Measuring Diversity of Preferences in a Group[1]

Vahid Hashemi Ulle Endriss

Institute for Logic, Language and Computation (ILLC)
University of Amsterdam

Abstract

We introduce a general framework for measuring the degree of diversity in the preferences held by the members of a group. We formalise and investigate three specific approaches within that framework: diversity as the range of distinct views held, diversity as aggregate distance between individual views, and diversity as distance of the group's views to a single compromise view. While similarly attractive from an intuitive point of view, the three approaches display significant differences when analysed using both the axiomatic method and empirical studies.

1 Introduction

Preferences are ubiquitous in AI [2]. Examples for application domains include recommender systems, planning, and configuration. Of particular interest is the case of preference handling in multiagent systems, where several agents each have their own individual preferences and we need to take decisions that are appropriate in view of such a profile of preferences. The normative, mathematical and algorithmic aspects of this problem are studied in the field of (computational) social choice [1].

Intuitively speaking, one might expect that the less *diverse* the preferences in a group of agents are, the easier it should be to come to mutually acceptable decisions. For example, in the most extreme case where all agents share the exact same preference order, it will be trivial to make collective decisions. *Vice versa*, the more diversity we find in a group, the more we should expect to encounter paradoxes, i.e., situations in which different social choice-theoretic principles would lead to opposing conclusions. In recent work we have proposed a new formal model of preference diversity to study such phenomena [3]. In this extended abstract we sketch this formal model and report on some of the results obtained.

2 Preference Diversity Indices

Let $\mathcal{X}$ be a finite set of m *alternatives*. We model *preferences* as (strict) *linear orders* over $\mathcal{X}$ and write $\mathcal{L}(\mathcal{X})$ for the set of such all preference orders. Let $\mathcal{N} = \{1, \ldots, n\}$ be a finite set of *agents*. A *profile* $\boldsymbol{R} = (R_1, \ldots, R_n) \in \mathcal{L}(\mathcal{X})^n$ is a vector of preference orders, one for each agent. The *support* of a profile $\boldsymbol{R} = (R_1, \ldots, R_n)$ is the set of preference orders occurring in it: $\textsc{Supp}(\boldsymbol{R}) = \{R_1\} \cup \cdots \cup \{R_n\}$. We call a profile $\boldsymbol{R}$ *unanimous* if $|\textsc{Supp}(\boldsymbol{R})| = 1$, i.e., if it is of the form $(R, \ldots, R)$.

Given two profiles $\boldsymbol{R}$ and $\boldsymbol{R'}$ (with the same number of agents n expressing preferences over the same number of alternatives m), we want to be able to make judgments about which of them we consider to be more diverse. To this end, we introduce the concept of *preference diversity index*.

Definition 1. *A* ***preference diversity index*** *(PDI) is a function* $\Delta : \mathcal{L}(\mathcal{X})^n \to \mathbb{R}^+ \cup \{0\}$, *mapping profiles to the nonnegative reals, such that* $\Delta(\boldsymbol{R}) = 0$ *for any unanimous profile* $\boldsymbol{R} \in \mathcal{L}(\mathcal{X})^n$.

[1] The full paper appears in the *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI-2014).*

Given a particular PDI Δ, we say that profile $\boldsymbol{R}$ is more diverse than profile $\boldsymbol{R}'$ if $\Delta(\boldsymbol{R}) > \Delta(\boldsymbol{R}')$. That is, any given PDI Δ induces a weak order on the set of all profiles (which we call a *preference diversity order*), ranking them from most to least diverse. We now introduce three specific families of PDI's:

(1) The *simple support-based PDI* counts the number of distinct preference orders in a profile. This idea can be generalised to counting, for a given $k \leqslant m$, the number of distinct ordered k-tuples of alternatives appearing in a profile (suitably normalised to ensure unanimous profiles map to 0).

(2) Under a *distance-based PDI*, we measure the distance (e.g., the *Kendall tau distance* or the *discrete distance*) between any two preference orders in a given profile and then aggregate the values obtained (e.g., by computing their sum or their maximum).

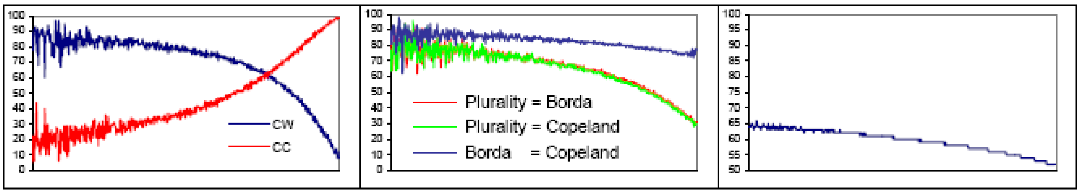
(3) Under a *compromise-based PDI*, we first aggregate the individual preferences (e.g., using the *Borda rule* or the *majority rule*), then compute the Kendall tau distance of each individual preference to that "compromise", and finally aggregate (e.g., add) the values thus obtained.

Other instantiations of the concept of PDI are certainly possible, and we consider identifying further specific PDI's an important direction for future work.

In the full paper we propose several *axioms* capturing certain desirable properties of PDI's and analyse which specific PDI's satisfy which of these properties [3].

3 Experimental Results

We have conducted several experiments, using both synthetic preference data and profiles sampled from real preference data, to obtain a better understanding of the concept of PDI, and in particular to gain insights into the effects of diversity on certain social choice-theoretic phenomena. As an example, consider the three graphs below, which concern scenarios with 5 alternatives and 50 voters (averaged over 1 million profiles drawn from the synthetic distribution defined by the so-called *impartial culture assumption*, under which every logically possible preference profile is equally likely to occur in practice):



For all three graphs, the x-axis corresponds to the diversity of profiles as measured by the distance-based PDI computing the sum of Kendall tau distances between pairs of individual preference orders. On the y-axis we see the frequency (in percent) of a particular phenomenon occurring. The first graph shows that, as diversity increases, the chance of encountering a profile with a *Condorcet winner* decreases, while the chance of encountering a profile with a *Condorcet cycle* increases. The second graph plots the frequency of two voting rules agreeing on the election winner against the diversity of the preferences of the voters. The third graph shows how average voter satisfaction (measured in terms of the proportion of alternatives a voter ranks above the election winner, in this case determined using the Borda rule) decreases as diversity increases.

Thus, we see a clear correlation between increases in diversity and both decreases of desirable effects and increases of undesirable effects. These results match our intuition well, thereby giving credence to our claim that PDI's are a useful formal model for capturing the important but elusive notion of diversity.

References

[1] F. Brandt, V. Conitzer, and U. Endriss. Computational social choice. In G. Weiss, editor, *Multiagent Systems*, pages 213–283. MIT Press, 2013.

[2] C. Domshlak, E. Hüllermeier, S. Kaci, and H. Prade. Preferences in AI: An overview. *Artificial Intelligence*, 175(7):1037–1052, 2011.

[3] V. Hashemi and U. Endriss. Measuring diversity of preferences in a group. In *Proceedings of the 21st European Conference on Artificial Intelligence (ECAI-2014)*. IOS Press, 2014.