

Guy Hoffman – Fall 2003

A Reading of
John A. Searle's
Collective Intentions and Actions

The main assertion of Searle's paper is that joint intention and collective intentionive behavior cannot be described as a function of the collective singular intentions and intentionive behaviors of the agents engaging in the joint activity.

He first lays out intuitively why this is true, building on a variety of failures inherent to previous attempts to define joint intention as a function of the individual intentions involved. One of his main claims is that actually a reverse dependency takes place, namely that the individual intentions do only get a meaning as part of the joint intention ("I'm pushing this car as part of our collective pushing").

Further, Searle tries to define a notation to support his observation. This can be summarized as the introduction of a new type of intention-to-action, namely a "collective i.a. by means of singular i.a.". This formalism's goal is mainly to suggest that we do indeed need a new type of i.a. to describe collective intent.

Personally, I did not find much value in the described notation, and I found that its solution is rather artificial. While supporting his main claim (which may be true), the formal solution seems to stem from the particular way intentions were formalized by him in the first place -- in terms of the action that satisfies them. This definition is actually what causes the most intuitive notation to be rejected by Searle .

I do find his final section insightful, though. In it, Searle suggests that society, rather than being an outcome of collective intent and action, is actually a prerequisite of the above. This is interesting since it says something about our innate drive towards collaboration, our instinctive wiring for formulating our own intentions in terms of collective intentions.

There is also a thread going back to work we have seen in developmental psychology demonstrating our instinctive appreciation of, and attention to goals rather than actions. It seems clear to me that as a collective-intent creature, our understanding of goals is vital.

Artificial collaborative agents, then, must also be goal-oriented to display

true collective intent – and as a result, collaborative sensibility. Naturally, machines do not have such an innate social sense. But, building collaborative machines, we might want to consider defining an overarching "We"-goal that both determines the derived "I"-goals and also aids in analyzing the human collaborator's actions in terms of his derived "I"-intents.