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

Thanks

Collaborators

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

Competitors

Martin Nowak
François Rousset
and many others

CRSNG
NSERC

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

Cooperate	Reward	Sucker's payoff
Defect	Temptation	Punishment
against	Cooperate	Defect

$$T > R > P > S$$

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

$a > c > d > b$ 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

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

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

$$\boxed{w_A(x) > w_B(x) \text{ if and only if } x > \frac{d-b}{a-b-c+d} = x^\star \downarrow 0 \text{ as } n \uparrow \infty}$$

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

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

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

$$\boxed{x(t) \uparrow 1 \text{ if } x(0) > x^\star \text{ and } x(t) \downarrow 0 \text{ if } x(0) < x^\star}$$

4. Fixation probability in a finite population

- N parents chosen at random to produce the next generation (assumed large)
- $(\pi_1, \ldots, \pi_N)$: 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

4. Fixation probability in a finite population

- N parents chosen at random to produce the next generation (assumed large)
- $(\pi_1, \ldots, \pi_N)$: 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
- $u(0) = X(0) = N^{-1}$

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

- $u(s) = E_s[X(\infty)]$: probability of ultimate fixation of A
- $u(0) = X(0) = N^{-1}$

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

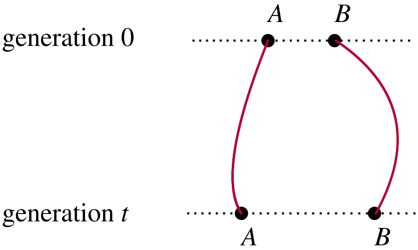
$u(s) > u(0)$ for $s > 0$ small enough: weak selection favors A replacing B

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

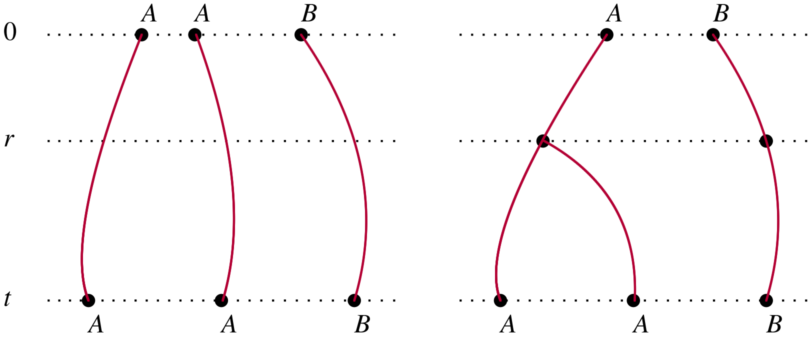
5. Generalized one-third law of evolution

$$\begin{aligned}\sum_{t\geq 0} E_0[X(t)(1-X(t))] &= \sum_{t\geq 0} P_0(A,B \text{ in generation } t) \\ &= \sum_{t\geq 0} p_{22}(t) P_0(A,B \text{ in generation } 0) \\ &= \frac{X(0)(1-X(0))}{1-p_{22}}\end{aligned}$$

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



$$\begin{aligned}
\sum_{t\geq 0} E_0[X(t)^2(1-X(t))] &= \sum_{t\geq 0} P_0(A,A,B \text{ in generation } t) \\
&= \sum_{t\geq 0} p_{33}(t) P_0(A,A,B \text{ in generation } 0) \\
&+ \sum_{t\geq 0} \frac{p_{32}(t)}{3} P_0(A,B \text{ in generation } 0)
\end{aligned}$$



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

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

with equality to $1/3$ (one-third law, Nowak *et al.* 2004), and then the weakest condition for cooperation to be favored, if and only if 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.

Λ-coalescent (e.g., Eldon and Wakeley 2006)

- 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 $(1-\psi)(N-1)^{-1}$

Λ-coalescent (e.g., Eldon and Wakeley 2006)

- 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 $(1-\psi)(N-1)^{-1}$

$$\boxed{\hat{x} \approx \frac{1-\psi}{3-2\psi} < \frac{1}{3} \text{ if } \alpha < 1}$$

which means a more stringent condition for cooperation to be favored if contributions of parents in offspring are highly skewed.

6. Projected average excess in payoff

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

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

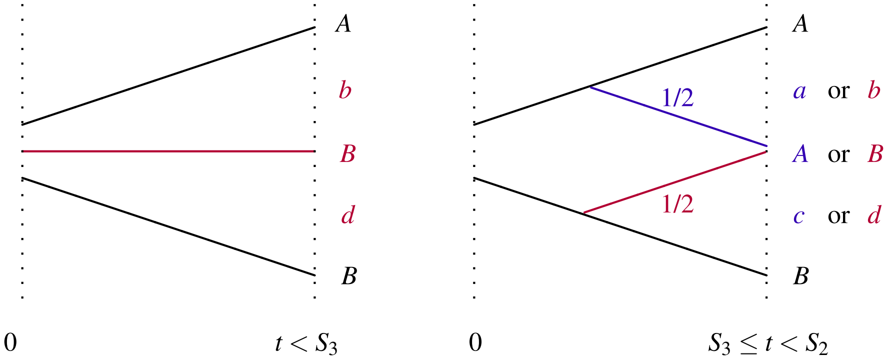
with S_j for a time with j lineages, and one-third law if $E_0(S_2) = 3E_0(S_3)$

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 within groups after dispersal

7. Group-structured population

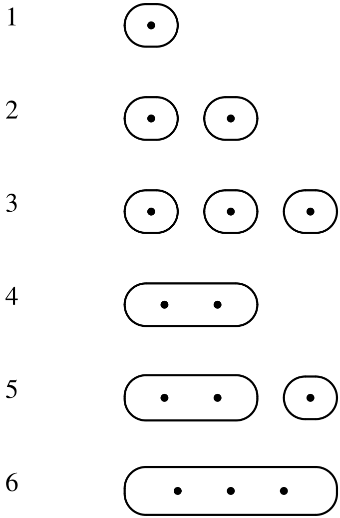
- D groups of N parents producing equal proportions of offspring (D assumed large)
- m : proportion of offspring in each group that disperse uniformly (Wright's island model)
- selection within groups after dispersal
- $X_i(t)$: frequency of A in group i before selection in generation t
- $\overline{X(t)} = D^{-1} \sum_{i=1}^{D} X_i(t)$: frequency of A in the population in generation t
- $\overline{X(0)} = (ND)^{-1}$

Then 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

$$\mathbf{P}^{\lfloor ND f_{22}^{-1} \tau \rfloor} \to \begin{pmatrix} e^{\tau \mathbf{G}} & 0 \\ \mathbf{F} e^{\tau \mathbf{G}} & 0 \end{pmatrix}$$

as $D \to \infty$, where

$$\mathbf{F} = \begin{pmatrix} f_{21} & f_{22} & 0 \\ 0 & f_{21} & f_{22} \\ f_{31} & f_{32} & f_{33} \end{pmatrix}$$

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

$$\mathbf{G} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 3 & -3 \end{pmatrix}$$

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

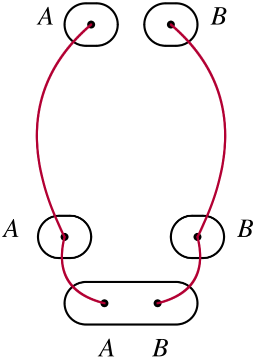
Two-time-scale argument for large D

- 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}.

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

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

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



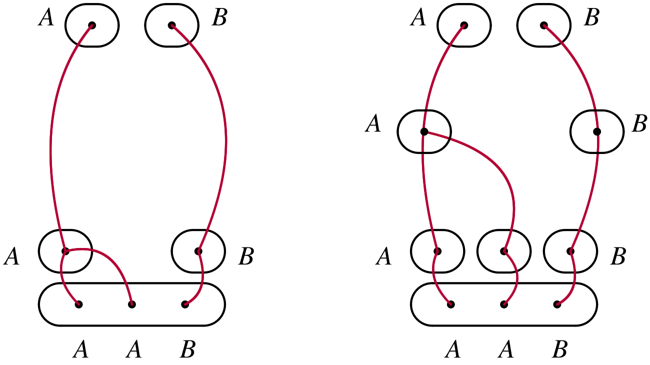
state 2 in generation 0

state 4 in generation t

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

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

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



state 2 in generation 0

state 6 in generation t

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

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 cooperation to be favored in the case of a group-structured population.

8. Differential contributions of groups

- Wright island model with selection before dispersal

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

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

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

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

which means an even less stringent condition for cooperation to be favored.

9. Exchangeable 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 $\beta < 1$

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

where μ_{lk} is the rate of transition from l to k lineages in different groups backward in time with ND^{β} generations as unit of time as $D \to \infty$

9. Exchangeable 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 $\beta < 1$

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

where μ_{lk} is the rate of transition from l to k lineages in different groups backward in time with ND^{β} generations as unit of time as $D \to \infty$

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

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

Rates of transition for lineages in different groups

$$\boxed{\mu_{lk} = \sum_{l \geq j \geq n \geq k-l+j \geq 1} \lambda_{lj} p_{jn} f_{n,k-l+j}}$$

where

$$\lambda_{lj} = \binom{l}{j} (\chi m)^j (1-\chi m)^{l-j}$$

$$p_{jn} = \frac{(N)_n S_j^{(n)}}{N^j}$$

$$f_{nk} = \sum_{j=0}^{k} \sum_{l=k-j}^{(n-1)\wedge(n-j)} \binom{n}{j} \frac{(Nm)^j (1-m)^{n-j} (N)_l S_{n-j}^{(l)}}{N^n - (1-m)^n (N)_n} f_{l,k-j}$$

with $(N)_n = N(N-1)\ldots(N-n+1)$, and

$$S_j^{(n)} = \frac{1}{n!} \sum_{\substack{j_1,\ldots,j_n \geq 1 \\ j_1+\ldots+j_n = j}} \binom{j}{j_1,\ldots,j_n}$$

Approximation for m small and N large

$$f_{nk} \approx \frac{M^k |S_n^k|}{M(M+1)\dots(M+n-1)}$$

where $M = 2Nm$ and

$$|S_n^k| = \text{coefficient of } x^k \text{ in } x(x+1)\dots(x+n-1)$$

in agreement with the Ewens sampling formula.

10. Diffusion approximation

- Wright island model with selection after dispersal and $s = \sigma(ND)^{-1}$
- ND generations as unit of time as $D \to \infty$

Applying Ethier and Nagylaki (1980) conditions, the infinitesimal mean and variance of the limiting diffusion are

$$\begin{aligned} m(x) &= \sigma(a-b-c+d)x(1-x)(xf_{33}-x^\star f_{22}+f_{21}-f_{31}) \\ v(x) &= x(1-x)f_{22} \end{aligned}$$

and the probability of fixation of A given a small initial frequency x_0

$$\int_0^{x_0} \exp\left\{-2\int_0^y \frac{m(x)}{v(x)}dx\right\}dy \approx x_0 + \sigma(a-b-c+d)x_0(1-x_0)f_{22}\left(\frac{1-f_{31}}{3f_{22}} - x^\star\right)$$

11. Summary and comments

- IPD in an infinite population can explain the spread of cooperation but only from a frequency $x > x^\star$.
- IPD in a finite population can explain that cooperation is favored to go to fixation from a low frequency under the condition $x^\star < \hat{x}$.
- In a large population, $\hat{x} \leq 1/3$ with $\hat{x} < 1/3$ leading to a more stringent condition if the contributions of parents in offspring are highly skewed.
- In a group-structured population with a large number of groups, the condition is less stringent with $\hat{x} > 1/3$ unless the contributions of groups in offspring are skewed enough.

- The first-order effect of selection on the probability of fixation is given by a projected average excess in payoff.
- 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, but without constraints on the reproduction scheme.
- An alternative approach under the assumption that the intensity of selection is of the same order of magnitude as the other evolutionary forces is a diffusion approximation, if the contributions of parents and groups in offspring are not too highly skewed.