erate stimuli to other units (or the system's "environment"), these stimuli being not necessarily responses to stimuli of the elementary component.

4. If conditions 1 and 3 are fulfilled, one can finally demand that the threshold for trial making be lowered, if the environment of the elementary component becomes energetically depleted.

By going carefully through this list of properties you may have spotted three important features. Number one, that these properties may well be attributed to a cortical neuron, if only some of its outstanding functional properties are taken into consideration. This should not come as a surprise, since we know that most of these two-legged admirable automata—in the Aristotelian sense—are equipped with about 10^{10} such elementary components. Number two, that condition 4 stipulates—if not explicitly—a certain "personal" or "miscroscopic" goal which the elementary component "seeks" to attain by its trial making activity. This goal is, of course, the maintenance of an energetically resourceful environment, in spite of the component's metabolic activity. As we shall see in a moment, this paradox is resolved by the component's ability to communicate with other elements in its environment. And, finally, number three, that I can be accused of building an automaton by using as elementary components automata. I am not going to refuse this argument; on the contrary, I wholeheartedly agree, with one reservation however, namely, that I have given the necessary and sufficient prescriptions to construct such elementary components. Indeed, there are many versions of electronic realizations of such components in existence today, from sizes of a couple of cubic inches down to about 2 cubic millimeters, and in costs ranging from $50.00 for very sophisticated devices down to a couple of cents apiece, barely fulfilling the points mentioned above. It is, however, not the particular component which is worth mentioning here. A single component in itself has no value whatsoever. Only due to the fact that they are capable of communicating with each other are they in a position to form coalitions by which these elements can achieve jointly what all elements separately would never be able to accomplish.

The secret behind the advantage for the individual to join a coalition lies, of course, in a super additive composition rule applicable for communicating elements. By this I refer to the old—but unfortunately inaccurate—saying that "the whole is more than the sum of its parts." Although this statement has been under heavy attack by positivists, operationalists, *etc.*, if put properly it emerges as a most important guiding principle in the theory and technology of self organizing and adaptive systems. Properly formulated we would say today that to a set of communicating elements we have to apply a superadditive composition rule, because "a *measure* of the sum of its parts is larger than the sum of the *measure* of its parts." Consider the function Φ as a measure; then for the two parts x, y we have:

$$\Phi(x + y) > \Phi(x) + \Phi(y)$$

This equation is nothing else but my definition of a coalition being put into precise, mathematical language. If you have any doubt as to the existence of such a measure which will satisfy the above equation, I suggest, *e.g.*, using for $\Phi = (\quad)^2$, that is "taking the square." We have

$$(x + y)^2 > x^2 + y^2$$

which is obviously true, because $(x + y)^2 = x^2 + y^2 + 2xy$, hence the left hand side of the inequality always exceeds the right hand side by an amount of $2xy$. Perhaps the following biological example will make my point of a super-additive composition rule even clearer:

The most important example of such a measure in connection with my topic is one which has been developed in information-theory(4). It is the concept of "certainty" or, as it is often referred to, as "neg-entropy," symbolized by $-H$. One of the most important findings in this theory is that the certainty of a joint event $-H(x + y)$ is always larger or equal to the sum of the certainties of the individual events, $-H(x)$ $-H(y)$, equality being the case only for