Evolution of cooperation

in finite populations

Sabin Lessard
Université de Montréal

Contents

Thanks

Collaborators

        Véronique Ladret
        Philippe Lahaie
        Samuel Langevin
        David Lasalle-Ialongo
        Géraldine Martin

Competitors

        Martin Nowak
        François Rousset
        and many others

1. Examples of cooperation in nature

Cooperation is widespread...

but emergence of cooperation is problematic!

- A game-theoretic framework based on pairwise interactions is a first step toward a better understanding of a complex phenomenon.

- This framework can provide clues about favorable conditions for the evolution of cooperation.

2. Prisoner's Dilemma (PD)

Payoff Matrix

*C*ooperate	*R*eward	*S*ucker's payoff
*D*efect	*T*emptation	*P*unishment
against	*C*ooperate	*D*efect

$T > R > P > S$

C	$R = 5$	$S = 1$
D	$T = 14$	$P = 3$
against	C	D

Iterated Prisoner's Dilemma (IPD)

▶ PD repeated *n* times between the same players with additive payoffs

Tit-for-Tat (*A*)	$a = Rn$	$b = S + P(n-1)$
Always-Defect (*B*)	$c = T + P(n-1)$	$d = Pn$
against	*A*	*B*

$$a > c > d > b \text{ as soon as } n > \frac{T-P}{R-P}$$

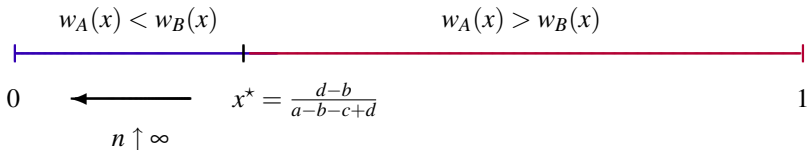| A	$a = 50$	$b = 28$
B	$c = 41$	$d = 30$
against	A	B

for $n = 10$ in the previous example

Expected payoffs in an infinite population

▶ random pairwise interactions

▶ x : frequency of A

$$w_A(x) = ax + b(1-x)$$

$$w_B(x) = cx + d(1-x)$$

$w_A(x) < w_B(x)$ $w_A(x) > w_B(x)$

0   ← $x^\star = \frac{d-b}{a-b-c+d}$ 1

      $n \uparrow \infty$

3. Evolutionary dynamics in an infinite population

- discrete, non-overlapping generations
- fitness : $1 + s \times$ payoff for some intensity of selection $s \geq 0$
- $x(t)$: frequency of A in offspring in generation t before selection

$$x(t+1) = \frac{x(t)(1 + sw_A(x(t)))}{1 + s\bar{w}(x(t))}$$

$$x(t+1) - x(t) = \frac{s(a - b - c + d)x(t)(1 - x(t))(x(t) - x^\star)}{1 + s\bar{w}(x(t))}$$

0	$x^\star$	1

4. Fixation probability in a finite population

- N parents chosen at random to produce the next generation (assumed large)

- $(\pi_1, ..., \pi_N)$: proportions of offspring produced in large numbers (assumed exchangeable; $\pi_i = N^{-1}$ for the Wright-Fisher model)

- $X(t)$: frequency of A in offspring in generation t before selection

In virtue of the ergodic theorem for Markov chains

$$X(T) \to X(\infty) = X(0) + \sum_{t \geq 0}(X(t+1) - X(t))$$

$X(\infty) = 1$ with probability $u(s)$, and 0 otherwise

- $u(s) = E_s[X(\infty)]$: probability of ultimate fixation of A

- $u(0) = X(0) = N^{-1}$

$$
\begin{aligned}
u(s) &= X(0) + \sum_{t \geq 0} E_s[X(t+1) - X(t)] \\
&= u(0) + s(a - b - c + d) \sum_{t \geq 0} E_s\left[\frac{X(t)(1-X(t))(X(t) - x^\star)}{1 + s\bar{w}(X(t))} \right] \\
&= u(0) + s(a - b - c + d) \sum_{t \geq 0} E_0[X(t)(1-X(t))(X(t) - x^\star)] + o(s)
\end{aligned}
$$

Definition: Weak selection favors A replacing B if $u(s) > u(0)$ for $s > 0$ small

$$
\boxed{x^\star < \frac{\sum_{t \geq 0} E_0[X(t)^2(1-X(t))]}{\sum_{t \geq 0} E_0[X(t)(1-X(t))]} = \hat{x}}
$$

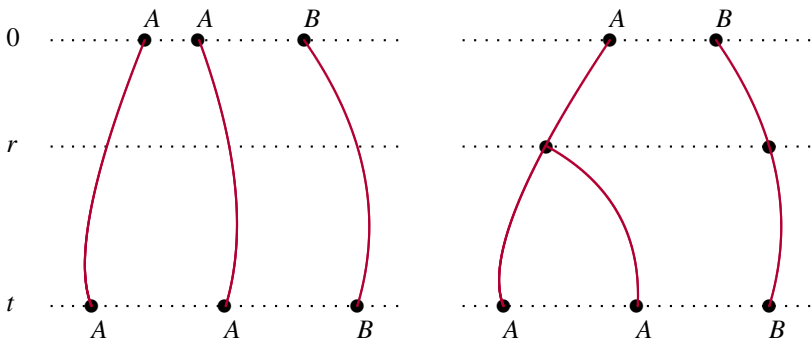# 5. One-third law of evolution

$$\sum_{t \geq 0} E_0[X(t)(1 - X(t))] = \sum_{t \geq 0} P_0(A, B \text{ in generation } t)$$

$$= \sum_{t \geq 0} p_{22}(t) P_0(A, B \text{ in generation } 0)$$

$$= \frac{X(0)(1 - X(0))}{1 - p_{22}}$$

with $p_{22}(t) = p_{22}^t$ the probability that two offspring chosen at random in generation t descend from two distinct ancestors in generation 0
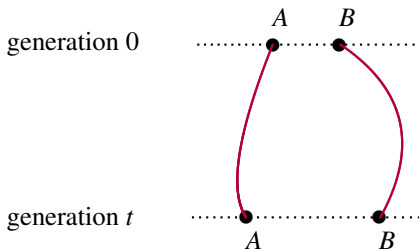
$$\sum_{t \geq 0} E_0[X(t)^2(1-X(t))] = \sum_{t \geq 0} P_0(A,A,B \text{ in generation } t)$$

$$= \sum_{t \geq 0} p_{33}(t) P_0(A,A,B \text{ in generation } 0)$$

$$+ \sum_{t \geq 0} \frac{p_{32}(t)}{3} P_0(A,B \text{ in generation } 0)$$

with $p_{33}(t) = p_{33}^t$ the probability that three offspring chosen at random in generation t descend from three distinct ancestors in generation 0, and

$$p_{32}(t) = \sum_{r=0}^{t-1} p_{33}^{t-r-1} p_{32} p_{22}^r = p_{32} \left(\frac{p_{33}^t - p_{22}^t}{p_{33} - p_{22}} \right)$$

the probability that they descend from two distinct ancestors in generation 0, from which

$$
\begin{aligned}
\sum_{t \geq 0} E_0[X(t)^2(1-X(t))] &= \frac{X(0)^2(1-X(0))}{1-p_{33}} + \frac{p_{32}X(0)(1-X(0))}{3(1-p_{22})(1-p_{33})} \\
&\approx \frac{p_{32}X(0)(1-X(0))}{3(1-p_{22})(1-p_{33})}
\end{aligned}
$$

for N large enough

For *N* large enough

$$\boxed{\hat{x} \approx \frac{p_{32}}{3(1-p_{33})} \leq \frac{1}{3}}$$

Definition (Nowak et al. *Nature* 2004): One-third law of evolution if $\hat{x} = 1/3$

$$w_A(1/3) > w_B(1/3)$$

0        $x^\star$ 1/3   1

Result (L. & Ladret JMB 2007): One-third law of evolution if and only if $p_{32} = 1 - p_{33}$, which means that at most 2 lineages out of 3 coalesce at a time backward in time with probability 1.

This characterizes the Kingman coalescent for a wide range of reproduction schemes as $N \to \infty$ with N generations as unit of time: then each pair of lineages coalesces at rate 1 independently of all others, so that

$$\lambda_{21} = 1, \lambda_{32} = 3, \lambda_{31} = 0,$$

and more generally

$$\lambda_{k,k-1} = \frac{k(k-1)}{2}$$

where λ_{kj} is the rate of death from k to j lineages backwards in time.

Skewed contributions of parents

- probability $1 - N^{-\alpha}$ that every parent produces the same proportion N^{-1} of offspring

- probability $N^{-\alpha}$ that one parent at random produces a proportion ψ of offspring and every other parent a proportion $(1-\psi)(N-1)^{-1}$

- in the case $\alpha < 1$ with N^{α} generations as unit of time as $N \to \infty$, $\lambda_{21} = \psi^2$, $\lambda_{31} = \psi^3$, $\lambda_{32} = 3\psi^2(1-\psi)$, according to a Λ-coalescent

$$\hat{x} \approx \frac{\lambda_{32}}{3(\lambda_{31}+\lambda_{32})} = \frac{1-\psi}{3-2\psi} < \frac{1}{3}$$
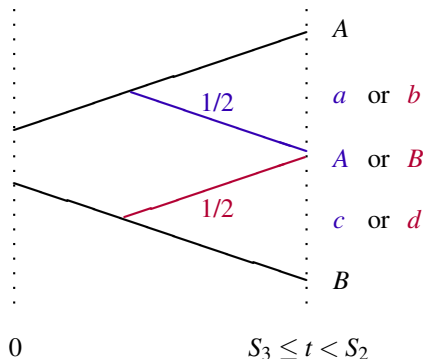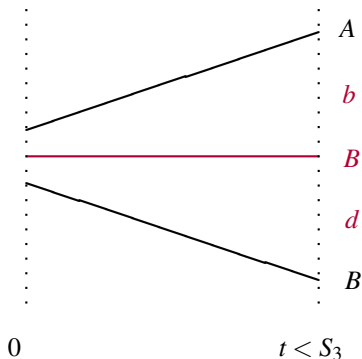
which means a more stringent condition for A to be favored.
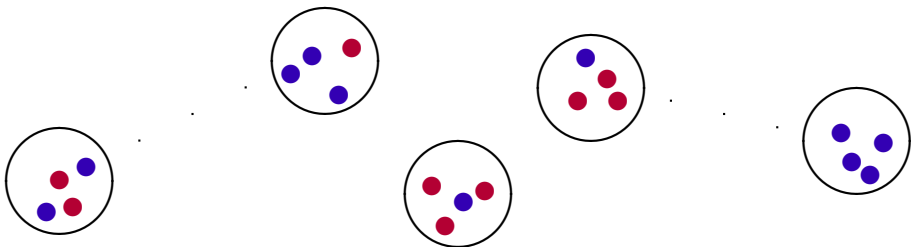
6. Projected average excess in payoff

Difference between the marginal payoff to *A* and the mean payoff to a competitor in all generations $t \geq 0$

$$\frac{u'(0)}{X(0)} \approx (b-d)E_0(S_3) + \left[\frac{a-c}{2} + \frac{b-d}{2}\right](E_0(S_2) - E_0(S_3))$$

with S_j for a time with j lineages, and one-third law if $E_0(S_2) = 3E_0(S_3)$

7. Group-structured population

- ▶ *D* groups of *N* parents producing equal proportions of offspring (*D* assumed large)

- ▶ *m* : proportion of offspring in each group that disperse uniformly (Wright's island model)

- ▶ selection among offspring within groups after dispersal

- $X_i(t)$: frequency of A in group i before selection in generation t

- $\overline{X(t)} = D^{-1} \sum_{i=1}^{D} X_i(t)$: frequency of A in the population in generation t

- $\overline{X(0)} = (ND)^{-1}$

In this case weak selection favors A replacing B if $x^\star < \hat{x}$

$$\hat{x} = \frac{\sum_{t \geq 0} E_0 \left[\overline{X(t)^2(1-X(t))} \right]}{\sum_{t \geq 0} E_0 \left[\overline{X(t)(1-X(t))} \right]}$$

$$= \frac{\sum_{t \geq 0} P_0(A,A,B \text{ in the same group in generation } t)}{\sum_{t \geq 0} P_0(A,B \text{ in the same group in generation } t)}$$

States for the ancestors of 3 offspring
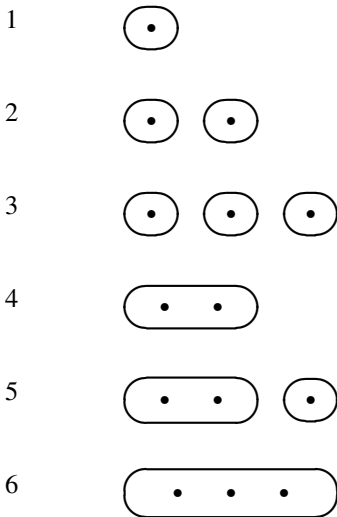
Transition matrix backward in time

Applying Möhle (1998) lemma

$$\mathbf{P}^{\lfloor NDf_{22}^{-1}\tau \rfloor} \rightarrow \left(\begin{array}{cc} e^{\tau \mathbf{G}} & 0 \\ \mathbf{F}e^{\tau \mathbf{G}} & 0 \end{array} \right)$$

as $D \rightarrow \infty$, where

$$\mathbf{F} = \left(\begin{array}{ccc} f_{21} & f_{22} & 0 \\ 0 & f_{21} & f_{22} \\ f_{31} & f_{32} & f_{33} \end{array} \right)$$

with f_{nk} the probability for n offspring in the same group *after dispersal* to have k ancestors in different groups in the case of an infinite number of groups,

$$\mathbf{G} = \left(\begin{array}{ccc} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 3 & -3 \end{array} \right)$$

the generator for the Kingman coalescent in a well-mixed population.

Two-time-scale argument for large D

▶ Fast scattering of lineages in the same group; slow collecting of lineages in different groups.

▶ Times spent with lineages in the same group can be neglected compared to times spent with lineages in different groups.

▶ The expected time spent with 2 lineages in different groups before coalescence in number of generations is approximately NDf_{22}^{-1}.

$$\sum_{t \geq 0} P_0(A, B \text{ in the same group in generation } t)$$

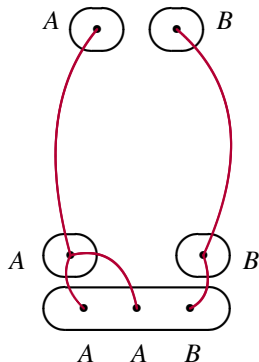$$\approx \sum_{t \geq 0} P_{42}(t) P_0(A, B \text{ in different groups in generation } 0)$$

$$\approx f_{22} \times ND f_{22}^{-1} \times \overline{X(0)}(1 - \overline{X(0)})$$
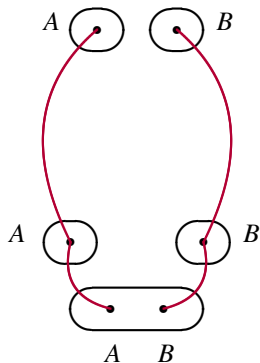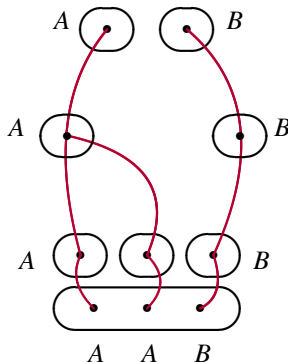

state 2 in generation 0

state 4 in generation t

A B

$$\sum_{t \geq 0} P_0(A, A, B \text{ in the same group in generation } t)$$

$$\approx \frac{1}{3} \sum_{t \geq 0} P_{62}(t) P_0(A, B \text{ in different groups in generation } 0)$$

$$\approx \frac{1}{3} \times (f_{32} + f_{33}) \times NDf_{22}^{-1} \times \overline{X(0)}(1 - \overline{X(0)})$$

state 2 in generation 0

state 6 in generation t

$$\hat{x} \approx \frac{1-f_{31}}{3(1-f_{21})} > \frac{1}{3}$$

with

$$f_{21} = \frac{(1-m)^2}{Nm(2-m)+(1-m)^2}$$

$$f_{31} = f_{21}\left[\frac{N(1-m)+2(N-1)(1-m)^3}{N^2m(3-3m+m^2)+(3N-2)(1-m)^3}\right]$$

which means a less stringent condition for *A* to be favored in the case of a group-structured population.

8. Differential contributions of groups

► Wright island model with selection among offspring before dispersal

$$
\begin{aligned}
u'(0) &= (b-d)\sum_{t\geq 0} E_0\left[\overline{X(t)(1-X(t))}\right] \\
&+ (a-b-c+d)\sum_{t\geq 0} E_0\left[\overline{X(t)^2(1-X(t))}\right] \\
&+ m(2-m)(b+c-2d)\sum_{t\geq 0} E_0\left[\overline{X(t)^2}-\overline{X(t)}^2\right] \\
&+ m(2-m)(a-b-c+d)\sum_{t\geq 0} E_0\left[\overline{X(t)^3}-\overline{X(t)}\ \overline{X(t)^2}\right]
\end{aligned}
$$

$$E_0\left[\overline{X(t)^2} - \overline{X(t)}^2\right] \approx P_0(A, B \text{ in different groups})$$
$$- P_0(A, B \text{ in the same group})$$
$$E_0\left[\overline{X(t)^3} - \overline{X(t)}\ \overline{X(t)^2}\right] \approx P_0(A, A, B \text{ with } B \text{ in a different group})$$
$$- P_0(A, A, B \text{ in the same group})$$

$$\boxed{\hat{x} \approx \frac{1 - \tilde{f}_{31} - m(2-m)(\tilde{f}_{21} - \tilde{f}_{31})}{3(1 - \tilde{f}_{21})} + \frac{(a-d)(N-1)^{-1}}{(a-b-c+d)} > \frac{1 - f_{31}}{3(1 - f_{21})}}$$

where $\tilde{f}_{nk}$ is the probability for n offspring in the same group *before dispersal* to have k ancestors in different groups with an infinite number of groups. This means an even less stringent condition for A to be favored.

9. Skewed contributions of groups

- selection after dispersal
- probability $D^{-\beta}$ that a group at random provides a fraction χ of migrants and every other a fraction $(1-\chi)(D-1)^{-1}$ for $\beta < 1$

$$
\begin{aligned}
u'(0) \;\approx\; & \left[(b-d)f_{22} + \left(\frac{a-b-c+d}{3} \right) \left(f_{32} + f_{33} \frac{\lambda_{32}}{\lambda_{32}+\lambda_{31}} \right) \right] \\
\times \;\; & \lambda_{21}^{-1} D^{\beta} \times \overline{X(0)}(1-\overline{X(0)})
\end{aligned}
$$

where λ_{kj} is the rate of transition from k to j lineages in different groups backward in time with D^{β} generations as unit of time as $D \to \infty$

$$
\boxed{\; \hat{x} \approx \frac{1-f_{31}-f_{33}\frac{\lambda_{31}}{\lambda_{32}+\lambda_{31}}}{3(1-f_{21})} < \frac{1-f_{31}}{3(1-f_{21})} \;}
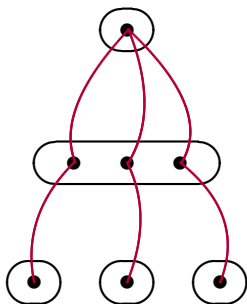$$

which means a more stringent condition for A to be favored.

Rates of transition for lineages in different groups

$$\lambda_{32} = 3(\lambda_{21} - \lambda_{31})$$

$$\lambda_{21} = (\chi m)^2 \tilde{f}_{21} = \left(\frac{\chi m}{1-m}\right)^2 f_{21}$$

$$\lambda_{31} = (\chi m)^3 \tilde{f}_{31} = \left(\frac{\chi m}{1-m}\right)^3 f_{31}$$

same ancestor in the case
of infinite number of groups

before dispersal

after dispersal

10. Summary and comments

- IPD in an infinite population predicts the increase in frequency of Tit-for-Tat (*A*) against Always-Defect (*B*), and therefore can explain the spread of cooperation, but only from a frequency $x > x^\star$.

- IPD in a finite population favors a rare mutant *A* replacing *B*, and therefore can explain the advantage of cooperation to go to fixation, but only under the condition $x^\star < \hat{x}$.

- In a large population, $\hat{x} = 1/3$ in the domain of the Kingman coalescent, but $\hat{x} < 1/3$ leading to a more stringent condition for cooperation to be favored if the contributions of parents in offspring are highly skewed.

- The first-order effect of selection on the probability of fixation is given by a projected average excess in payoff.

- In a group-structured population with uniform dispersal in the limit of a large number of groups, $\hat{x} > 1/3$ which means a less stringent condition for cooperation to be favored.

- The condition is weaker if dispersal occurs after selection rather than before selection, so that there are differential contributions of groups. On the other hand the condition is stronger if the contributions of groups in offspring are more skewed.

- The results obtained from the first-order effect of selection are ascertained only under very weak selection, actually as long as the intensity of selection is small compared to the intensity of the other evolutionary forces (drift, dispersal), but without constraints on the reproduction scheme.

Thanks!