Daniel Dennett is famous for the concept of the "intentional stance" which he introduces in his essay *Intentional Systems*. In the following paragraphs I will elaborate on Dennett's notion of the "intentional stance", how it is applied to "intentional systems" and which problems and advantages one barters for this application.

While other theories try to reduce intentional vocabulary ("beliefs", "desires" and also "rationality") to either physical (Smart's identity theory, for example) or functional terms (like Putnam's functionalism), Dennett grants the intentional vocabulary its own right of existence. His main point is that to describe a system, we can choose from three different stances: The physical, the design and the intentional stance. The purpose of the description and knowledge about the system determine which of the three description levels is appropriate for the viewer. If the viewer decides in favor of the intentional level he adopts the "intentional stance", for the physical and functional level he adopts the physical or design stance respectively. The viewer does not claim that there *is* anything like in the system that makes it work in a certain way, he just views it as if there *was*.

There are some peculiarities about the "higher" two stances, which I will now go into. Dennett uses the example of a chess computer to explain most of his points. If the chess computer is not connected to the power socket, for example, predictions about its behavior may fail for reasons on the physical level, even if the functional organization of a system is totally understood. Predictions on the intentional level can fail for reasons with origin in both, the physical and the design level (the latter being a programming error of the chess computer, for example). As mentioned above, the intentionality and the functional organization are only ascribed to the system, that is, there is not necessarily anything in it that makes it work as predicted.

Reasons to ascribe intentionality to a system can be found on an empirical basis – one can predict the move of the chess computer by assuming it has the game state and the rules as beliefs and the winning as a desire. In other cases – in man and animal, basically, but also in aliens – the reasons can also be inferred from the existence of a system's evolutionary history: Only if the system acts goal directedly (pursues desires rationally) on useful information on its environment (beliefs) evolution can ensure survival and propagation (p. 8).

Still it has to be asked why Dennett, contrary to his predecessors, not only allows the error-prone intentional stance, but also claims that its adoption may be useful in certain cases. As I understand there are two reasons to adopt the intentional stance: First, to predict the behavior of complex systems, and second, closely related, to verify theories worded on the design stance. I will now explain both uses in detail.

Human beings, animals and even complex machineries as chess computers are hard to describe on a level based on inputs and outputs (as tried by behaviorists) or just abstract states characterized by their inputs (as functionalists argue it can be done). The intentional stance, on the other hand, is empirically warranted – sciences as game theory and economics, for example, successfully depend on it for their predictions and it is has proven useful to adopt it towards man and animal. Thus it is at least empirically useful to adopt the intentional stance to predict behavior. But with this adoption we presume something to be in the system which is not necessarily there: Intelligence and rationality. Behaviorists would claim that this presumption is dangerous and should thus be avoided in the first place. Dennett argues for his theory by proving that the presumption (or "loan" as he puts it) is neither empirically vacuous nor irreducible (the loan cannot

be repaid):

Rationality assumptions are not empirically vacuous (in the sense that they can be easily adjusted to the data) because of their normative nature: Since the systems we use the intentional stance on are either produced for a certain goal or evolved with the goal of survival and propagation, we can assume that they will behave optimally. Systems with the same ascribed beliefs and desires thus should all act similar. This also holds true for not apparently survival-relevant behavior, as the predicted answer to questions like “What is five times seven?”, since language must have evolved to transport believes and to convince others of their truth. As a result there would be more true things believed, which is an evolutionary advantage and the only reason why language would evolve and persist at all.

Rationality “loans” can be “repaid” according to Dennett, since they can be used for a positivistic theory building approach: A theory that is supposed to predict behavior on another than the intentional level, say the design level, has to stand the comparison test with the predictions of the intentional stance. That way it can be (and is being, as Dennett points out) utilized for artificial intelligence, for example, to model automata that are similar to human beings.

To summarize the above claims about the intentional stance, when treating a system as an intentional one we ascribe beliefs, desires and rationality to it and abstract from what it is that causes these attributes. Empirical evidence shows that we can then assume optimal behavior and get reasonably good predictions. Occasional prediction errors can be attributed to false beliefs or system failures on a “lower” (design or physical) level.

I like Dennett’s approach, since he can give reasons for rational, intelligent behavior, and thus why we get so far by predicting from the intentional stance. With this argument he can establish the intentional stance as a serious alternative to the other stances. On the other hand he does not claim it to be the *non plus ultra* of stances, he rather points out its usefulness for researching the design stance. If this research would succeed for the system “human being”, then we could physically model a system as we wish, give it the same design, and would thus create artificial life that would pass the turing test. However, the intentional stance is inherently very superficial and gives rise to conflict when predictions do not match the actual system behavior: It is not clear whether the expectations of the system should be modified by adjusting our assumptions about its beliefs and desires, or whether this should be considered to be a fault in its logic. The consequences of these two interpretations are very different: If we choose the second, we drop the assumption of rationality. Since this situation is very common, I suppose the empirical content of the intentional stance is potentially very small.