Improving the Odds for Concensus

Whitman Richards

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: Decision-making among a collection of interacting entities, or agents, entails aggregating the beliefs of the constituent agents to guide the behavior of the system as a whole. If the agents' beliefs are unconstrained, then cyclic outcomes are highly likely, in the worst cases reaching probability one. Often the systems are driven into chaotic behavior. How, then, can one improve the odds for reaching a stable concensus?

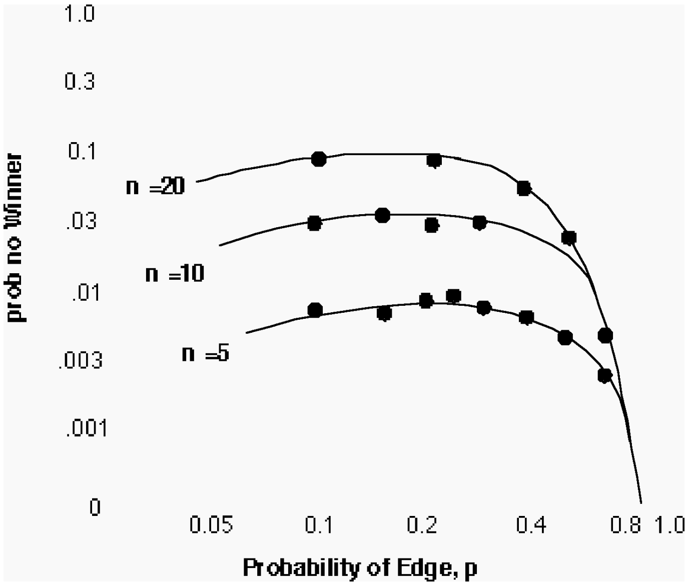
Motivation: Collective behavior based on the aggregation of individual preferences is widespread. One obvious application is in societal decision-making when a group of individuals (or nations) must aggregate their diverse preferences. Another application is to understand the human brain, which can be viewed as consisting of a large number of semi-autonomous neural modules acting as "agents" (Minsky, 1986; van Essen et al. 1992.) If we assume these kinds of aggegation processes are not subject to a "dictator" (such as a homunculus), then there is always the possibility of unstable outcomes, meaning that there may be cycles among alternatives (Arrow, 1963; McKelvey, 1979; Saari & Haunsperger, 1991.) Yet most collective decisions, espcecially our own percepts, are stable, singular and typically compelling. Hence we seek some simple, plausible constraints on the voting behaviors of any society of agents, such that concensus is almost always reached. [3].

Previous Work: The key constraint is that an agent's choices should be consistent with a "mental model" that is shared by all the agents (Richards et al, 1998.) This shared model forces the pereference orderings for an agent's alternative choices to be correlated with the preference orderings of the other agents, rather than being just a random order. These "mental models" express the relations between the alternatives, and hence can be represented as a graph, with the vertices of the graph being the alternatives, and the edges showing the relationships between alternatives. Given the vertex representing an agent's first choice, his second choices will then be those alternatives represented by the adjacent vertices, etc. If all agent's choices are consistent with the shared "mental model", then concensus will be reached better than 90% of the time for mental models generated as random graphs. [1]

Approach: Mental Models depicted as random graphs have two important variables that affect the probability of reaching concensus: the number of vertices, n, and the probability of an edge, p. If p = 1, then we can show that concensus will always be achieved. (A weaker condition for concensus is that the graph simply be covered.) As p -¿ 0, the odds for concensus decreases. This result is shown in Fig. 1, where the ordinate shows the probability of no agreement. These results are for random graphs only. For certain types of graphs, such as those with covers, k-partite constructions, or unrooted trees, we can show that concensus will always be reached if the graph structure is respected in thepreference orderings.

Impact: Although concensus occurs roughly 90% of the time or better when a mental model is shared among agents, an important lesson is that the form of the mental model is an important variable in collective stability. Furthermore, contrary to findings that emphasize the stability-inducing properties of probabilistic choice, such as the Condorcet Jury Theorem or probabilistic voting, our results suggest that a failure to reach concensus becomes more likely if agents choose randomly from alternatives over which they have no firm knowledge.

Future Work: Our own Mental Models are not arbitrary, and tend to have certain forms. For example, heirarchical trees, "subway maps", linear orderings are very common. Which of these forms are likely to achieve concensus, and which not ? There are also severe limits on how many alternatives an agent can handle at once (e.g. "magical number seven"), which restricts the scope of a model, or at least its "active" region. The effects of unintentional misorderings of preferences (ie "noise") is unexplored, and the related issue of "mis-matched" mental models. Alternatives may not be symmetically related; hence directed graphs should be understood. Certain asymptotic conditions remain unresolved. Finally, we would like to design a neural net that would perform a pair-wise comparison of weighted alternatives, related by arbitrary graphical forms.



References:

[1] D. Richards, B. McKay, and W. Richards. Collective Choice and Mututal Knowledge Structures. *Adv. in Complex Systems*, 1, 221-236, 1998.

[2] W. Richards, B. McKay, and D. Richards. The Probability of Choice with Shared Knowledge Structures. *AI Memo*, 1690, MIT, July, 2000.

[3] M. Minsky Society of Mind. *Simon and Schuster*, NY, 1986.