1 Introduction

Modeling knowledge with partitions is very common and almost unique in game theory and information economics. In this construction each individual is endowed with a partition of the set of possible states which can be interpreted as the knowledge of the people. One interpretation of such model is that a partition, or in particular the cells of a partition, represents the state of the mind (See, for example Zamir 2008) for the individual. The aim of this note is to offer a metric that measures partitions in the light of this interpretation. According to the proposed metric the distance between two partitions is the weighted average of the non-empty symmetric differences of each cell that contain each element of the set by excluding double counting.

Although it is possible to use this metric as an index for other purposes such as cluster analysis, categorization theory or even data mining given the common usage of partitions in these areas (See Wagner and Wagner 2007 for a thorough coverage), our primary motivation is to measure partitions for game theory or decision theory with the interpretation given above.

The rest of the paper is organized in the following way: next section gives the relevant notations and definitions, and introduces the proposed metric. After that we compare some important distance measures defined in the literature with the proposed one through examples.

2 Notations and the metric

Let Ω be a finite set with n members. We say that a collection $\mathcal{P} = \{P(\omega)\}_{\omega\in\Omega}$ of sets is a partition of Ω, and call the $P(\omega)$ the atoms of the partition, if

$$P(\omega) \cap P(\omega') = \emptyset \quad \text{for} \quad \omega \neq \omega', \quad \text{and} \quad \bigcup_{\omega} P(\omega) = \Omega.$$

In the interpretation, Ω is the set of all possible states ω that is relevant for the situation at hand. When some state ω_0 is realized a player's knowledge will be represented by the atom $\mathcal{P}$ that contains ω_0. Any member ω in $\mathcal{P}$ is indistinguishable from ω_0 from the viewpoint of the player.

Let $\mathscr{P}$ be the set of all partitions of Ω. Consider two arbitrary members $\mathcal{P}$ and $\mathcal{Q}$ of this set. We can interpret each partition as the mental state of each player. Define the symmetric difference of two arbitrary sets A and B, denoted by $A\Delta B$, as

$$A\Delta B := (A \setminus B) \cup (B \setminus A).$$

Now consider the collection $\mathcal{S}$ of symmetric differences given by,

$$\mathcal{S} := \{P(\omega_i)\Delta Q(\omega_i) \neq \emptyset : P(\omega_i) \in \mathscr{P}, Q(\omega_i) \in \mathscr{Q}, i = 1, \dots, n\} = \{S(\omega_i)\}_{i=1}^{n}.$$

The collection $\mathcal{S}$ contain symmetric differences of atoms in each partition that contain ω. It is not difficult to see that

$$P(\omega_i)\Delta Q(\omega_i) = P(\omega_j)\Delta Q(\omega_j) \quad \text{for some} \quad i \neq j.$$

Note that $\mathcal{S}$ contains only one of those such repetitive sets and this precludes double counting.

Definition 1. *Define $\rho : \mathscr{P} \times \mathscr{P} \to \mathbb{R}$*

$$\rho(\mathcal{P}, \mathcal{Q}) := \sum_{i=1}^{n} \frac{r_i}{\binom{n}{s_i}} - \sum_{r_i \in D} \frac{r_i}{\binom{n}{s_i}}$$

where s_i is the cardinality of the atom $S(\omega_i)$ that contains ω_i in $\mathcal{S}$ and r_i is 0 if $s_i = 0$, otherwise 1. D is the set of items that show more than one. We can write in a compact form by excluding double counting

$$\rho(\mathcal{P}, \mathcal{Q}) := \sum_{S(\omega) \in \mathcal{S}} \frac{r}{\binom{n}{s}}$$

where $\mid S(\omega) \mid = s$, r is the total number of atoms in $\mathcal{S}$ with cardinality s, and r_i is the cardinality of sets in $\mathcal{S}$ with i element.

To make sense out of the definition consider the following example.

Example 1. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with two partitions $\mathcal{P} = \{\{\omega_1, \omega_3\}, \{\omega_2\}\}$ and $\mathcal{Q} = \{\{\omega_1, \omega_2\}, \{\omega_3\}\}$. Then $\mathcal{S} = \{S(\omega_1), S(\omega_2), S(\omega_3)\} = \{\{\omega_2, \omega_3\}, \{\omega_1\}, \{\omega_1\}\}$. Note that $S(\omega_2) = S(\omega_3)$ even though $\omega_2 \neq \omega_3$. The distance $\rho(\mathcal{P}, \mathcal{Q})$ is then $2/3$.

Proposition 1. *The function ρ is a metric on the set $\mathscr{P}$ of all partitions for a given Ω.*

Proof. (1) $\rho(\mathcal{P}, \mathcal{Q}) = 0$ implies $r = 0$ for each atom in $\mathcal{S}$. In other words collection of symmetric differences contain empty sets. That is $P(\omega)\Delta Q(\omega) = \emptyset$ which implies $P(\omega) = Q(\omega)$. Hence every omega is assigned in the same atoms in both $\mathcal{P}$ and $\mathcal{Q}$. Thus, $\mathcal{P} = \mathcal{Q}$.

Conversely, it is immediate to conclude that $\mathcal{P} = \mathcal{Q}$ implies $\rho(\mathcal{P}, \mathcal{Q}) = 0$.

(2) Since symmetric differences in sets are by definition satisfy symmetry, we have $\rho(\mathcal{P}, \mathcal{Q}) = \rho(\mathcal{Q}, \mathcal{P})$.

(3) To finish the proof we need to show for arbitrary $\mathcal{P}, \mathcal{Q}$ and $\mathcal{R}$

$$\rho(\mathcal{P}, \mathcal{Q}) \leq \rho(\mathcal{P}, \mathcal{R}) + \rho(\mathcal{R}, \mathcal{Q})$$

We prove the claim by induction on $\mid \Omega \mid = n$. The claim is trivial when $n = 1$.

Suppose now the claim is true for some n i.e., $\mid \Omega \mid = n$. Consider now the set $\Omega \cup \{\omega_0\}$ so that the cardinality is $n + 1$. Consider arbitrary partitions $\mathcal{P}$, $\mathcal{Q}$ and $\mathcal{R}$ of $\Omega \cup \{\omega_0\}$. These partitions can be obtained by partitions $\mathcal{P}'$,$\mathcal{Q}'$ and $\mathcal{R}'$ of Ω, by including ω_0 to them, respectively.

There are two possibilities for adding a new element to a given partition. It can either be added as a singleton - the atom contains only the new element - or it can be included into one of the existing atoms. With the abuse of notation, let us denote the former case $\mathcal{X} \oplus x$ and the latter case $\mathcal{X} \uplus x$ for an arbitrary partition $\mathcal{X}$ and an arbitrary new element x.

(3.1) Suppose we obtain each partition by adding ω_0 separately. That is

$$\mathcal{P} := \mathcal{P}' \oplus \omega_0, \quad \mathcal{Q} := \mathcal{Q}' \oplus \omega_0, \quad \mathcal{R} := \mathcal{R}' \oplus \omega_0$$

So by induction hypothesis we have

$$\rho(\mathcal{P}', \mathcal{Q}') \leq \rho(\mathcal{P}', \mathcal{R}') + \rho(\mathcal{R}', \mathcal{Q}').$$

Equivalently,

$$\sum_{i=1}^{n} \frac{r_i}{\binom{n}{s_i}} - \sum_{r_i \in D} \frac{r_i}{\binom{n}{s_i}} \leq \sum_{i=1}^{n} \frac{\tilde{r}_i}{\binom{n}{\tilde{s}_i}} - \sum_{\tilde{r}_i \in \tilde{D}} \frac{\tilde{r}_i}{\binom{n}{\tilde{s}_i}} + \sum_{i=1}^{n} \frac{\hat{r}_i}{\binom{n}{\hat{s}_i}} - \sum_{\hat{r}_i \in \hat{D}} \frac{\hat{r}_i}{\binom{n}{\hat{s}_i}}$$

Then since $S(\omega_0) = \emptyset$ - $r_{n+1} = 0$ - in every collection of symmetric differences, and the other $S(\omega_i)$ would not change, the above equation implies

$$\begin{aligned}\sum_{i=1}^{n} \frac{r_i}{\binom{n+1}{s_i}} - \sum_{r_i \in D} \frac{r_i}{\binom{n+1}{s_i}} \leq & \sum_{i=1}^{n} \frac{\tilde{r}_i}{\binom{n+1}{\tilde{s}_i}} - \sum_{\tilde{r}_i \in \tilde{D}} \frac{\tilde{r}_i}{\binom{n+1}{\tilde{s}_i}} \\ & + \sum_{i=1}^{n} \frac{\hat{r}_i}{\binom{n+1}{\hat{s}_i}} - \sum_{\hat{r}_i \in \hat{D}} \frac{\hat{r}_i}{\binom{n+1}{\hat{s}_i}}\end{aligned}$$

which is equivalent to

$$\rho(\mathcal{P}, \mathcal{Q}) \leq \rho(\mathcal{P}, \mathcal{R}) + \rho(\mathcal{R}, \mathcal{Q})$$

(3.2) Consider now ω_0 is included one of the existing atoms of the given partitions, say $\mathcal{P}'$. That is

$$\mathcal{P} := \mathcal{P}' \uplus \omega_0, \quad \mathcal{Q} := \mathcal{Q}' \oplus \omega_0, \quad \mathcal{R} := \mathcal{R}' \oplus \omega_0$$

Now suppose $P'(\omega_i)$ is the atom that contains ω_0 in $\mathcal{P}$. So the only difference between $\mathcal{P}'$ and $\mathcal{P}$ is that the atom $P'(\omega_i)$ includes ω_0 in addition to other elements. Then symmetric difference $P'(\omega_i)\Delta Q'(\omega_i) = S(\omega_i)$ will include ω_0 which causes change in the left-hand side of the triangle inequality. Also, $S(\omega_0)$ may cause a change. However there will be identical changes in left hand side of the triangle inequality through a change in $P'(\omega_i)\Delta R'(\omega_i)\tilde{S}(\omega_i)$. This is also true for $\tilde{S}(\omega_0)$.

Finally, the last component will not change because $\mathcal{Q}$ and $\mathcal{R}$ contain ω_0 as an atom so symmetric differences will not change and $\hat{S}(\omega_0) = \emptyset$. As in the previous case the conclusion follows.

Note that with similar argument we can show similar cases where ω_0 is added two partitions as a singleton and added into the remaining one as part of the existing atoms. Formally, by the symmetry of arguments, this case implies the same conclusion for the following cases.

$$\mathcal{P} := \mathcal{P}' \oplus \omega_0, \quad \mathcal{Q} := \mathcal{Q}' \uplus \omega_0, \quad \mathcal{R} := \mathcal{R}' \oplus \omega_0,$$
$$\mathcal{P} := \mathcal{P}' \oplus \omega_0, \quad \mathcal{Q} := \mathcal{Q}' \oplus \omega_0, \quad \mathcal{R} := \mathcal{R}' \uplus \omega_0.$$

(3.3) Consider now another case where ω_0 is included only in one partition, say $\mathcal{P}$, as a separate atom. That is

$$\mathcal{P} := \mathcal{P}' \oplus \omega_0, \quad \mathcal{Q} := \mathcal{Q}' \uplus \omega_0, \quad \mathcal{R} := \mathcal{R}' \uplus \omega_0.$$

Now suppose $Q'(\omega_i)$ and $R'(\omega_j)$ are the atoms including ω_0 in $\mathcal{Q}$ and $\mathcal{R}$, respectively. This will cause a change in $P'(\omega_i)\Delta Q'(\omega_i) = S(\omega_i)$ and symmetric differences of every component in these atoms. So left hand side of the triangle inequality will change, however, symmetrical changes will happen on the right hand side because of symmetrical differences of atoms of $\mathcal{P}$ and $\mathcal{R}$. If we include symmetric differences of atoms of $\mathcal{Q}$ and $\mathcal{R}$ the result follows.

(3.4) Note that the last case in which ω_0 added to existing atoms of $\mathcal{P}'$, $\mathcal{Q}'$ and $\mathcal{R}'$ follows the same logic with initial case where ω_0 added as a separate atom to each of the partitions.

The list covers all possible cases and the claim follows by induction. □

The logic of the above proof is that adding a new element to create new partitions would create similar effects on the each side of the inequality. The only change then is to have one more element in the denominator of the formula, but this means only rescaling without effecting the direction of the inequality.

The metric is weighting the total number of the atoms of $\mathcal{S}$ with the same cardinality. Since each atom is a subset of Ω, we use the probability of obtaining that particular atom as our weight. Intuitively, the symmetric difference measures how far apart the sets are[1] and gives us subsets of the original set Ω. We weighted each subset with the probability of obtaining it from Ω.

3 Comparison with other metrics

The distance measures for partitions proposed in the literature are mostly distance indices which can be categorized into three groups[2]: indices constructed with combinatoric approach,

[1]In a game theoretic context this refers to distance between knowledge of agents when the true state is ω.

[2]Note that this classification is not common nor uniform.

indices constructed with informational approach and metrics that rely on tools from probability (or, measure theoretic metrics). Roughly speaking, the indices in the former group have been constructed by counting the number of pairs that agree in different partitions such as Mirkin and Chernyi (1970) and William (1971) or by counting the number of pairs that disagree in different partitions such as Arabie and Boorman (1973). Note that indices constructed by counting the agreed pairs such as Rand index are not a metric since the distance between two identical partitions is 1 according to them. Counting disagreed pairs, however, solves this problem and establishes a metric.

The construction of the indices in the second group is based on Shannon (1948) entropy such as De Mántras (1991) and Simovici and Jaroszewicz (2003). The logic behind these indices is to make a random variable by using partition structure to employ entropy which is introduced for a random variable distribution. The desired random variable is obtained by taking the ratio of cardinality of each atom to the cardinality of the original set.

The metrics in the last group are more common in economic theory. One of the earliest example of such (semi) metric is due to Boylan (1971). The (semi) metric is defined on sub-sigma-algebras and it allows to measure differences between information structures. Allen (1983), Stinchcombe (1990) and Monderer and Samet (1996) use this metric to measure informational differences and to topologize abstract space of information. Recently, Mohlin (2015) proposed another metric for the same purpose. This metric weights symmetric differences of each cell with their intersection.

We now compare our metric with the ones in the last group. The main reason for this is that the metrics in the last group have wide usage in economic theory or are proposed for economic theory in mind. The metrics in the other groups are designed for cluster analysis or data mining purposes and measuring distances with these metrics generally produce counter-intuitive results in the context of game theory or decision theory. Also, metric proposed by Mohlin (2015) have a very close relation with the metrics in the first two groups. So assessing one of them would give enough idea about the implications of the other.

Before proceeding let us first give the definitions of these two metrics.

Definition 2 (Boylan (1971))**.** *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. The function ρ_b given by*

$$\rho_b(\mathcal{P}, \mathcal{Q}) := \sup_{P\in\mathcal{P}} \inf_{Q\in\mathcal{Q}} \mu(P\Delta Q) + \sup_{Q\in\mathcal{Q}} \inf_{P\in\mathcal{P}} \mu(P\Delta Q)$$

defined on sub-sigma-algebras of $\mathcal{B}$ is a semi-metric.

Definition 3 (Mohlin (2015))**.** *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space. The function ρ_m given by*

$$\rho_m(\mathcal{P}, \mathcal{Q}) := \sum_{P(\omega)\in\mathcal{P}} \sum_{Q(\omega)\in\mathcal{Q}} \mu\Big(P(\omega) \cap \mu(Q(\omega)\Big)\mu\Big(P(\omega)\Delta Q(\omega)\Big)$$

defined on $\mathscr{P}$ is a metric.

In the following example consider μ as the cardinality of each set at hand[3].

[3]With some tedious calculations, it is possible to show that no measure μ with Boylan metric allows the symmetry property we wanted.

Example 2. Consider again $\Omega = \{\omega_1, \omega_2, \omega_3\}$ with the partitions $\mathcal{P} = \{\{\omega_1, \omega_3\}, \{\omega_2\}\}$, $\mathcal{Q} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}\}$ and $\mathcal{R} = \{\{\omega_1, \omega_2, \omega_3\}\}$. Intuitively moving from $\mathcal{P}$ to $\mathcal{Q}$ and to $\mathcal{R}$ should have the same distance because of the symmetry of the situation. That is, moving from $\mathcal{P}$ to $\mathcal{Q}$ means distinguishing one more thing and moving from $\mathcal{P}$ to $\mathcal{Q}$ means mixing up one more thing so that the distance from moving $\mathcal{P}$ to $\mathcal{Q}$ and to $\mathcal{R}$ should be the same. The following table summarizes the distance for each of the metrics.

Metric	$\mathcal{P}, \mathcal{Q}$	$\mathcal{P}, \mathcal{R}$	$\mathcal{R}, \mathcal{Q}$
ρ	2/3	2/3	1
ρ_b	2	3	4
ρ_m	2	4	2

Table I: Distances according to different metrics

Observe that the equality $\rho_m(\mathcal{P}, \mathcal{Q}) = \rho_m(\mathcal{R}, \mathcal{Q})$ is highly counterintuitive.

Note that the results given above can be checked for different partitions and different finite sets as well. The reason that we propose our metric is to measure informational differences as intuitively as possible. Also our measure offers a unique way of measurement. The other two metrics are sensitive to the measure μ. It is also difficult to make sense out of these metrics if we allow two different measures μ and μ' for different individuals. So in that sense, it is difficult to use these metrics with heterogeneous or multiple priors.

4 Conclusion

This paper proposed a metric for partitions. Although it can be useful in data mining and clustering analysis, our hope is that it can be applied primarily to game theoretic and decision theoretic situations. One possible set up where this metric is useful would be information design problems in which a principal sets up an information structure to obtain a certain outcome. If the information design is a costly task then the metric proposed in this paper can be used to measure the cost of alternative designs.

References

Allen, B. (1983) "Neighbouring information and distributions of agents' characteristics under uncertainty" *Journal of Mathematical Economics* **12**(1), 63-101.

Arabie, P. and S.A. Boorman (1973) "Multidimensional scaling of measures of distance between partitions" *Journal of Mathematical Psychology* **10**(2), 148-203.

Boylan, E. S. (1971) "Equiconvergence of martingales" *The Annals of Mathematical Statistics* **42**(2), 552-559.

De Mántaras, R. (1991) "A distance-based attribute selection measure for decision tree induction" *Machine Learning* **6**(1), 81-92.

Mirkin, G.B. and L.B. Chernyi (1970) “Measurement of the distance between distinct partitions of a finite set of objects” *Automation and Remote Control* **31**(5), 786-792.

Mohlin, E. (2015) “A metric for measurable partitions” *Journal of Mathematical Psychology* **67**, 39-44.

Monderer, D. and D. Samet (1996) “Proximity of information in games with incomplete information” *Mathematics of Operation Research* **21**(3), 707-725.

Shannon, C.E. (1948) “A mathematical theory of communication” *Bell System Technical Journal* **27**(3), 379-423.

Simovici, D.A. and S. Jaroszewicz (2003) “Generalized conditional entropy and decision trees” in *Journees Francophones d'Extraction et de Gestion de Conneissances (EGC2003)* pp. 369-380 Lyon:France.

Stinchcombe, M.B. (1990) “Bayesian information topologies” *Journal of Mathematical Economics* **19**(3), 233-253.

Wagner, S. and D. Wagner (2007) *Comparing Clustering: An Overview* Universität Karlsruhe.

William, M.R. (1971) “Objective criteria for the evaluation of clustering methods” *Journal of the American Statistical Association* **66**(336), 846-850.

Zamir, S. (2008) “Bayesian games: games with incomplete information” The Hebrew University of Jerusalem discussion paper series 486.