Procedural Mixture Spaces

Rommeswinkel, Hendrik

National Taiwan University

5 March 2019

Procedural Mixture Spaces

Hendrik Rommeswinkel

March 5, 2019

Abstract

This paper provides a representation theorem for procedural mixture spaces. Procedural mixture spaces are mixture spaces in which it is not necessarily true that a mixture of two identical elements yields the same element. Under the remaining standard assumptions of mixture spaces, the following representation theorem is proven; a rational, independent, and continuous preference relation over mixture spaces can be represented either by expected utility plus the Shannon (1948) entropy or by expected utility under probability distortions plus the Rényi (1961) entropy. The entropy components can be interpreted as the utility or disutility from resolving the mixture and therefore as a procedural as opposed to consequentialist value.

1 Introduction

The expected utility representation by von Neumann and Morgenstern (1944) was initially stated using an "algebra of combining" of compound lotteries. Herstein and Milnor (1953) simplified their axioms greatly by introducing "mixture sets" which since have become known as mixture spaces. Mixture spaces can be given the interpretation of nested binary lotteries; we play one lottery to determine which lottery is played next, which determines which lottery is played afterwards, and so on. An important property of mixture spaces is that the "process" of mixing is irrelevant; a mixture between two identical outcomes is simply that exact outcome. There is no cost from resolving the lottery or joy from the thrill of its resolution.

In this paper, we drop the assumption $\mu a \oplus (1-\mu)a = a$, that any mixture between two identical elements of the mixture space is simply that same element. We maintain all other assumptions on the mixture space introduced by Herstein and Milnor (1953) along with the classical preference axioms of expected utility: Weak Order, Continuity, and Independence. These axioms are then shown to hold if and only if the preference can be represented by one of two functional forms. The first is expected utility plus a weight times the Shannon (1948) entropy. The second is expected utility under power-form probability distortion plus a weight times the Rényi (1961) entropy of the probabilities. In both cases, the entropy can be interpreted as a function representing the procedural gain or loss from going through the motions of resolving a mixture with a particular probability.

Entropy functions have been used in many applications in the sciences.[1] In this paper, we interpret the procedural mixture space as a lottery space in which the process of the resolution of the lottery matters. Based on the wide range of uses of entropy representations, we expect the concept of procedural mixture spaces to also be useful in other contexts.

We introduce the notation in Section 2, and present the axioms and the representation theorem in Section 3. The proof of the theorem is given in Section 4. In Section 5 we provide an application to the Allais (1953) paradox. In 6, we conclude with a short discussion of the method of the proof and state similar results that follow in almost the same fashion.

2 Notation

We follow the notation of Herstein and Milnor (1953). A capital Latin script letter denotes a set, a lowercase italic letter denotes an element of a set, a lowercase greek letter denotes an element of the real interval $[0, 1]$ and $\mathcal{S} = \{x|P\}$ denotes a set of elements having the property P. To avoid confusion, we will use $\oplus$ as our mixture operator instead of $+$, which will be only used for the addition of real numbers. $\mathbb{R}$ is the set of real numbers, $\mathbb{R}_{++}$ are strictly positive real numbers and we define $0 \ln 0 = 0$.

3 Axioms and Representation Theorem

The axioms of Herstein and Milnor (1953) for a mixture space $\mathcal{S}$ are given as follows:

$$\mu a \oplus (1-\mu)b \in \mathcal{S} \tag{1}$$
$$1a \oplus (1-1)b = a, \tag{2}$$
$$\mu a \oplus (1-\mu)b = (1-\mu)b \oplus \mu a, \tag{3}$$
$$\lambda[\mu a \oplus (1-\mu)b] \oplus (1-\lambda)b = (\lambda\mu)a \oplus (1-\lambda\mu)b] \tag{4}$$

where each axiom holds for all $a, b \in \mathcal{S}$ and all μ, λ. We may call these axioms, respectively, Closure, Connectedness, Commutativity, and Reduction of Compound Lotteries.

The last axiom however conflates two economically distinct properties, Associativity and Reducibility, which are, respectively:

$$\lambda[\mu a \oplus (1-\mu)b] \oplus (1-\lambda)c = (\lambda\mu)a \oplus (1-\lambda\mu)\left[\frac{\lambda(1-\mu)}{1-\lambda\mu}b \oplus \frac{(1-\lambda)}{1-\lambda\mu}c\right] \tag{5}$$
$$\mu a \oplus (1-\mu)a = a. \tag{6}$$

[1] The origins of entropy lie in thermodynamics and information theory. In economics, entropy has been suggested as a measure of inequality (Shorrocks, 1980), freedom (Suppes, 1996), and diversity (Nehring & Puppe, 2009). The literature on rational inattention following Sims (2003) employs entropy as a measure of uncertainty in macroeconomic models, a characterization is given in Caplin, Dean, and Leahy (2017). Frankel and Volij (2011) suggest entropy based indices of school segregation.

for all $a, b \in \mathcal{S}$ and all μ, λ.

In this paper, we remove the assumption of Reducibility. We give a possible interpretation. If the mixture space consists of lotteries, the decision maker may or may not treat a randomization between a and a as simply a occurring with certainty. This is because before a is resolved, something happens: a roulette wheel is spun, a dice thrown, etc.. The removal of the Reducibility assumption therefore means that we allow for a small procedural component –that only depends on the mixture probability– to play a role.

A procedural mixture space $\mathcal{S}$ therefore fulfills:

$$\mu a \oplus (1-\mu)b \in \mathcal{S} \tag{7}$$

$$1a \oplus (1-1)b = a, \tag{8}$$

$$\mu a \oplus (1-\mu)b = (1-\mu)b \oplus \mu a, \tag{9}$$

$$\lambda[\mu a \oplus (1-\mu)b] \oplus (1-\lambda)c = (\lambda\mu)a \oplus (1-\lambda\mu)\left[\frac{\lambda(1-\mu)}{1-\lambda\mu}b \oplus \frac{(1-\lambda)}{1-\lambda\mu}c\right] \tag{10}$$

where each axiom holds for all $a, b \in \mathcal{S}$ and all μ, λ. To give intuition to procedural mixture spaces, consider a lottery space. A mixture space relates mixtures between two lotteries to another lottery. A procedural mixture space is only required to do so if the mixture yields one of the two lotteries with certainty.

Definition 1. $\succsim$ is a weak order on $\mathcal{S}$ if for all $a, b, c \in \mathcal{S}$:
i) $a \succsim b$ or $b \succsim a$ or both, and
ii) $a \succsim b$ and $b \succsim c$ implies $a \succsim c$.

We use the symbols $\sim$ and $\succ$ to denote the symmetric and asymmetric parts of $\succsim$ and $a \not\sim b$ means either $a \succ b$ or $b \succ a$.

Definition 2. A function $U : \mathcal{S} \to \mathcal{R}$ represents $\succsim$ if

$$a \succsim b \Leftrightarrow U(a) \geq U(b) \tag{11}$$

U is called a representation.

Let $\succsim$ be a relation on a procedural mixture space $\mathcal{S}$. We assume:

Axiom 1. $\succsim$ is a weak order.

Axiom 2. For any $a, b, c \in \mathcal{S}$, the sets $\{a | \alpha a \oplus (1-\alpha)b \succsim c\}$ and $\{a | c \succsim \alpha a \oplus (1-\alpha)b\}$ are closed.

Axiom 3. If $a, a', b \in \mathcal{S}$, $\mu \in (0,1)$ then $a \succsim a' \Leftrightarrow \mu a \oplus (1-\mu)b \succsim \mu a' \oplus (1-\mu)b$.

These axioms are commonly named Rationality, Continuity, and Independence. We required a slight strengthening of Axiom 3 compared to Herstein and Milnor (1953). Reducibility allows them to generate our Axiom 3 from a weaker assumption.

We obtain the following representation theorem:

Theorem 1. *The relation $\succsim$ on the procedural mixture space $\mathcal{S}$ fulfills Axioms 1-3 if and only if there exists a continuous, real valued representation $U : \mathcal{S} \to \mathbb{R}$ such that for some $q \in \mathbb{R}, r \in \mathbb{R}_{++}$*

$$U(\mu a \oplus (1-\mu)b) = \mu^r U(a) + (1-\mu)^r U(b) + q \cdot H_r(\mu) \tag{12}$$

$$H_r(\mu) = \begin{cases} -\mu \ln \mu - (1-\mu)\ln(1-\mu), & r = 1 \\ -\mu^r + -(1-\mu)^r + 1, & r \neq 1 \end{cases} \tag{13}$$

In other words, we have obtained two possible representations. Either we obtain the expected utility of the mixture plus the Shannon (1948) entropy. Alternatively, we obtain expected utility under power-form probability distortions plus the Rényi (1961) entropy.

4 Sufficiency Proof

Proof. Neccessity is straightforward. We prove sufficiency.

Lemma 1. *$\mathcal{S}$ is topologically connected under the order topology.*

Proof. By connectedness of the unit interval and Axiom 2, the order topology on any $\{a|a = \mu s' \oplus (1-\mu)s''\}$ is connected. If $\mathcal{S}$ is not connected, then it is the union of two nonempty disjoint open sets $\mathcal{S}'$ and $\mathcal{S}''$. Take any element $s' \in \mathcal{S}'$ and $s'' \in \mathcal{S}''$. The order topology on $\mathcal{S}''' = \{a|a = \mu s' \oplus (1-\mu)s''\}$ is disconnected by the nonempty open sets $S' \cap S'''$ and $S'' \cap S'''$, yielding a contradiction. □

Lemma 2. *$\succsim$ is coseparable, i.e.,*

$$\mu a \oplus (1-\mu)b \sim \bar{\mu}\bar{a} \oplus (1-\bar{\mu})\bar{b} \tag{14}$$

$$\mu a' \oplus (1-\mu)b \sim \bar{\mu}\bar{a}' \oplus (1-\bar{\mu})\bar{b} \tag{15}$$

$$\mu a \oplus (1-\mu)b' \sim \bar{\mu}\bar{a} \oplus (1-\bar{\mu})\bar{b}' \tag{16}$$

jointly imply

$$\mu a' \oplus (1-\mu)b' \sim \bar{\mu}\bar{a}' \oplus (1-\bar{\mu})\bar{b}' \tag{17}$$

Proof. Using Commutativity and Associativity it is straightforward to show that

$$1/2[\mu a \oplus (1-\mu)b] \oplus 1/2[\mu a' \oplus (1-\mu)b'] \tag{18}$$

$$= 1/2[\mu a' \oplus (1-\mu)b] \oplus 1/2[\mu a \oplus (1-\mu)b'] \tag{19}$$

for any μ, a, b, a', b'. Using Axiom 3 together with the assumptions stated above then guarantee the desired result. □

Lemma 3. *$\succsim$ can be represented by continuous U, F such that*

$$U(\mu a \oplus (1-\mu)b) = F(a,\mu) + F(b, 1-\mu) \tag{20}$$

Proof. We either obtain the representation trivially, if $a \sim b$ for all $a, b \in \mathcal{S}$ or using the main theorem of Qin and Rommeswinkel (2018) which provides a representation theorem for weak orders on (open subsets of) $X \times Y \times Z$ with the representation $f(x,z) + g(y,z)$. Here we choose $\mathcal{X} = \mathcal{S}$, $\mathcal{Y} = \mathcal{S}$, and $\mathcal{Z} = (0,1)$ and endow the space with the product topology of the order topologies and the subspace topology of the reals. Thus, we will first obtain the representation on $\mu \in (0,1)$ and then extend it to $[0,1]$ using Axiom 2. To apply the main theorem of Qin and Rommeswinkel (2018), we require the following conditions.

Since we have a product space, the well-behavedness assumptions of Qin and Rommeswinkel (2018) are not needed and we also only need essentiality instead of strict essentiality. Essentiality requires that for at least some μ and some a, then there exist some b, b' such that $\mu a \oplus (1-\mu)b \not\sim \mu a \oplus (1-\mu)b'$ and for some a, b there exist some μ, μ' such that $\mu a \oplus (1-\mu)b \not\sim \mu' a \oplus (1-\mu')b$. The former is guaranteed by Axiom 3 and the exclusion of the case $a \sim b$ for all $a, b \in \mathcal{S}$. The latter is guaranteed by Axiom 2 and the exclusion of the case $a \sim b$ for all $a, b \in \mathcal{S}$.

Next, we need conditional independence of the $\mathcal{X}$ and $\mathcal{Y}$ dimensions for fixed $\mathcal{Z}$ dimension. This holds by Axiom 3.

Further, coseparability of the $\mathcal{X}$ and $\mathcal{Y}$ dimension given $\mathcal{Z}$ has been shown above.

Continuity of $\succsim$ holds in the order topology on $\mathcal{S}$. However, we require continuity in the product topology on $\mathcal{S} \times \mathcal{S} \times (0,1)$. By Axioms 2 and 3 the product topology is finer than the order topology on $\mathcal{S}$, guaranteeing continuity in the product topology.

Topological connectedness of the product topology follows from the connectedness of its components X, Y, and Z. The interval $(0,1)$ is obviously connected and each component S is connected in the order topology.

From Qin and Rommeswinkel (2018) then follows the existence of functions F and E such that $\succsim$ can be represented by

$$U(\mu a \oplus (1-\mu)b) = F(a,\mu) + E(b,\mu) \tag{21}$$

Commutativity of the mixture space guarantees that we can set $E(b,\mu) = F(b, 1-\mu)$. $\square$

Lemma 4. *$F(a,\mu) = A(\mu)U(a) + B(\mu)$ for all μ and all $a \in \mathcal{S}$.*

Proof. For fixed μ, $F(a,\mu)$ is a monotone transformation of U:

$$\begin{aligned}
& F(a,\mu) \geq F(b,\mu) && (22)\\
\Leftrightarrow\quad & F(a,\mu) + F(c,1-\mu) \geq F(b,\mu) + F(c,1-\mu) && (23)\\
\Leftrightarrow\quad & \mu a \oplus (1-\mu)c \succsim \mu b \oplus (1-\mu)c && (24)\\
\Leftrightarrow\quad & a \succsim b && (25)\\
\Leftrightarrow\quad & U(a) \geq U(b) && (26)
\end{aligned}$$

Therefore, we can write $F(a,\mu) = G(U(a),\mu)$. We obtain from Associativity:

$$\begin{aligned} & U(\mu a \oplus (1-\mu)[\lambda b \oplus (1-\lambda)c]) && (27) \\ =& G(U(a),\mu) + G(G(U(b),\lambda) + G(U(c),1-\lambda),1-\mu) && (28) \\ =& G(U(b),(1-\mu)\lambda) + G(G(U(a),\lambda) + G(U(c),1-\lambda),1-\mu) && (29) \end{aligned}$$

Noting that we have two continuous additive representations over $\mathcal{S} \times \mathcal{S}$ (specifically here the elements a and b), by the uniqueness of additive representations, we have that G must be positively affine in its first argument. Therefore $G(U(a),\mu) = A(\mu)U(a) + B(\mu)$. □

Lemma 5. $A(\mu) = \mu^r$, $r \in \mathbb{R}_{++}$.

Proof. We define $H(\mu) = H(1-\mu) = B(\mu) + B(1-\mu)$ Using Associativity, we can derive that

$$\begin{aligned} & A(\lambda)\left[A(\mu)U(a) + A(1-\mu)U(b) + H(\mu)\right] + A(1-\lambda)U(c) + H(\lambda) && (30) \\ =& A(\lambda\mu)U(a) + H(\lambda\mu) \\ & + A(1-\lambda\mu)\left[A\left(\frac{\lambda(1-\mu)}{1-\lambda\mu}\right)U(b) + A\left(\frac{1-\lambda}{1-\lambda\mu}\right)U(c) + H\left(\frac{\lambda(1-\mu)}{1-\lambda\mu}\right)\right] && (31) \end{aligned}$$

Consider a substitution a' for a under which the above condition needs to still hold. If $\Delta U = U(a) - U(a')$, then it follows that

$$A(\lambda)A(\mu)\Delta U = A(\lambda\mu)\Delta U \tag{32}$$

and therefore A is multiplicative. Using Cauchy's functional equation it is straightforward to derive that $A(\mu) = \mu^r$, $r \in \mathbb{R}$. By Axiom 3, $r > 0$. □

We obtain

$$\lambda^r H(\mu) + H(\lambda) = (1-\lambda\mu)^r \left[H\left(\frac{\lambda(1-\mu)}{1-\lambda\mu}\right)\right] + H(\lambda\mu) \tag{33}$$

and substitute: $\lambda = 1-x$ and $\lambda\mu = y$. Using $H(x) = H(1-x)$ we obtain:

$$(1-x)^r H\left(\frac{y}{1-x}\right) + H(x) = (1-y)^r H\left(\frac{x}{1-y}\right) + H(y) \tag{34}$$

with two types of solutions[2] (Ebanks et al., 1987):

$$\begin{aligned} A(\mu) = \mu; \quad & H(\mu) = -(\mu \ln \mu + (1-\mu)\ln(1-\mu))q + s && (35) \\ A(\mu) = \mu^r; \quad & H(\mu) = -(\mu^r + (1-\mu)^r - 1)q + s && (36) \end{aligned}$$

where $q, s \in \mathbb{R}$. From Axiom 2 and Connectedness, we also have that in both representations $s = 0$. We have therefore obtained the desired representation:

$$\begin{aligned} U(\mu a \oplus (1-\mu)b) =& A(\mu)U(a) + A(1-\mu)U(b) + q \cdot H(\mu) && (37) \\ \text{with } A(\mu) = \mu; \quad H(\mu) =& -\mu\ln\mu - (1-\mu)\ln(1-\mu) && (38) \\ \text{or } A(\mu) = \mu^r; \quad H(\mu) =& -\mu^r - (1-\mu)^r + 1 && (39) \end{aligned}$$

□

[2]We provide a more general reference than necessary here. Indeed, Ebanks, Kannappan, and Ng (1987) allow for each of the H components to be different functions and allows $A(\mu)$ to be a multiplicative vector-valued function. This is useful for potential extensions.

5 Example

Staying with the context of preferences over lotteries, we can give a small interesting application. Allais (1953) famously argued against the normative appeal of the independence axiom. As it turns out, even in the representation without probability distortions –expected utility plus weighted Shannon entropy– we can rationalize the Allais (1953) choices over the following lotteries:

Situation A:	Receive 100 millions with certainty.
Situation B:	Receive 500 millions with 10% probability. Receive 100 millions with 89% probability. Receive nothing with 1% probability.
Situation C:	Receive 100 millions with 11%. probability. Receive nothing with 89% probability.
Situation D:	Receive 500 millions with 10% probability. Receive nothing with 90% probability.

The Allais (1953) choices are to prefer A over B and D over C which is incompatible with expected utility theory. Suppose according to expected utility, a decision maker prefers B over A and D over C. Notice that the Shannon (1948) entropy of Situation A is much lower than the entropy of B. However, the entropy of C is almost the same as the entropy of D. Therefore, if a negative weight on the entropy is sufficiently high, the cost associated with the higher entropy may outweigh the expected utility in the choice over A and B without being pivotal in the choice between C and D, thus yielding the Allais (1953) choices.

6 Discussion

In this paper, the representation theorem has been obtained by changing the axioms and mixture space structure of Herstein and Milnor (1953) as little as possible. Using the methods of this paper, variations of its theme can be derived easily. In particular:

— Instead of weakening Reduction of Compound Lotteries to Associativity, one can alternatively require that the supports of a, a' are disjoint from the support of b in Axiom 2. This weakening of the Independence axiom may be more intuitive in some contexts, for example when measuring the quantitative diversity of the support of a lottery.

— Instead of a binary operator, it is possible to use any n-ary operator to compound more than two lotteries at once. Under a symmetric triadic mixture operator one can derive a stronger result, obtaining expected utility plus the Shannon (1948) entropy.

— We only used a special symmetric case of Qin and Rommeswinkel (2018). Moreover, the fundamental equation of information that was used in the last step of the proof has already been solved for much more general solutions in Ebanks et al. (1987). The more general results of these papers allow for weakening Commutativity, Independence, and Closure and even replacing $\mu \in [0,1]$ by vectors without changing much of the proof.

References

Allais, M. (1953, October). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica*, *21*(4), 503. doi:10.2307/1907921

Caplin, A., Dean, M., & Leahy, J. (2017, August). *Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy* (No. w23652). National Bureau of Economic Research. Cambridge, MA. doi:10.3386/w23652

Ebanks, B. R., Kannappan, P., & Ng, C. T. (1987, December). Generalized fundamental equation of information of multiplicative type. *Aequationes Mathematicae*, *32*(1), 19–31. doi:10.1007/BF02311295

Frankel, D. & Volij, O. (2011). Measuring school segregation. *Journal of Economic Theory*, *146*(1), 1–38.

Herstein, I. N. & Milnor, J. (1953, April). An Axiomatic Approach to Measurable Utility. *Econometrica*, *21*(2), 291–297. doi:10.2307/1905540. JSTOR: 1905540?origin=crossref

Nehring, K. & Puppe, C. (2009). Diversity. In *Handbook of Rational and Social Choice*. Oxford, UK: Oxford University Press.

Qin, W.-z. & Rommeswinkel, H. (2018). Quasi-separable Preferences.

Rényi, A. (1961). On measures of information and entropy. In *Proceedings of the fourth Berkeley Symposium on Mathematics, Statistics and Probability 1960* (pp. 547–561).

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal*, *27*(3), 379–423, 623–656.

Shorrocks, A. F. (1980, April). The Class of Additively Decomposable Inequality Measures. *Econometrica*, *48*(3), 613. doi:10.2307/1913126. JSTOR: 1913126?origin=crossref

Sims, C. A. (2003, April). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690. doi:10.1016/S0304-3932(03)00029-1

Suppes, P. (1996). The nature and measurement of freedom. *Social Choice and Welfare*, *13*(2), 183–200.

von Neumann, J. & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.