**The problem of suboptimization**

[**Principia Cybernetica Web - 021031**](http://pespmc1.vub.ac.be/)

|  |
| --- |
| **Optimizing the outcome for a subsystem will in general not optimize the outcome for the system as a whole. This intrinsic difficulty may degenerate into the "tragedy of the commons": the exhaustion of shared resources because of competition between the subsystems.** |

When you try to optimize the global outcome for a system consisting of distinct subsystems (e.g. maximizing the amount of prey hunted for a pack of wolves, or minimizing the total punishment for the system consisting of the two prisoners in the Prisoners game), you might try to do this by optimizing the result for each of the subsystems separately. This is called "suboptimization". The [principle of suboptimization](http://pespmc1.vub.ac.be/ASC/PRINCI_SUBOP.html) states that suboptimization in general does not lead to global optimization. Indeed, the optimization for each of the wolves separately is to let the others do the hunting, and then come to eat from their captures. Yet if all wolves would act like that, no prey would ever be captured and all wolves would starve. Similarly, the suboptimization for each of the prisoners separately is to betray the other one, but this leads to both of them being punished rather severely, whereas they might have escaped with a mild punishment if they had stayed silent.

The principle of suboptimization can be derived from the more basic systemic principle stating that "[the whole is more than the sum of its parts](http://pespmc1.vub.ac.be/RECSYSCO.html)". If the system (e.g. the wolf pack) would be a simple sum or "aggregate" of its parts, then the outcome for the system as a whole (total prey killed) would be a sum of the outcomes for the parts (prey killed by each wolf separately), but that is clearly not the case when there is interaction (and in particular cooperation) between the parts. Indeed, a pack of wolves together can kill animals (e.g. a moose or a deer), that are too big to be killed by any wolf in separation. Another way of expressing this aspect of "non-linearity" is to say that the interaction the different wolves are engaged in is a *non*[*zero-sum game*](http://pespmc1.vub.ac.be/ZESUGAM.html), that is, the sum of resources that can be gained is not constant, and depends on the specific interactions between the wolves.

As a last example, suppose you want to buy a new car, and you have the choice between a normal model, and a model with a catalyzer, that strongly reduces the poisonous substances in the exhaust. The model with catalyzer is definitely more expensive, but the advantage for you is minimal since the pollution from your exhaust is diffused in the air and you yourself will never be able to distinguish any effect on your health of pollution coming from your own car. Rational or optimizing decision-making from your part would lead you to buy the car without catalyzer. However, if everybody would make that choice, the total amount of pollution produced would have an effect on everybody's health, including your own, that will be very serious, and certainly worthy the relatively small investment of buying a catalyzer. The suboptimizing decision (no catalyzer) is inconsistent with the globally optimizing one (everybody a catalyzer). The reason is that there is interaction between the different subsystems (owners and their cars), since everybody inhales the pollutants produced by everybody. Hence, there is also an interaction between the decision problems of each of the subsystems, and the combination of the optimal decisions for each of the subproblems will be different from the optimal decision for the global problem.

The problem of suboptimization underlies most of the problems appearing in evolutionary [ethics](http://pespmc1.vub.ac.be/ETHICS.html). Indeed, ethics tries to achieve the "greatest good for the greatest number", but the greatest good (optimal outcome) for an individual is in general different from the greatest group for a system (e.g. society) of individuals.

Another, more dramatic implication of the problem of suboptimization is what Garrett Hardin has called the "[tragedy of the commons](http://www.dieoff.com/page95.htm)". The example is simple: imagine a group of shepherds who let their animals graze on a common pasture. Each animal that is added will bring additional profits to its shepherd. However, it will also diminish the overall profits of the group, since the grass eaten by that animal will no longer be available to the other animals. Yet, the loss of profit for the owner because of reduced grass will always be smaller than the gain because of an additional animal. Thus, for each individual shepherd, the optimal decision is to increase his herd. For the system consisting of all shepherds together, however, this strategy will result in an overgrazing of the pasture and the eventual exhaustion of the common resource.

**Reference:** Heylighen F. (1992): "Evolution, Selfishness and Cooperation", Journal of Ideas, Vol 2, # 4, pp 70-76.