Biotechnology Ethics
or
Towards a Generalized Theory of Justice
or
My cause is more important than your cause
Or
How independent unrelated causes become dependent depending upon who gets their way first

We will use as much of standard terminology from game theory as possible.
Let P_i denote the i-th player in the game.

We start off our model by assuming that there exist only two modes of action of player i upon player j, that is, P_i can only either kill or imprison P_j or P_i can leave P_j alone. We will later build on this model to incorporate cooperative (mutualistic) actions. In reality it is common for P_i to decide upon killing P_j based upon whom – which other players -- P_j thinks should be killed.

In subsequent models, we will show how to *amplify* the number of players greatly for *certain* players while keeping the number and identities of other players small and known, say, P_k = {me for $k=1$, Professor Roselli for $k=2$, Americans for $3 \le k \le 3 \cdot 10^8$, *Escherichia coli* bacteria for $3 \cdot 10^8 + 1 \le k \le 10^{10}$ }

Let P_1 = me, P_2 = an American female human soldier who murders and eats chickens and pigs, P_3 = a chicken, P_4 = a pig, P_5 = an *Escherichia coli* bacterium, P_6 = an Iraqi human who murders and eats chickens but not pigs who wants to suppress women, P_7 = an Iraqi human vegetarian who wants to free women, P_8 = any politician who supports the policies of George Bush. We form the matrix whereby player P_i in the i-th row kills or imprisons player P_j in the j-th column if there is a -1 in the (i, j)-th entry, and 0 otherwise.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	0	0	0	0	0	−1	0	−1
P_2	0	0	−1	−1	−1	−1	−1	0
P_3	0	0	0	0	−1	0	0	0
P_4	0	0	0	0	−1	0	0	0
P_5	−1	−1	−1	−1	0	−1	−1	−1
P_6	−1	−1	−1	0	−1	0	−1	−1
P_7	0	0	0	0	−1	0	0	−1
P_8	0	0	−1	−1	−1	−1	−1	0

If P_8 acts first, then we are left with

	P_1	P_2	P_8
P_1	0	0	−1
P_2	0	0	0
P_8	0	0	0

. So then if P_1 acts next we are left with

	P_1	P_2
P_1	0	0
P_2	0	0

, only two survivors.

Suppose instead P_1 acted first. Then

	P_1	P_2	P_3	P_4	P_5	P_7
P_1	0	0	0	0	0	0
P_2	0	0	−1	−1	−1	−1
P_3	0	0	0	0	−1	0
P_4	0	0	0	0	−1	0
P_5	−1	−1	−1	−1	0	−1
P_7	0	0	0	0	−1	0

. If P_2 gets her way next, then we take out P_3, P_4, P_5, P_7. This leaves us with

	P_1	P_2
P_1	0	0
P_2	0	0

again. However, suppose instead P_3 gets her or his way next after P_1. Then

	P_1	P_2	P_3	P_4	P_7
P_1	0	0	0	0	0
P_2	0	0	−1	−1	−1
P_3	0	0	0	0	0
P_4	0	0	0	0	0
P_7	0	0	0	0	0

.

The only one left to act is P_2. Again, if P_2 acts next, then

	P_1	P_2
P_1	0	0
P_2	0	0

will result. This is a stable situation. However, this is a stable situation which entailed a *non-minimum amount of killing*.

Suppose instead P_1 decides to kill or imprison P_2. I.e. Suppose we started with the matrix

$$\begin{array}{c|cccccccc} & P_1 & P_2 & P_3 & P_4 & P_5 & P_6 & P_7 & P_8 \\ \hline P_1 & 0 & -1 & 0 & 0 & 0 & -1 & 0 & -1 \\ P_2 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & 0 \\ P_3 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ P_4 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ P_5 & -1 & -1 & -1 & -1 & 0 & -1 & -1 & -1 \\ P_6 & -1 & -1 & -1 & 0 & -1 & 0 & -1 & -1 \\ P_7 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 \\ P_8 & 0 & 0 & -1 & -1 & -1 & -1 & -1 & 0 \end{array}$$

and that P_1 had his way first again.

Then

$$\begin{array}{c|ccccc} & P_1 & P_3 & P_4 & P_5 & P_7 \\ \hline P_1 & 0 & 0 & 0 & 0 & 0 \\ P_3 & 0 & 0 & 0 & -1 & 0 \\ P_4 & 0 & 0 & 0 & -1 & 0 \\ P_5 & -1 & -1 & -1 & 0 & -1 \\ P_7 & 0 & 0 & 0 & -1 & 0 \end{array}$$

. If now P_7 acts, we will eliminate P_5 and have

$$\begin{array}{c|cccc} & P_1 & P_3 & P_4 & P_7 \\ \hline P_1 & 0 & 0 & 0 & 0 \\ P_3 & 0 & 0 & 0 & 0 \\ P_4 & 0 & 0 & 0 & 0 \\ P_7 & 0 & 0 & 0 & 0 \end{array}$$

. Finally, we will arrive in a *stable* situation with *four* surviving individuals: the two vegetarians, the pig and the chicken… who *deserve* to be alive *precisely because by their choices they made the amount of killing minimal*.

How do we distinguish between "doing something" vs "not doing something"? Regardless of whether the "something" is bad or good to another player, "doing something" will *by definition* (hence, *always*) inflict a *negative* upon the do-er. It will always take *physical energy* to do so. Hence, the *degree of causality* will be proportional to the energy spent. Another way to say this: if the do-er suffers no negative effect, or practically no negative effect, then *they* cannot (and should not) take credit for nor be blamed for causing an action.

Example. A warden W of a prison puts a prisoner P into their prison unjustly for 20 years. Meanwhile, outside, activists A take time from their day protesting in cold and rain for P's freedom. After the first day of imprisonment, W signs the release for P.
It is wrong to say that W caused P's release and to take credit for delivering +20 years of life to P. W should not have put P into prison in the first place. Some wardens may arrogantly think *they* are the ones to being doing the prison this great "favor". It was the activists A who suffered –1 in terms of their unpaid time wasted who secured P's release. P owes A, not W.

Now consider quality of life. Still consider only 1 strategy of one player upon another.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	–0.3	–1.1	–1.5	–0.4	0	–2	0	–9
P_2	0.1	0.2	–5	–4	–1.4	–2.4	–3.3	0
P_3	0	0	0	0	–2	0	0	0
P_4	0	0	0	0	–2	0	0	0
P_5	–2.1	–1.9	–0.9	–0.8	0.2	–2	–2	–2.5
P_6	–1.5	–1	–1	–1.9	–1	–1.4	–1.3	1
P_7	–1	0	0	0	–1	0	0	–1
P_8	–0.8	–0.9	1	1	1	1	1	0

Let the number –2 be the cutoff point below which if P_j suffers more than –2, then the j-th column will be eliminated. The final stable matrix will be one in which all entries are greater than –2.
Let the energy sacrifice by P_1 be given by

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	0	–1.1	–1.5	–0.9	–0.8	–0.4	–0.2	–1

In fact, we will have *two* square matrices. One which records the *effect* P_i has on P_j. The second is the negative effect P_i has on itself by their actions upon P_j. We say P_i *causes* something to happen to P_j if there exists some *nonzero* negative effect upon P_i. In other words, P_i must make *some sacrifice*. In place of the entry of the matrix, why not create an XML link to describe the sacrifice and effect?

Each of these matrices themselves is something which player P_i thinks *is*.

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8
P_1	–0.3/–0.3	–1.1/–4	–1.5/–5	–0.4/–0.01	0/–0.04	–2/–0.9	0/–0.7	–9/–1
P_2	0.1	0.2	–5	–4	–1.4	–2.4	–3.3	0
P_3	0	0	0	0	–2	0	0	0
P_4	0	0	0	0	–2	0	0	0
P_5	–2.1	–1.9	–0.9	–0.8	0.2	–2	–2	–2.5
P_6	–1.5	–1	–1	–1.9	–1	–1.4	–1.3	1
P_7	–1	0	0	0	–1	0	0	–1
P_8	–0.8	–0.9	1	1	1	1	1	0

Example: War. We partition the players into two subsets in which no other alliances are made. We will assume a no-friendly-fire model whereby no A hurts an A and no B hurts a B.

	A_1	A_2	A_3	A_4	B_1	B_2	B_3	B_4
A_1	0	0	0	0	0	−2	−12	−9
A_2	0	0	0	0	−1	−2	−3	0
A_3	0	0	0	0	−20	−8	0	−15
A_4	0	0	0	0	−20	−18	0	−77
B_1	−2	−1	−44	−108	0	0	0	0
B_2	−1	−3	−71	−44	0	0	0	0
B_3	−66	−77	−99	−107	0	0	0	0
B_4	−71	−11	0	−5	0	0	0	0

.We will try to avoid making objective claims about cause and effect. Instead, we check for consistency. I.e. {If A_1 causes B_1 a certain amount of damage, then B_1 cannot do damage to A_2} implies that {If B_2 causes A_2 a certain amount of damage, then A_2 cannot do damage to B_3} by considering various levels of isomorphism.

If no two actions occur simultaneously, then there are (theoretically) 8! possible temporal orders in which the A s and B s can attempt to kill each other. In reality, because of the prior action of an A , some of the B s will be unable to carry out their actions, and vice versa. Let us see how this plays out with the game above.

Let us say that a kill is attained if the negative effect of an action is less than –10. Suppose A_1 acts first. Then A_1 merely injures B_2 and B_4 but kills B_3. Hence, B_3 has failed to do any damage to the A s – directly.

Theorem. At least *one* individual in a war has to suffer an injury or death without themselves having caused an injury or death first. Hence, an "innocent" individual must always suffer or die.

However, B_3 has managed to divert the energy of A_1. We are left with

	A_1	A_2	A_3	A_4	B_1	B_2	B_4
A_1	0	0	0	0	0	−2	−9
A_2	0	0	0	0	−1	−2	0
A_3	0	0	0	0	−20	−8	−15
A_4	0	0	0	0	−20	−18	−77
B_1	−2	−1	−44	−108	0	0	0
B_2	−1	−3	−71	−44	0	0	0
B_4	−71	−11	0	−5	0	0	0

If B_4 acts next, then B_4 will kill A_1 and A_2 but leave A_3 alone and A_4 only injured by –5 units. So A_2 never made a kill or injury. We are now left with $\begin{bmatrix} & A_3 & A_4 & B_1 & B_2 & B_4 \\ A_3 & 0 & 0 & -20 & -8 & -15 \\ A_4 & 0 & 0 & -20 & -18 & -77 \\ B_1 & -44 & -108 & 0 & 0 & 0 \\ B_2 & -71 & -44 & 0 & 0 & 0 \\ B_4 & 0 & -5 & 0 & 0 & 0 \end{bmatrix}$. Let A_3 act next. So A_3 kills B_1 and B_4. This leaves $\begin{bmatrix} & A_3 & A_4 & B_2 \\ A_3 & 0 & 0 & -8 \\ A_4 & 0 & 0 & -18 \\ B_2 & -71 & -44 & 0 \end{bmatrix}$. If B_2 acts next, then only B_2 remains.

Is-should. Two players. Suppose A thinks $\begin{bmatrix} & A & B \\ A & i_{A,A} & i_{A,B} \\ B & i_{B,A} & i_{B,B} \end{bmatrix}$. The letter i indicates that this is the way A thinks the world is. More specifically, these are the effects that A believes that A and B have upon each other with their single choice of strategies. Denote this matrix by $M_{A,i}$. We wish to create sequences of logical statements based upon such matrices. Namely, B: if A thinks $M_{A,i}$, then B thinks $M_{B,s} = \begin{bmatrix} & A & B \\ A & s_{A,A} & s_{A,B} \\ B & s_{B,A} & s_{B,B} \end{bmatrix}$. The matrix $M_{B,s}$ denotes the way B thinks the effects should be. We may denote this logical consequence as $M_{A,i} \Rightarrow M_{B,s}$.

How can we notate the concept that $M_{A,i} \Rightarrow M_{B,s}$ and $M_{C,i} \Rightarrow M_{D,s}$ are (roughly) equal? One way could be $(M_{A,i} \Rightarrow M_{B,s}) \Leftrightarrow (M_{C,i} \Rightarrow M_{D,s})$.

How can we relate the magnitudes of the entries of the matrices to these logical consequences? For example, if $i_{A,B} > i_0$ and $s_{A,B} > s_0$, then $j_{C,D} > j_0$ and $t_{A,B} > t_0$.

Example. Let A = a patient. Let B = doctors. Let C = me. The patient thinks that "doctors are arrogant because doctors don't tell them the medical reasons for what they do". So, this patient thinks the negative effect of B upon A is $i_{B,A}$ = "arrogance, not revealing information" for some selfish reason $i_{B,B}$ that benefits the doctor. This patient thinks $s_{B,A}$ = "doctors should tell them the medical reasons for what they do" even if it costs the doctor their reputation a bit $s_{B,B}$. I think "doctors are arrogant because doctors don't care about the cost of their services to the patient and do not reveal the total cost in writing to the patient before the service".