

A Reading of
Michael E. Bratman's
What is Intention?

In his article, Bratman tries to address the philosophical basis of intention from the point of view of the spiritual/material divide. The existence of intentions - mental constructs directed at future action - poses the following problem: intentions either influence future action in a "ghostly" fashion (metaphysically unsound), or exist irrevocably between the time of intent and the time of the action (rationally unsound) or do not have any importance to the action they relate to (functionally unsound).

The author then goes on to counter this problem by backing intention's *raison d'être* on two main accounts: its role in practical reasoning, i.e. that actions and reaching them require time-consuming planning and deliberation, as well as coordination with other actions of both the agent and their collaborators; and its role in the formation of beliefs, both within the agent and within its collaborators, mainly as a result of the permanence and commitment inherent to intentions.

The paper then deals for a large part with what I found to be a rather irrelevant (at least to us) discussion of a problem in field of ethics, namely that of the "*package deal*". The only point of interest to us here is in the revelation of the twofold connectivity of intentions through both backward-projection towards reasoning leading up to an intention, and forward-projection towards planning and action, emphasizing the pivotal role of intent in autonomous agents.

I find Bratman's reasoning to stray from its original path, that of defining intention, and even from that of evading his so-called "*trilemma*". In addition, I feel that his main argument does not hold many insights into the understanding of intentions that cannot be derived from folk psychology and common sense. The threefold conclusion he starts with seems artificial to me, and I do not fully agree that intentions run this risk (why should intentions be irrevocable to have an influence on future actions?).

That said, some of his detailed analysis in support of his arguments are of interest to us, even if they're not his main point. This might be due to the fact that his way of thinking emerges from involvement in the design of artificial intelligence systems. Two points in particular strike me as interesting:

Similar to other authors, Bratman stresses the the role of intention in a hierarchical and forward-moving *structure of desire, intent and planning*. Early intentions are cross-valued with internal, and - more interestingly - external constraints and result in more practical intentions, closer related to actions. In the context of agent collaborations, this inverse tree-structure can be utilized in coordinating individual agents' plans into a larger collaborative plan, through cross-constraint solving. Prior intentions of many agents can be evaluated with one another and the agents can thus reach a common, more practical intent.

Another interesting point he raises is that of *defaulting the persistence* of intentions over their reevaluation. This can be a handy heuristic both for designing companion agents to be more life-like, by preferring them to keep their intentions unless radical new data is acquired, and for the design of belief systems of such agents as to their human counterpart's intentions. Collaborative agents should assume that their counterpart does not change their intention unless there is a good reason to do so. This persistence is crucial in building belief models, because - as Bratman notes - intentions strongly foster belief in the intention's action.