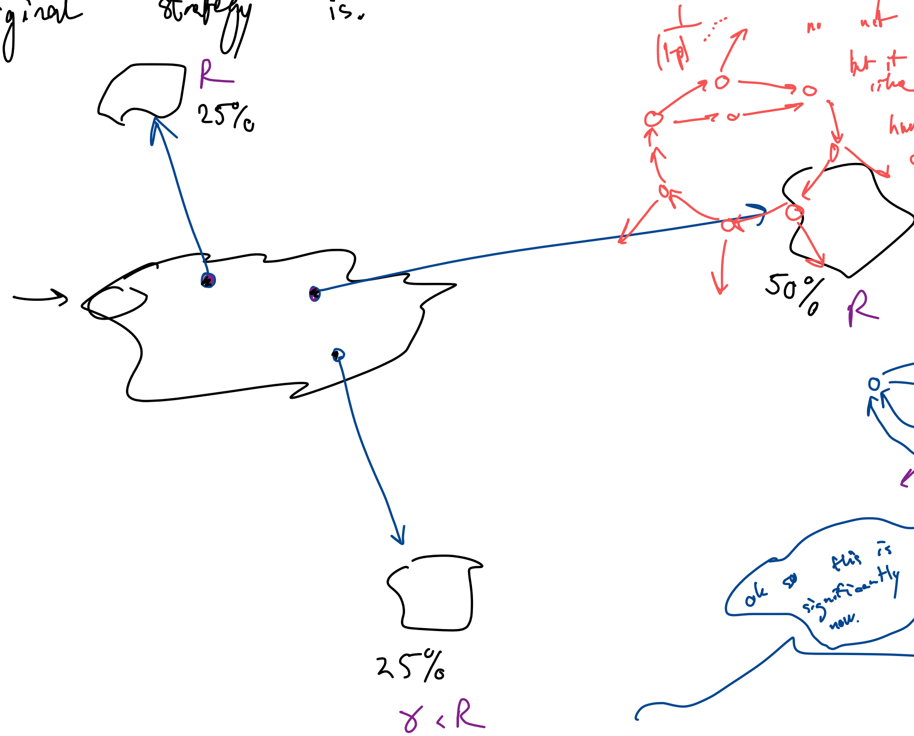


maybe need to/do TFT here? recursion?

FUCK. this is indeed a counter example

what is clear about this? if can choose 50% exact probability for small p which is cool

OK. … so the problem is my idea of the invading strategy to completely determine where the original strategy is.



or………
compare $s_1 \to s_2$ with $s_1 \to s_2$

hmm does this help? possibly yes

oh maybe it does…..

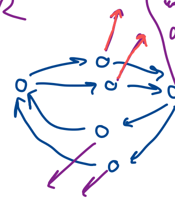
no not really…..

but it creates something like it!

hmm. there are …

cool.

ok cool. wait. yeah does not really do what you want.



ok so this is significantly harder now.

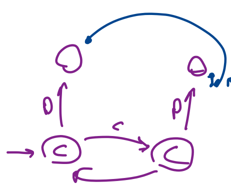
wait……… isn't this exactly what I wanted? this way it is possible to construct

high prob

low prob

annoying problem: interference

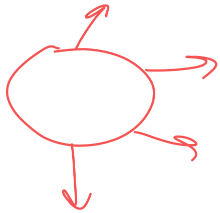
gates for arbitrary prob.



same construct: then:

needs to not be worse.

IDEA



put these in sequence

oh wow I think that works

for $s_2 - s_2$: 99% advances

1% revert to s_1

$s_2 - s_1$: 50% advances

50% revert to s

can then identify s_2 with itself

ALWAYS

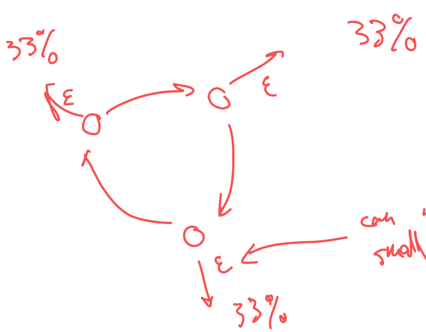
(but I think this might work on Pavlov too. It is pretty cool though)

ok so only problem is if s_1 locks into a bad connected component while we're doing this.

→ same problem as before

I think Pavlov can be invaded here

but this feels like magic and kinda sketchy......



can make extremely small by chaining $p \nearrow$ together

OK doesn't work......

but a nice thing to come out of this:

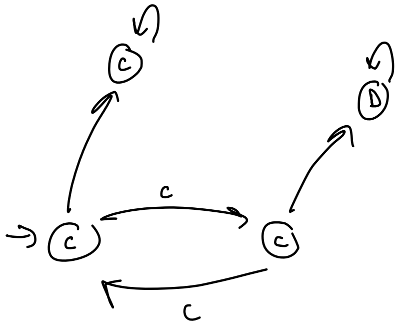
we can probably model probabilistic finite automata is simply normal ones with added

cycles

not obvious, but would be very cool if it worked.

PLAN:

* proceed as before.
* investigate:
    * copy everything, replace just in connected component that's bad
    * I think that's my only idea....

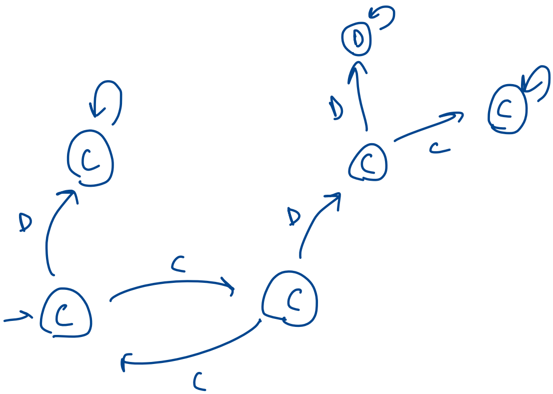


how invade this?

this one is really interesting, because it's just playing two strategies against each other

can I use that somehow?

or fake news!



how invade D ?

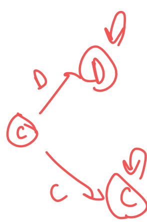
ok I'm pretty sure this works.

$\boxed{\text{how generalize???}}$

$$(1-\varepsilon)\left(v_{s_2}(s_1) - v_{s_1}(s_1)\right) + \varepsilon\left(v_{s_2}(s_2) - v_{s_1}(s_2)\right)$$

need to be equal to all D.

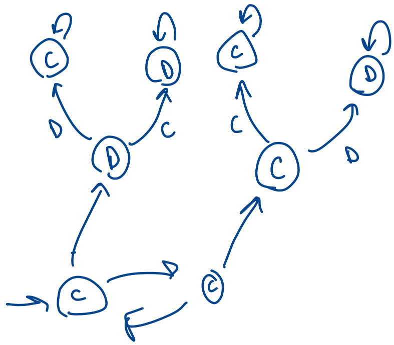
better against itself.



kinda works, but kinda lame.

or actually very reasonable....

two nice strategies bad against each other?



if feels like one can reduce the multistrategy scenario to one with multiple SCCs so they

oh. not obvious at all, because we can't guarantee starting times align.

look ~~only~~ at SCCs!

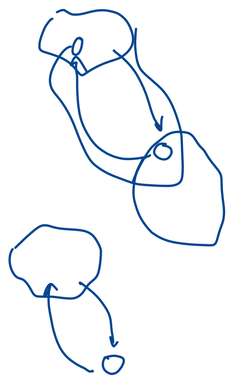
note: when run very long, will always end up in an SCC.

if only modify SCCs themselves, then we know relative pairing of SCCs always the same, which is nice, also always start in the same place of each SCC ahhhh nice.

ok. so thing is, given a p, can find an N s.t. w prob really high, we can identify in which state we are.

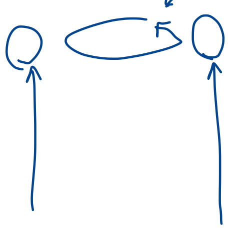
theoretically, with prob as close to 1 as we want, we should be able to determine where we are. eh. not sure.

so, idea:



key property: once in SCC, can make any number of moves without screwing up anything, because has unique limiting distribution

oh might contribute bad things.



IDEA:

replicate everything, except SCCs.
inside each SCC, can determine
which SCC opponent is in.
(wait...... noise, anyone?)

not obvious, and
ugly.

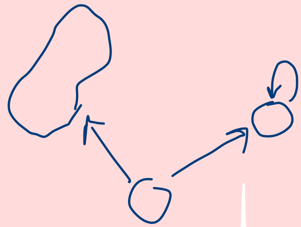
not obvious.
oh, obvious as $p \to 0$.

(almost) endless sequence of deterministic Cs?
nah?

oh, we can identify,
with probability

oh. can we just do original
idea, after getting to SCC?

$\boxed{\text{oh. I think so.}}$



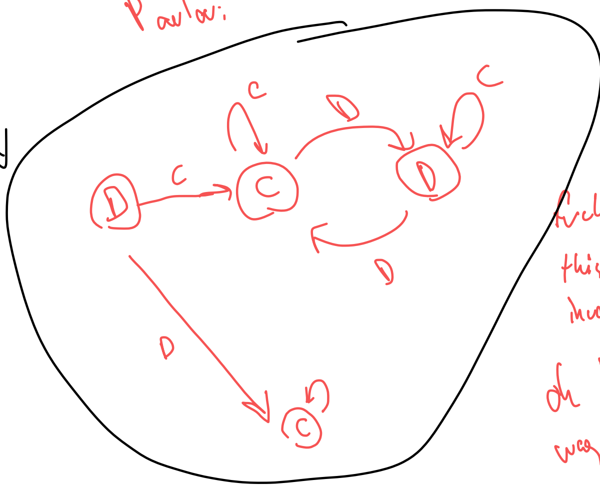
noooooo
how do you know
what to expect??
need to identify
which one it
is!

which seems almost
impossible.

take this further?

~~THIS CRUSHES PAVLOV~~
~~[illegible] and obviously~~

Pavlov:



Arrrrrrrr.
this will indeed
invade Pavlov.

oh this what
was I calculated
previously!

~~NOPE. TFT Pavlov is better than TFT~~

OK (i'm fairly certain), THIS WORKS:

although, luckily, it is suddenly more interesting.

aughhh this is is super ugly now....

replicate s_1 up to SCCs.

for every pair of SCCs in s_1, find a string b s.t. they produce different output if deterministic. then, run over all SCC^2 bs in s_2, thereby determining which SCC s_1 is in. then, we also know where in the SCC s_1 is, which we then can use for our original strategy, which then chooses the SCC copy.

what happens when we run s_1 against s_2? will be in some state with high probability, thus reach  !

this relies on: SCCs have unique limiting dist, it is possible to talk about SCCs, in SCC, starting position does not matter.

can I write this concisely?

do I really need epsilon??
yes! need it fixed

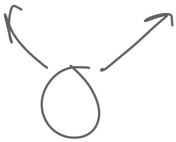
$v_{s_1}(s_1) - \gamma$.

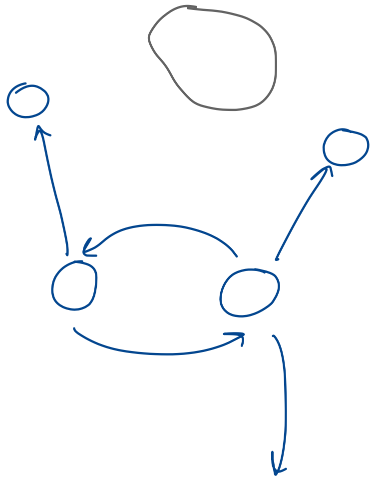
$p_{c_1} \to c_1$.

s_2:

* find action s_1 would take at finite time T

*





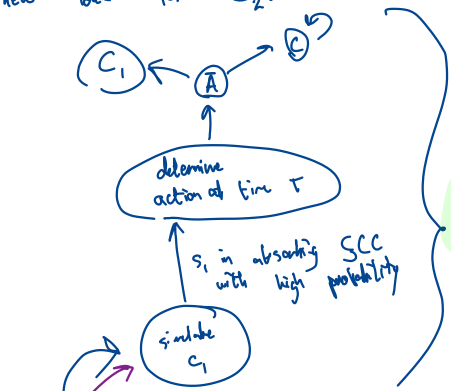
oh wait, a simpler proof?

first, check if you're against s_1 or s_2. then, either transition to a copy of s_1 or to c?. Thus, we then only add 1 new absorbing SCC.

nope nope nope. not necessarily same probabilities p_c then!

$\frac{R+\gamma}{2}$.

new idea for C_2:



schematic

yeah, works nicely

problem: s_1 stuck in non-absorbing SCC.

non-obvious how to do this.

take C_1:

key claims:

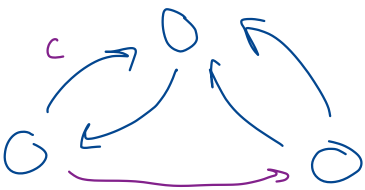
→ when s_1 is in absorbing SCC, can do whatever you want

→ before that, if completely simulate s_1, then all is good with the world.

one full cycle of the SCC: prob $p' > 0$ of exiting.
time to exit = geometric r.v.

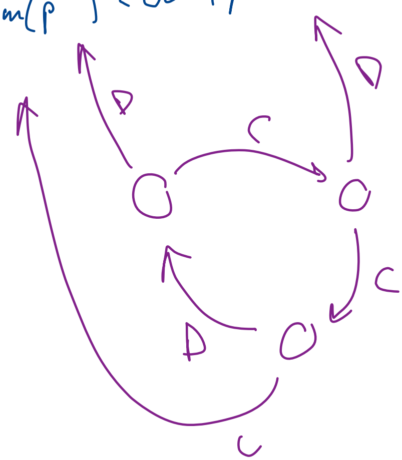
oh wait. can do prob $\left(\frac{1}{2}\right)^N$ not sure can do better. oh — require N consecutive mistakes of exiting → p'^2

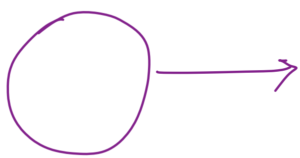
$P(C_2 \text{ exit before } s_1) = p'^2$ happens before p'
$= \text{Geom}(p'^2) < \text{Geom}(p')$



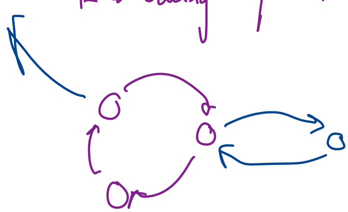
replace every state D

$P(\text{bad}) = \text{Geom}(p'^2) < \text{Geom}(p')$





oh what if they do interlocking cycles



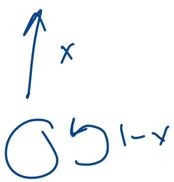
$P(\text{Geom}(x^2) < \text{Geom}(x)) \stackrel{?}{=} O(x)$

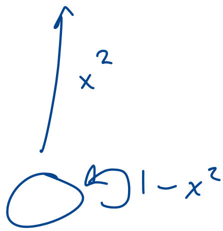
then: happy!

when $x \neq 1-O(p) \Rightarrow x = O(p)$

$\dfrac{\frac{1}{2}}{\frac{1}{2}+\frac{1}{2}\cdot p}$ $\quad \dfrac{1}{1+p} \quad \dfrac{p}{1+p}$

ok: to exit first,





possible problem: s_2 takes ∞ steps until getting to pivot state. not necessarily follow itself!!! ...
in fact: s_2 only advances when it makes a mistake.

SO: key problem: recognize itself again

even bigger problem: clone of s_2 might end up in an exact copy of an absorbing SCC of s_1.

(list all nonabs SCCs N times)

$\frac{1}{2}$, $\frac{1}{2}\cdot x$, $\frac{1}{2}\cdot(1-x)$

lose win go back

all cycles can be broken by just make sure that they are !!! (no zero.)

ah, yes, my friend. good work my friend.

OVERALL IDEA

* simulate s_1 until the real s_1 falls into absorbing SCC with prob $1-O(p)$.

→ now, s_1 will always converge to the same thing!

* determine what action s_1 will take at time T.
* self identify.

how do we make sure s_2 & s_2 are in the same state??

ok then s_2 is only a finite

A steps ahead of s_2-clone.

but don't know how many hmmmm.

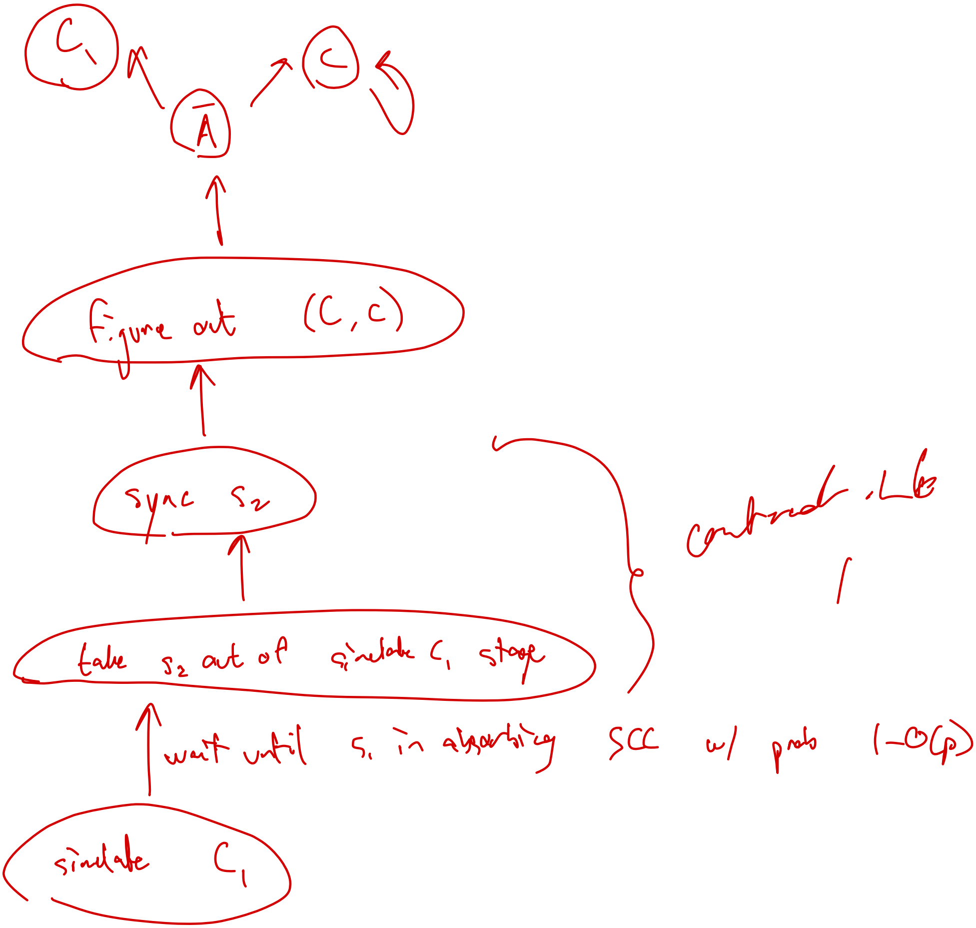
ahhh

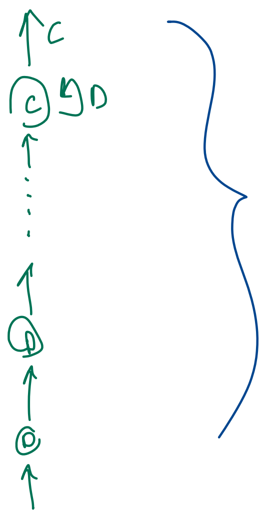
LOL this is too fucking intricate but it is actually kinda cool.

then: ① determine s_1's A at time T > max possible gap

② transition into ~~long~~ chain of $\overline{A}$

③ profit :)





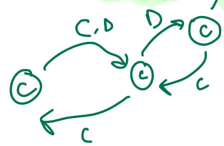
waiting for s_2 to catch up ☺

need to know what cycle it is

problem! the virtual version is not necessarily synced!

ahhh, depending on state of D we're in!

oh ok, so compare Geom(x) and Geom(px)



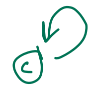
idea:

take H:

what if lags behind?

ahh. nice. I think this is prob 1? nop.

could we do something simpler? nah this is a simple idea, just hard to write down



a messy idea is worth 100% if works, 0% if not works.

$\boxed{\text{FML don't panic.}}$

replace every node by a copy of a massive graph D. The output at every state in D is dictated by the state it replaces, and ~~transitions~~ between copies of D also.

there should be a name for this

Kronecker product? yup. cool

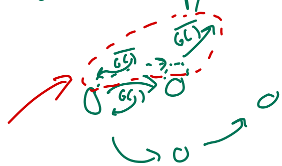
<u>within D:</u> (this is the hard part)

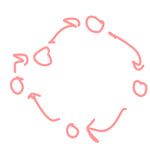
take all transient SCCs as list. (N of them)

repeat list N times. also for every start state in the SCC, repeated M times,

each is a double copy of the SCC

NEED EXAMPLE





<u>proof</u>: s_1 will take path $\underbrace{s_1 \to s_2 \to \dots \to s_K}_{SCCs}$ ← absorbing

then,

potential lemma: if X, Y absorbing $\Rightarrow X \times Y$ absorbing

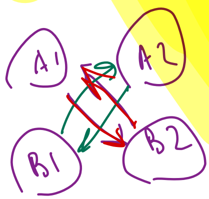
we could probably do something else.

just close the most advantageous one

Fix: we can make s_1 end up where we want it to end up in the strongly connected component

don't panic.

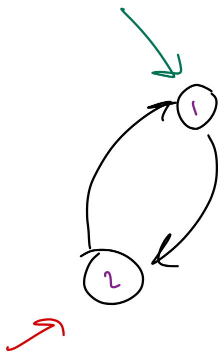
uhh. rather define bijection between (H, K) and S?

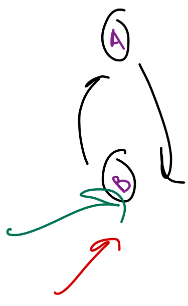


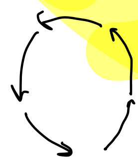
fuck

?SCCs? yeahhh fuck does this ruin my whole proof?

i think it might.......







$$\pi = \sum_{S \in \Omega} P(s_1 - s_2 \to S) \cdot \pi^{(S)}$$

we can't just choose the best.

↳ need most similar to γ!

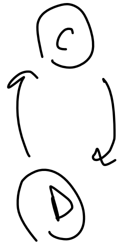
$E[\text{fraction in } (c_1, c_2)] =$

$= E[\text{fraction}$

possible lemma: each (H, h), (H', h') pair maps to one ascc.

R+P
T+S

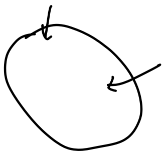
who knows.



oh we can choose the best.
but additional complications.

ok, we could also potentially recreate probabilities.

→ feels very hard.



so, choose the best one it is.

prediction: fixing it takes 2 hours

OK, so new plan

think thru!

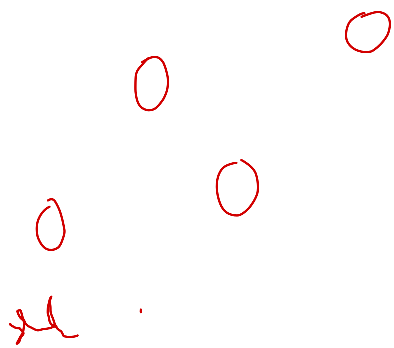
lemma 4.13:

either $V_{s_2}(s_1) = R \pm O(p)$

or $V_{s_2}(s_1) \geq \gamma \pm O(p)$

$\lim_{p \to 0} V_{s_1}(s_2) < R$

ez to prove



R R R R P R R R R P

TS TS TS TSS1TTT

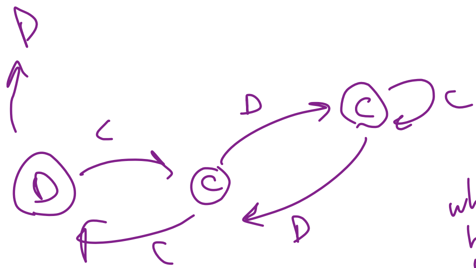
need $V_{s_1}(s_2) < R$!

I really do think there's a way to fix this.

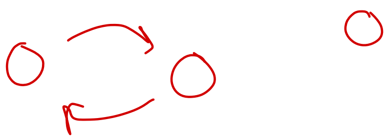
could we just

ahh. NICE SOLUTION

problem:



what happens here likso



evaluate in which cycle it is in, before leaving.

so like:

suppose this SCC.
evaluate if currently is a cycle
have a grace period

coffee, and pondering life, it is...

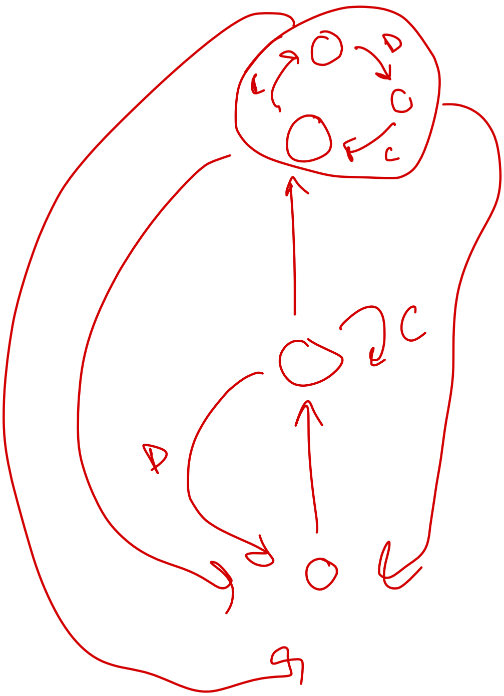
im a genius ← a TRUE genius.

↙ need care about where im cycle.

s_1 has some cycles: $\{ C, \; CDC, \; DDC, \ldots \}$

isn't this crazy

do:



pumping lemma <3<3<3

↳ then I can refer to cipser!

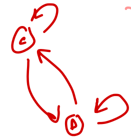
OK, so drastically simpler way:

take long word that just never appears.

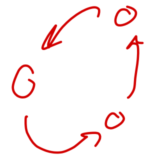
but! we'll get to it with prob $\sim p^N$.

so eventually we'll be there.

so: from every transient SCC
there is a way out using
at most n edges.
so if input is exactly those, then
$\to$ out, so $P(\text{out}) \geq p^n$
for a fixed n for each SCC.



can generate every string



or, even simpler: there is a way of getting to some absorbing comp in n steps, so
$P(\text{absorb in} \leq n \text{ steps}) \geq p^n$.

n = diameter of graph.

true of every time step
$\Rightarrow \text{geometric}(p^n)$

need: $\exists$ word it cannot recognize without an error

construct a word of length
$N > n \Rightarrow \text{geometric}(p^N)$.

↳ just repeat that N times.

use pumping lemma... !!!

ignore

states 'cause absorb.

$x y^n z$

no $\ddot{n}$ what am I dealing with.

next: repeat every cycle twice, the second time making 1 mistake what will the payoffs be?

I think, exactly what I wrote there.

I need new notation for disregarding

$\underbrace{p(\ldots)}$ terms though

maybe use?

$\boxed{\text{can use } O(p) \text{ as } p \to 0}$

---

what do I need to prove?

→ smth smth SCCs

defined simply as strongly connected components!

→ we care only about SCCs which have no outgoing edges

unique time average distribution

↳ "absorbing SCCs"

→ every pair of ACCSCCs in s_1 and s_2 ⟺ every ACCSCCs of their Markov chain

$$V_{s_1}(s_2) = \sum_{\text{absorbing SCCs}} p_i \cdot V_i(s_1, s_2)$$

"obvious" by Markov chain thinking.

yeah, I think I need the time average mode of thinking

just skip the expectation thingy

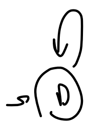
why can Pavlov not be invaded?

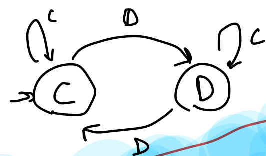
↳ good question. important question.

if better against itself ⇒ worse against Pavlov

Strategies:

* suppose invader existed
* any strat against Pavlov will get same or worse performance.
*



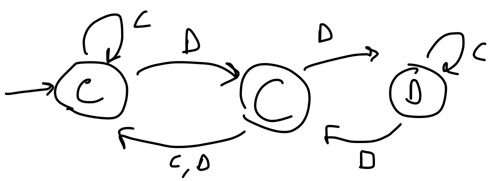


oh ahhh. that one fares worse against Pavlov!!!

possible invaders?

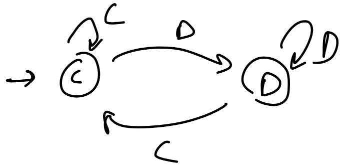
problems with Pavlov: spends too much time in D-D state

better?



Overwhelmingly seems like Pavlov can't be invaded by this....

I believe this one may be able to invade Pavlov.



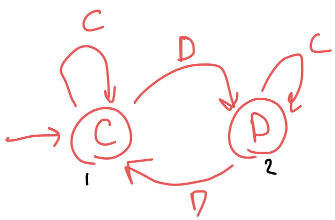
do I need $\boxed{2R > T+P}$??

that would be sad

Pavlov vs allD: D-D, D-C vs C-C. hmmm....

PLAN: tonight, try to find an ES strategy.
tomorrow, actually write down theorem 1, and state the existence question as an open problem.

OK, I believe Pavlov is ES. (assuming $2R > T+P$)

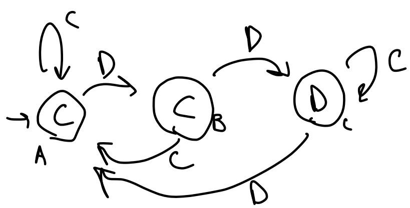


why would it be ES?

uhhhh smth optimal share of D state?

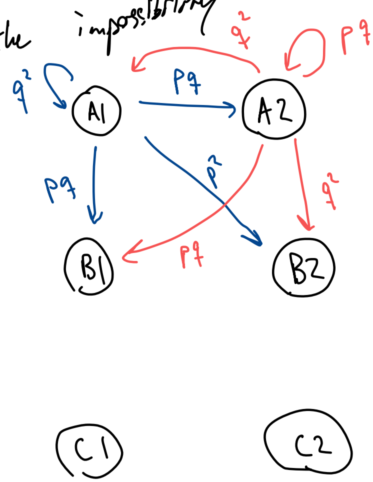
$$V_{s_2}(s_1) \overset{?}{>} V_{s_1}(s_1)$$

seems possible to prove of it.

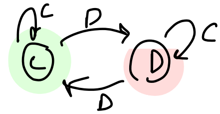


$q = 1-p$

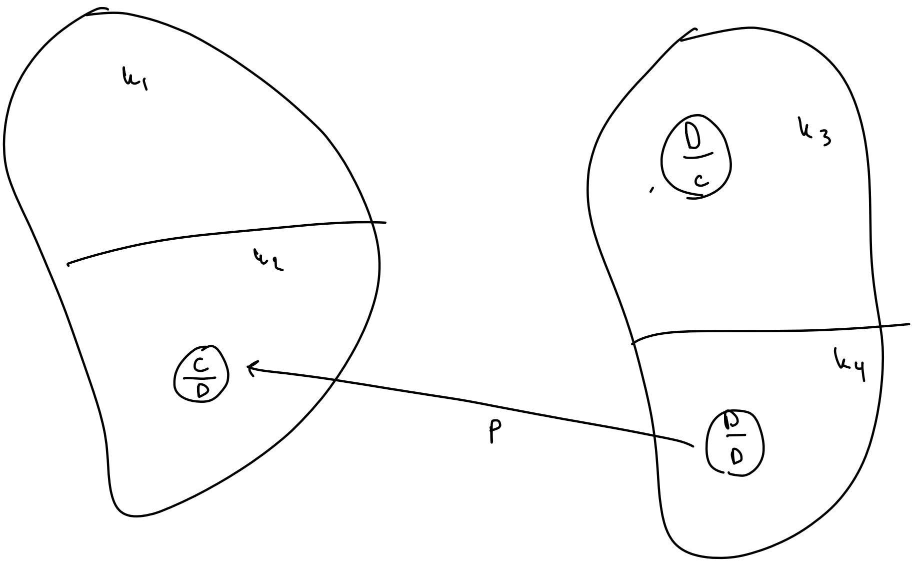
the impossibility



look at Pavlov-invader Markov.



Color states depending on Pavlov.

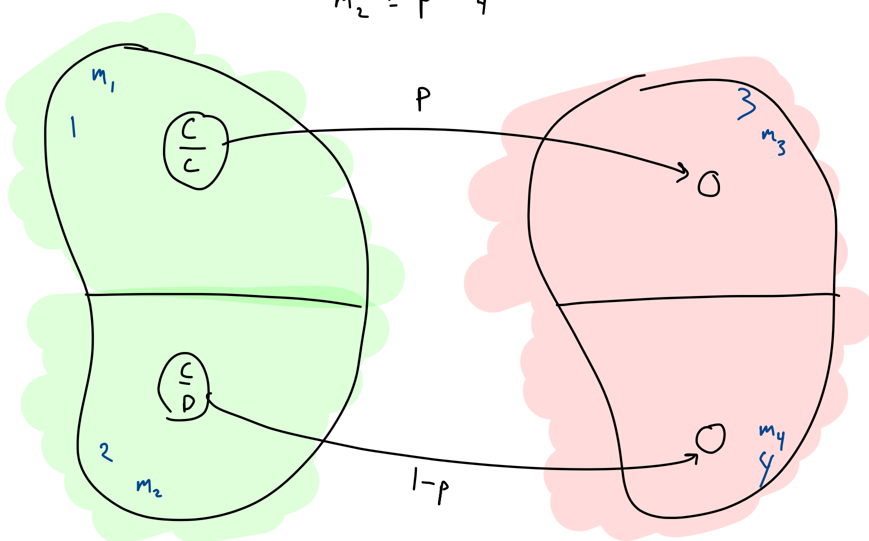


$m_3 \geq p \cdot m_1$.

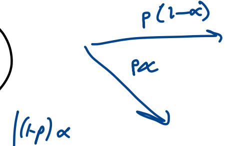
$m_4 \geq (1-p) \cdot m_2$

$\Rightarrow R \geq p m_1 + (1-p) m_2 \geq p m_1 + p(1-p) k_4$

$m_2 \geq p \cdot k_4$



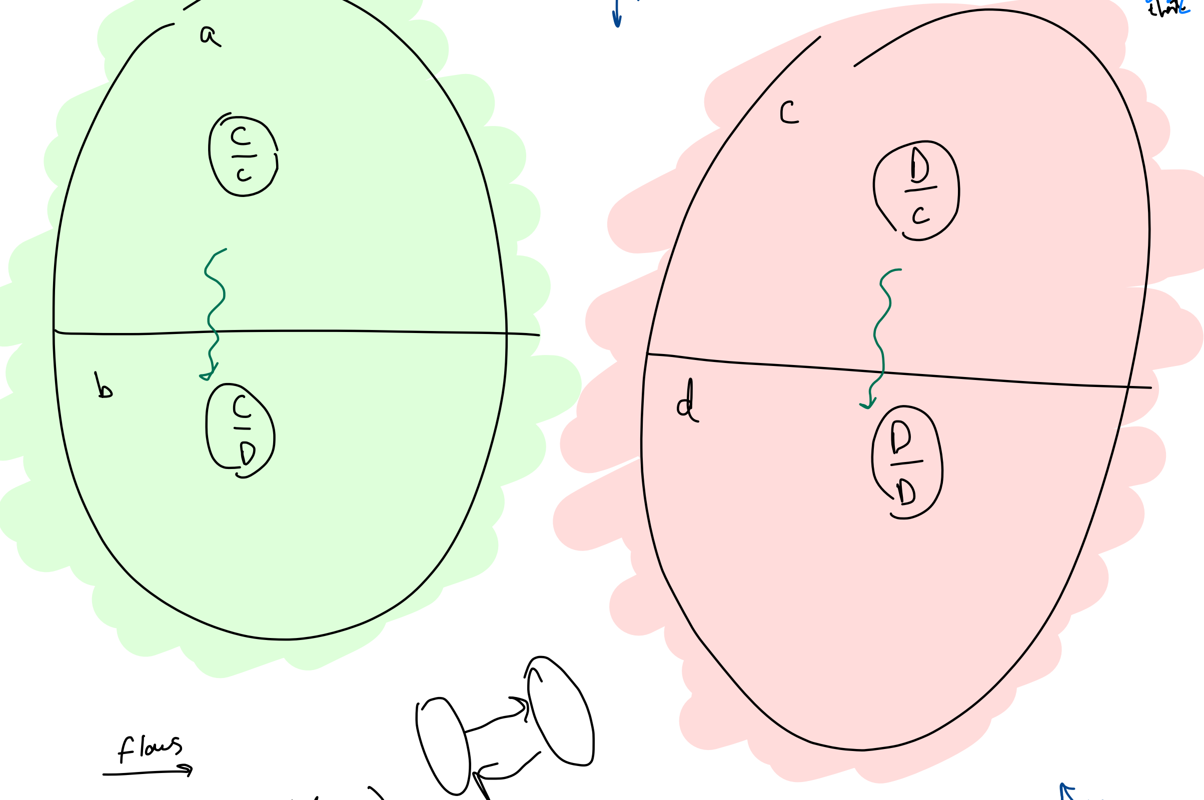
Pavlov invader

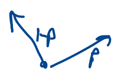


$c \geq \frac{p}{1-p} a$

$d \geq \frac{p}{1-p} b$

because





flows →

$= a \cdot p + b(1-p)$

$a+b+c+d = 1$

SO: $ap + b(1-p) = cp + d(1-p)$

flows ←

$= d \cdot (1-p) + cp$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ p & 1-p & -p & -1-p \\ & & & \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 0 \\ \\ \end{bmatrix}$$

want: maximize b, maximize $a+b$, $a+b+d$

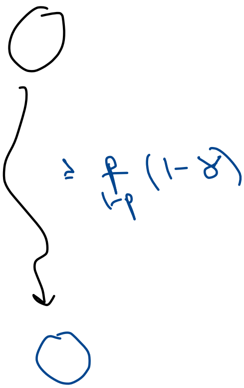
or rather:

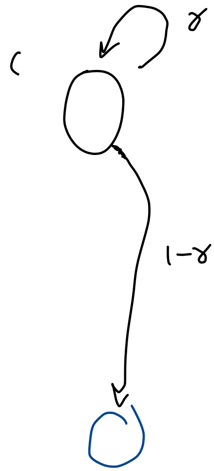
maximize $R \cdot a + T \cdot b + S \cdot c + P \cdot d$

may assume wlog $S=0$ (just shift all)

then: maximize $R \cdot a + T \cdot b + P \cdot d \leq Ra + Pd + T\left(\frac{c+d-ap}{1-p}\right)$

bound γ from below





$$x(\quad \dots) = 1$$

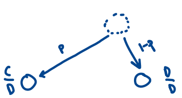
$$\alpha \geq \frac{p}{1-p}(1-\gamma)$$

$$x = \frac{1}{(\dots)}$$

$$1-\gamma \leq \frac{1-p}{p} \cdot \alpha$$

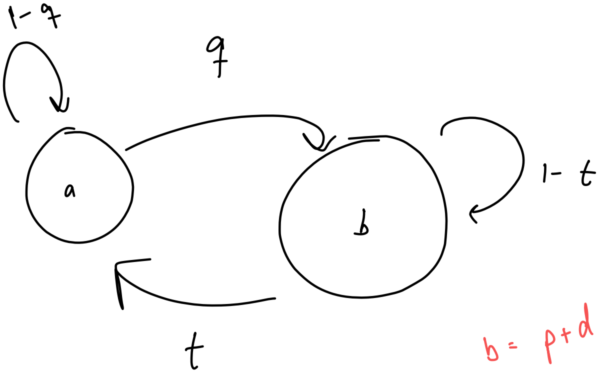
$$\gamma \geq 1 - \frac{1-p}{p} \cdot \alpha$$

$$d \cdot \frac{p}{1-p} \leq b$$



$$\underbrace{\text{entire input } \frac{D}{D}}_{= d} \cdot \frac{p}{1-p} \leq \text{entire input } \frac{C}{D}$$

makes sense. think about it!



$a + p + 2d \leq 2ap + 2dp + 1$

$p + d \geq 2ap + 2dp$

$a(1-2p) + d(2-2p) \leq 1-p$

$a \cdot 2p \leq d(1-2p) + p$

equality!!!

$d = \frac{(1-p) - a(1-2p)}{(2-2p)}$

$b = p + d - 2ap - 2dp$

max: $Ra + Tp + Td - 2Tap - 2Tdp + Pd$

$= Ra + (T+P)d + Tp - 2Tap - 2Tdp$

$a \cdot 2p(2-2p) \leq (1-p)(1-2p) - a(1-2p)^2 + p(2-2p)$

$a(4p - 4p^2 + 1 - 4p + 4p^2) \leq$

$\leq 1 - 3p + 2p^2 + 2p - 2p^2$

$= 1-p$

$a(R - 2Tp) + d(T + P - 2Tp) + Tp$

$qa - tb = 0$

$qa = tb$

$\frac{qa}{t} = b$

$T(1-p) + P$

given a, obviously want to maximize d.

suppose a is not maximal. then increase a by δ. ???

then, decrease d by $\delta \cdot \frac{(1-2p)}{(2-2p)} < \delta$

nice :)

want $\geq q$
$\leq t$

SO: want a maximal!

$c = 1 - a - b - d$

$a + b + c + d = 1$

$ap + b(1-p) = cp + d(1-p)$

$ap + b - bp = p - ap - bp - dp + d - dp$

$2ap + b = p + d - 2dp$

$2ap + 2dp + b = p + d$

maximize $Ra + Tb + Pd$

(1) $a+b+c+d=1$

(2) $ap+b(1-p)=cp+d(1-p)$

(3) $c \geq \frac{p}{1-p} \cdot a$ not tight!

(4) $b \geq \frac{p}{1-p} \cdot d$ not tight!

goal: maximize $R\cdot a + T\cdot b + S\cdot c + P\cdot d$

may assume w.l.o.g. $S=0$, so

$\max(Ra + Tb + Pd)$

(2) $\Rightarrow ap + b - \cancel{bp} = p - ap - \cancel{bp} - dp + d - dp \Rightarrow$

$\Rightarrow b = p + d - 2ap - 2dp$

insert: $Ra + T(p+d-2ap-2dp) + Pd =$

$= a(R-2Tp) + d(T+P-2Tp) + Tp$ ← can ignore constants.

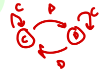
(4) $\Rightarrow p+d-2ap-2dp \geq \frac{p}{1-p} d \Rightarrow$

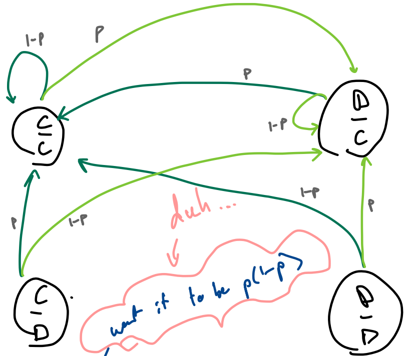
$\Rightarrow p+d-2ap-2dp-p^2-pd+2ap^2+2dp^2 \geq pd \Rightarrow$

$\Rightarrow a(-2p+2p^2)+d(1-4p+2p^2)+p-p^2 \geq 0$

$\Rightarrow a\cdot 2p\cdot(1-p) \leq d(2(1-p)^2-1)+p(1-p)$

what is it about $p(1-p)$?





$a = w_a\cdot(1-p) + w_b\cdot p + w_c\cdot p + w_d\cdot(1-p)$

$c = w_a\cdot p + w_b\cdot(1-p) + w_c\cdot(1-p) + w_d\cdot p$

$\frac{w_a}{1-p} + \frac{w_b+w_c}{p}\ldots$

key: w_a, w_c related

(3) $\Rightarrow 1-a-b-d \geq \frac{p}{1-p}\cdot a \Rightarrow 1-a-d \geq \frac{pa}{1-p}+p+d-2ap-2dp \Rightarrow$

$\Rightarrow a\left(1+\frac{p}{1-p}-2p\right)+d(2-2p) \leq 1-p$

w_a comes from: always c, sometimes c

fix a. then, want d as big as possible. so,

$d = \frac{(1-p)^2 - a(1-p+p-2p(1-p))}{2(1-p)^2} =$

$= \frac{1}{2} - a\left(\frac{1}{2}\cdot\frac{1}{(1-p)^2} - \frac{p}{1-p}\right)$

$w_a = a\cdot(\text{certain}_c + go_{cc}\cdot(1-p) + go_{DC}\cdot p)$

$w_c = c\cdot(\text{cer}_c + p\cdot go_{cc} + (1-p)\cdot go_{DC})$

too good to be true :(

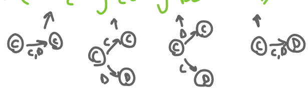
OH THIS WORKS

$\boxed{\text{IF can prove } c \geq p(1-p), \text{ then DONE}}$ (i think)

YAYYY

don't you dare not remember

$W_a = a\cdot(\overset{\cdot 1}{\text{cer}_c} + \overset{\cdot(1-p)}{go_{cc}} + \overset{\cdot p}{go_{DC}} + \overset{\cdot 0}{\text{cer}_D})$



$a-d = (1-2p)^3$:O

$(1-p)(2p-1)^2 + p(1-p)$ $\quad p(1-p)$

$p(2p-1)^2 + p(1-p)$

$p(1-p)$

$q(q-p)^2 + pq$ $\quad pq$

$p(q-p)^2 + pq$

pq

and so on

then minimize when all equal !!!

(which happens with $p(1-p)$!!!!!

that is, $\text{cer}_c + go_{cc} + go_{DC} + \text{cer}_D = 1$

can vary however you want,

but is minimized when $go_{cc} = 1$

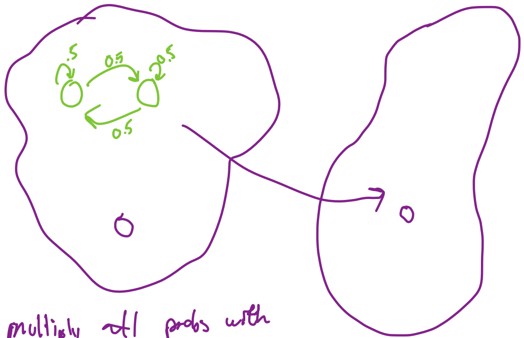
(problem: cer_D ....)

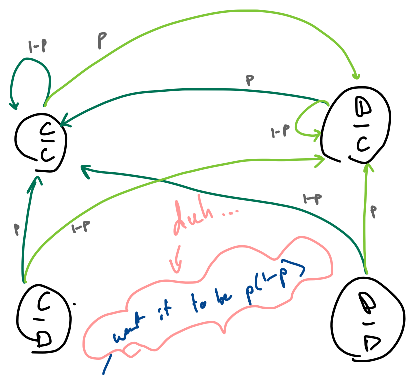
(relies on: $2p(1-p) \leq p^2 + (1-p)^2$

good old AM-GM ;))

$a = w_a \cdot (1-p) + w_b \cdot p + w_c \cdot p + w_d \cdot (1-p)$

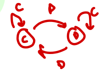
what is it about $p(1-p)$?





duh...

want it to be $p(1-p)$



$a = w_a \cdot (1-p) + w_b \cdot p + w_c \cdot p + w_d \cdot (1-p)$

$c = w_a \cdot p + w_b \cdot (1-p) + w_c \cdot (1-p) + w_d \cdot p$

$\frac{w_a}{1-p} + \frac{w_b + w_c}{p} \cdot \frac{w_d}{1-p}$

key: w_a, w_c related

multiply all probs with

$p - 2dp$

w_a comes from:

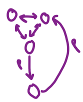
always c

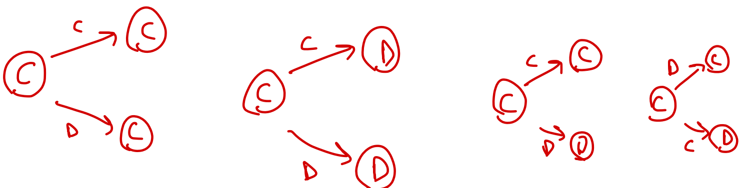
sometimes c

so,

$w_a = a \cdot (\text{certain}_c + go_{cc} \cdot (1-p) + go_{Dc} \cdot p)$

$w_c = c \cdot (cer_c + p \cdot go_{cc} + (1-p) \cdot go_{Dc}')$





fraction of mass in $\binom{C}{C}$

cer_c

cer_D

go_{cc}

go_{Dc}

too good to be true ¨

OH THIS WORKS (think)

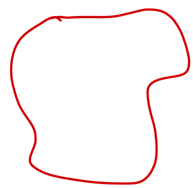
$w_a = a \cdot (1 \cdot cer_c + (1-p) \cdot go_{cc} + p \cdot go_{Dc} + 0 \cdot cer_D)$

$w_c = c \cdot (1 \cdot cer_c' + p \cdot go_{cc}' + (1-p) \cdot go_{Dc}' + 0 \cdot cer_D')$

fraction of mass in $\binom{D}{C}$

$cer_c + cer_D + go_{cc} + go_{Dc} = 1.$

C states



$C = w_b (1-p) + w_d p +$

$+ a(p \cdot cer_c + p(1-p) go_{cc} + p^2 go_{Dc})$

$+ c((1-p) cer_c' + p(1-p) go_{cc}' + (1-p)^2 go_{Dc}')$

suppose for a sec that we don't allow cer_D

$a = w_b \cdot p + w_d \cdot (1-p)$

$+ a((1-p) cer_c + (1-p)^2 go_{cc} + p(1-p) go_{Dc}) + c(p cer_c' + p^2 go_{cc}' + p(1-p) go_{Dc}')$

key question: is the fraction of the mass the same?

$a + p + 2d \leq 2ap + 2dp + 1$

$p + d \geq 2ap + 2dp$

$\to a(1-2p) + d(2-2p) \leq 1-p$

$\to a \cdot 2p \leq d(1-2p) + p$

equality!!!

$d = \frac{(1-p) - a(1-2p)}{(2-2p)}$

$a \cdot 2p(2-2p) \leq$

$(1-p)(1-2p) - a(1-2p)^2 + p(2$

$-2Tdp$

$a(4p - 4p^2 + 1 - 4p + 4p^2) \leq$

$) + Tp$

$\leq 1 - 3p + 2p^2 + 2p$

$= 1-p$

$+p$

$c = \frac{p}{1-p} a$

conditions are

$\Rightarrow \frac{pa}{1-p} + a + p + 2d \leq 2ap + 2dp + 1$

$\Downarrow$

$\cancel{pa} + a - \cancel{a}p + p - p^2 + 2d - 2dp \leq 2ap + 2dp + 1 - 2ap^2 - 2dp^2 - p$

$a(1 - 2p + 2p^2) + d(2 - 4p + 2p^2) \leq 1 - 2p + p^2$

$a((1-p)^2 + p^2) + 2d(1-p)^2 \leq (1-p)^2$

again, given a, want to maximize d

so, will want equality!

that is, $d = \frac{1}{2} - a\left(\frac{1}{2} + \frac{p^2}{2(1-p)^2}\right)$

$a = 1-p$?

$2d \leq 2p - 2p^2 + 2dp$

$d \leq p$

$\boxed{a = 1-p}$

$\boxed{d = p}$

$\Rightarrow d(1-2p) \geq 2p - 2p^2 - p$
$\quad p - 2p^2$

$a \cdot 2p(1-p)^2 \cdot 2 \leq (1-2p)\left((1-p)^2 - ap^2 - a(1-p)^2\right) + 2p(1-p)^2$

$a\Big(\underbrace{4p(1-p)^2 + (1-2p)(1-p)^2}_{(1+2p)(1-p)^2} + (1-2p)p^2\Big) \leq \underbrace{(1-2p)(1-p)^2 + 2p(1-p)^2}_{(1-p)^2}$

$+(1-2p)p^2$

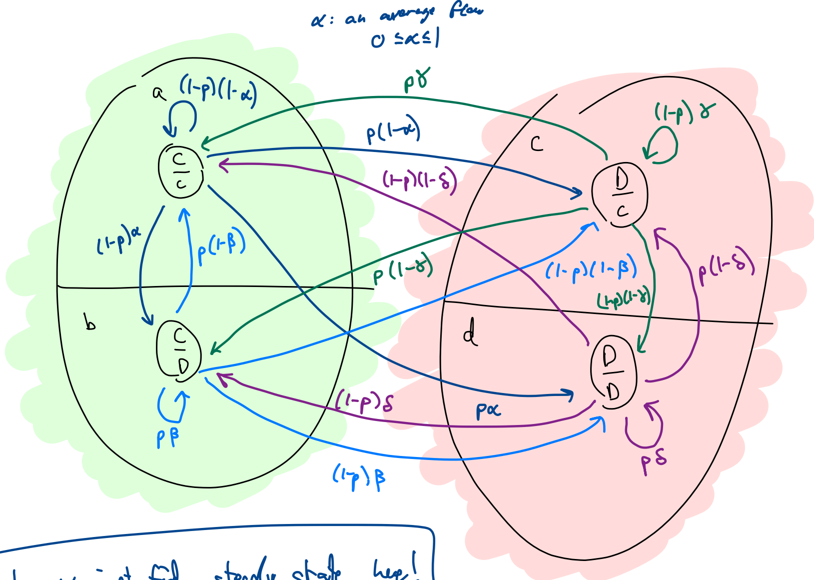
$1 + 2p - 2p - 4p^2 + p^2 + 2p^3 \quad + p^2 - 2p^3$

$a(1 - 2p^2) \quad \Rightarrow \quad a \leq \frac{(1-p)^2}{1-2p^2}$

$$a = \frac{1-2p+p^2}{1-2p^2} .$$

$$d = \frac{1}{2} - \frac{1}{2} \cdot \frac{((1-p)^2 + p^2)}{1-2p^2}$$

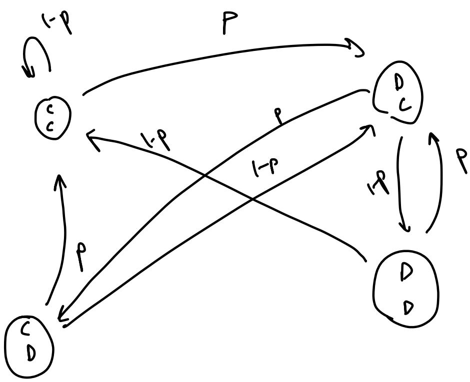
wow interesting. $d \cdot 2 - p^2$ is the real d!



unknown: 8

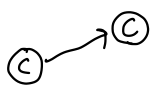
eqs: 5

ah wow. just find steady state here!



problem is: $\alpha, \beta, \gamma, \delta$ cannot all be 0.

but we can reduce α?



what does $\alpha = 0$ mean?

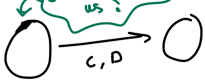
that from C, always go to C.

which is bad for us

in particular, $\alpha = 0 \Rightarrow \gamma = 1$

I suppose....

can these affect us?



can we formalize the relationship between α and γ?

$\begin{pmatrix} C \\ C \end{pmatrix} \xrightarrow{0}$ a C state $\iff$ $\begin{pmatrix} D \\ C \end{pmatrix} \xrightarrow{0}$ a C state

α = uhhh unclear

so the $\begin{pmatrix} D \\ C \end{pmatrix}$ weighing are always $\geq \frac{p}{1-p} \cdot \begin{pmatrix} C \\ C \end{pmatrix}$

$\begin{pmatrix} C \\ C \end{pmatrix} \xrightarrow{p}$ a C state $\iff$ $\begin{pmatrix} D \\ C \end{pmatrix} \xrightarrow{1-p}$ a C state

$\begin{pmatrix} C \\ C_C \end{pmatrix}$ $+$ $\quad \begin{pmatrix} C \\ C_D \end{pmatrix}$ $-$ $\quad \begin{pmatrix} D \\ C_C \end{pmatrix}$ $-$ $\quad \begin{pmatrix} D \\ C_D \end{pmatrix}$ $+$

$$\alpha = \varepsilon = \frac{\sum w_i \cdot o_i}{\sum w_i}$$

$1 - p \iff p$

$$\gamma = \frac{\sum w_{i'} \cdot \tilde{o}_i}{\sum w_{i'}}$$

$1 \iff 1$

$\rightarrow$

suppose $\alpha < p$. then certain weight in C_C state

$\sum w_i = W_{CC} + W_n \qquad \sum w_{i'} = W_{CC'} + W_{n'}$

$1-p$ $\qquad$ 1

$\frac{1}{2}$ $\qquad$ $\frac{1}{2}$

$\downarrow$ $\qquad$ $\downarrow$

p $\qquad$ 1

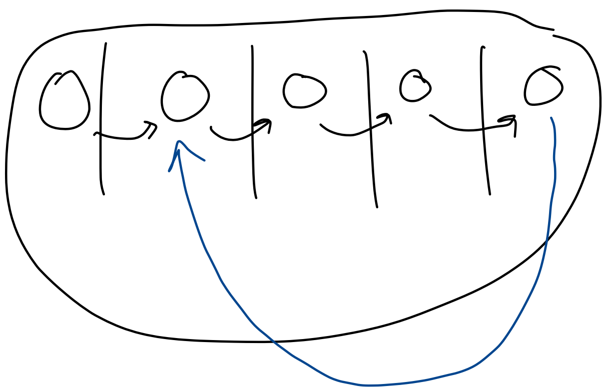
$$\frac{p^2}{(1-p)^2+p^2}$$

$$\frac{\frac{p}{1-p}}{\frac{1-p}{p}+\frac{p}{1-p}} = \frac{p^2}{(1-p)^2+p^2}$$

Fuck not strong enough bound.

$\frac{1-p}{p}x$ $\qquad$ $\frac{p}{1-p}x$

could bin based on distance to choice node:



lemma 4.1:

time average $\pi^{(a,b)}_{c_1,c_2} = E\left[\text{fraction of time in } (c_1,c_2) \mid \text{initial } (a,b)\right]$

$$\bigcirc \longrightarrow \bigcirc$$

$$M\pi \overset{?}{=} \pi$$

$$E\left[\text{fraction in } (c_1,c_2) \mid \text{initial } (a,b)\right] \overset{?}{=} \sum_{\substack{u \text{ s.t.} \\ (u,(c_1,c_2)) \in \text{Markov}}} p_{u\to(c_1,c_2)} \cdot E\left[\text{fraction in } u \mid \text{initial } (a,b)\right]$$

$$\overset{?}{=} E\left[\sum_u p_{u\to(c_1,c_2)} \cdot (\text{fraction in } u) \mid \text{initial } (a,b)\right]$$