Conveying Value Via Categories

Online Appendix [Not For Publication]

Paula Onuchic ⓡ Debraj Ray

October 2020

1. Alternative Approach to Optimal Categorization

We provide an alternative description of the sender's optimal categorization and its associated pooling and separating intervals.

For any a and a', say that a' is a *successor* of a (or $a' \succ a$) if $a' > a$ and

$$\frac{S(x) - S(a)}{R(x) - R(a)} \geq \frac{S(a') - S(a)}{R(a') - R(a)} \tag{a.1}$$

for all $x \in (a, a']$. To understand this definition, suppose $a' \succ a$. Then conditional on placing all weight in $[a, a']$, R is "more optimistic" than S: it puts more weight on higher quality levels. So the sender might profit from treating $[a, a']$ as a pooling interval. For instance, in Figure 1 of the main text, R is the uniform distribution on $[0, 1]$ and S is as pictured. The left panel has full revelation, which yields the expected value of a under S. In the right panel, there is pooling between t_0 and its successor t_1, and separation everywhere else. The resulting value is the expectation under the distribution in red, which follows R in the pooling region, and S thereafter. Since the red distribution is a first-order improvement over S, the expected value under this scheme is strictly higher. Hence, we can view intervals $[a, a']$, with $a' \succ a$, as those over which it is profitable to pool.Indeed, Panel B suggests that we can make further gains by moving t_1 to the right towards even larger successors of t_0. That motivates the following definition: $[a, a')$ is a *maximal interval* if $a' \succ a$ and there is no pair $\{b, b'\} \neq \{a, a'\}$ such that $b \leqslant a$ and $b' \geq a'$, and $b' \succ b$. The following observation guarantees that maximality is a well-defined concept, and — crucially — that maximal intervals must be disjoint.

Observation A.1. (i) *If a has a successor, it has a largest successor.*

(ii) *If $a_k \downarrow a$ and each a_k has a successor, then so does a.*

(iii) *Suppose that $a < b < a' < b'$, with $a' \succ a$ and $b' \succ b$. Then $b' \succ a$. In particular, any two maximal intervals must be disjoint.*

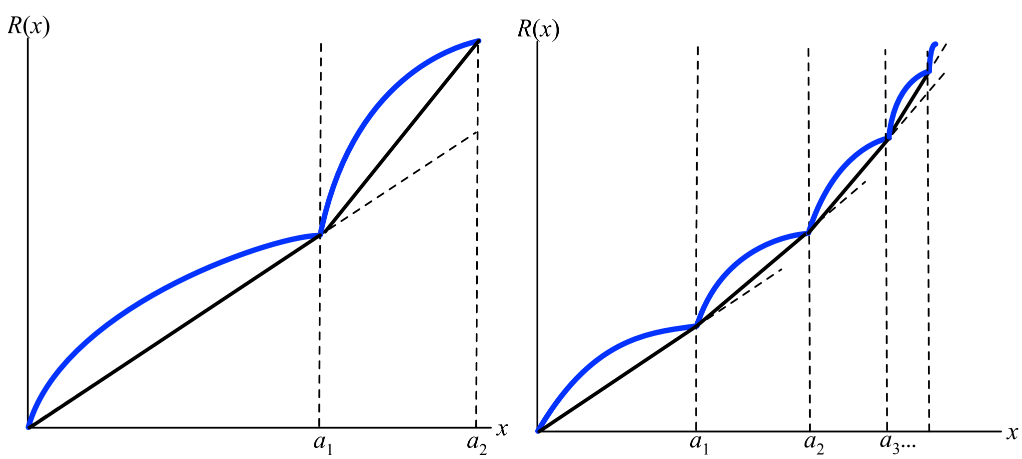
1.1. **A Characterization of A^*.** We can now state:

Theorem A.1. *Choose $A^* \in \mathcal{A}_R$ so that it pools in all maximal intervals and separates in all other intervals. Then A^* is well-defined, and solves the sender's problem.*

Theorem A.1 has two steps. First, A^* is well-defined. This requires Observation A.1, particularly part (iii), which guarantees that no two maximal intervals overlap. Second, the collection of maximal intervals chalk out precisely the lower convex envelope of the composite function $H = S \circ R^{-1}$. That is, the envelope is made out of straight-line segments joining the edges $(w, H(w))$ and $(w', H(w'))$, where $[w, w')$ is maximal, and by following $H(z)$ otherwise. The proof is completed by appealing to Theorem 1.

1.2. **A Procedure to "Compute" A^*.** The categorization A^* can be described by an algorithmic procedure if we assume:

[I] There are finitely many maximal intervals in $[\underline{a}, \bar{a}]$.



(a) Adjacent Pooling **(b)** Adjacent Pooling, Indefinitely

Figure A.1. R is uniform (not shown) and S is the blue curve. Panel A clarifies, using (a.1), why there could be two successive pooling intervals. Panel B shows that this could recur indefinitely, if S has the "appropriate" shape; Condition I rules this out.

Despite the absence of a precise genericity statement, it is clear that Condition I will fail only for singular configurations of R and S. In Figure A.1, R is uniform (not shown) and S is depicted by the "scalloped" blue line. In panel A, it is easy to see, using the definition in (a.1), that a_1 is the largest successor of $\underline{a}$, and by the same argument, that a_2 is the largest successor of a_1. But the situation could repeat itself, and could do so indefinitely, as depicted in Panel B. In this case Condition I fails. Conditional on R being uniform, this imparts a delicate asymptotic shape to S which would vanish on "small perturbations" of either function. This is why we view the Condition as "generic."

Under Condition I, the following procedure generates A^*, the verification of which is an easy consequence of Observation A.1 and Theorem A.1.

Proposition A.1. *Begin at $a_0 = \underline{a}$. If a_0 has a successor, then set a_1 equal to the* largest *such successor (well-defined by part (i) of Observation A.1), and designate $[a_0, a_1)$ to be a pooling interval. If a_0 has no successor, then a_1 is set to the first $a > a_0$ that* does *have a successor, or failing that, to $\bar{a}$. (This is guaranteed by part (ii) of Observation A.1.) In that case, designate $[a_0, a_1)$ to be separating. If $a_1 = \bar{a}$, close the open right parentheses in the intervals above, and end the procedure.*

Otherwise, repeat the same process starting from a_1 to find $a_2, a_s \dots$ Continue until $[\underline{a}, \bar{a}]$ is fully covered. Under Condition I, the procedure must end in finitely many steps.

1.3. **Connection of the Procedure to Rayo (2013).** Problem (3) in Rayo (2013) is stated as follows (using his notation):

$$\begin{aligned} &\max_{\varphi} \int_{\theta_L}^{\theta_H} \varphi(\theta) h(\theta) dF(\theta) \\ &\text{s.t. } \varphi \text{ is non-decreasing, and} \\ &\varphi(\theta) = \mathbb{E}\left[\theta' : \varphi(\theta') = \varphi(\theta)\right] \text{ for all } \theta. \end{aligned} \tag{a.2}$$

Say that h satisfies condition (NC) over the interval (θ_1, θ_2) if, for all $\theta \in (\theta_1, \theta_2)$:

$$\mathbb{E}\left[h(z) : z \in (\theta_1, \theta)\right] \geqslant \mathbb{E}\left[h(z) : z \in (\theta, \theta_2)\right] \tag{a.3}$$

Next, Rayo defines a collection of intervals $\mathcal{P}$ with typical element P, satisfying:

(1) For every interval (θ_1, θ_2) over which h is decreasing, there exists a $P \in \mathcal{P}$ such that $(\theta_1, \theta_2) \in \mathcal{P}$.

(2) Over every $P \in \mathcal{P}$, h satisfies condition (NC).

(3) Every $P \in \mathcal{P}$ is "maximal," i.e., there does not exist a larger interval $(\theta_1, \theta_2) \supset P$ over which h satisfies condition (NC).

He remarks that exactly one such collection exists and proves the following theorem:

Theorem A.2. *An optimal truthful filter is given by*

$$\varphi(\theta) = \begin{cases} \mathbb{E}[\theta \in P] \textit{ if } \theta \in P \textit{ for some } P \in \mathcal{P}, \\ \theta \textit{ otherwise.} \end{cases}$$

Now we want to show how to map this back into our model. Let the receiver prior be $R = F$, where F is the distribution introduced by Rayo above. Also, let the sender prior be S such that $dS(\theta) = h(\theta)dF(\theta)$. We can rewrite (a.3) as

$$\int_{\theta_1}^{\theta} h(z)\tfrac{dR(z)}{R(\theta)-R(\theta_1)} \geqslant \int_{\theta}^{\theta_2} h(z)\tfrac{dR(z)}{R(\theta_2)-R(\theta)} \;\Leftrightarrow\; \int_{\theta_1}^{\theta} \tfrac{dS(z)}{R(\theta)-R(\theta_1)} \geqslant \int_{\theta}^{\theta_2} \tfrac{dS(z)}{R(\theta_2)-R(\theta)}$$

$$\Leftrightarrow \tfrac{S(\theta)-S(\theta_1)}{R(\theta)-R(\theta_1)} \geqslant \tfrac{S(\theta_2)-S(\theta)}{R(\theta_2)-R(\theta)} \;\Leftrightarrow\; \tfrac{S(\theta)-S(\theta_1)}{R(\theta)-R(\theta_1)} \geqslant \tfrac{S(\theta_2)-S(\theta_1)}{R(\theta_2)-R(\theta_1)}$$

This establishes that for each h and F, there exists a sender and a receiver prior in our setup such that there is an equivalence between our definition of a successor and Rayo's definition of an interval that satisfies (NC). Therefore, the collection of intervals $\mathcal{P}$ defined above for some h and F coincides with the collection of pooling intervals picked up by the algorithm we proposed above for some S and R.

1.4. **Proofs for Alternative Characterization.** *Proof of Observation A.1*. Parts (i) and (ii) follows immediately from the assumed continuity of R and left-continuity of S. To show part (iii), we prove that

$$\frac{S(x) - S(a)}{R(x) - R(a)} \geq \frac{S(b') - S(a)}{R(b') - R(a)} \tag{a.4}$$

for every $x \in (a, b']$. We establish (a.4) in three steps.

Step 1: $\dfrac{S(b) - S(a)}{R(b) - R(a)} \geq \dfrac{S(b') - S(b)}{R(b') - R(b)}$.

Proof. Because $a' \succ a$ and $b \in [a, a']$, we have:

$$\frac{S(b) - S(a)}{R(b) - R(a)} \geq \frac{S(a') - S(a)}{R(a') - R(a)}. \tag{a.5}$$

At the same time, by elementary accounting,

$$\frac{S(a') - S(a)}{R(a') - R(a)} = \frac{R(b) - R(a)}{R(a') - R(a)} \left[\frac{S(b) - S(a)}{R(b) - R(a)}\right] + \frac{R(a') - R(b)}{R(a') - R(a)} \left[\frac{S(a') - S(b)}{R(a') - R(b)}\right] \tag{a.6}$$

Combine (a.5) and (a.6) and rearrange to get

$$\frac{S(b) - S(a)}{R(b) - R(a)} \geq \frac{S(a') - S(b)}{R(a') - R(b)}. \tag{a.7}$$

Because $b' \succ b$ and $a' \in [b, b']$,

$$\frac{S(a') - S(b)}{R(a') - R(b)} \geq \frac{S(b') - S(b)}{R(b') - R(b)} \tag{a.8}$$

Together, (a.7) and (a.8) imply Step 1.

Step 2: Inequality (a.4) holds for all $x \in (a, a']$.

Proof. Because $b' \succ b$ and $a' \in [b, b']$,

$$\frac{S(a') - S(b)}{R(a') - R(b)} \geq \frac{S(b') - S(b)}{R(b') - R(b)} \tag{a.9}$$

Invoking the accounting identity (a.6), and combining it with (a.9), we get:

$$\begin{aligned}
\frac{S(a') - S(a)}{R(a') - R(a)} &= \frac{R(b) - R(a)}{R(a') - R(a)} \left[\frac{S(b) - S(a)}{R(b) - R(a)}\right] + \frac{R(a') - R(b)}{R(a') - R(a)} \left[\frac{S(a') - S(b)}{R(a') - R(b)}\right] \\
&\geq \frac{R(b) - R(a)}{R(a') - R(a)} \left[\frac{S(b) - S(a)}{R(b) - R(a)}\right] + \frac{R(a') - R(b)}{R(a') - R(a)} \left[\frac{S(b') - S(b)}{R(b') - R(b)}\right] \\
&= \frac{R(b) - R(a)}{R(a') - R(a)} \left[\frac{S(b) - S(a)}{R(b) - R(a)}\right] + \left[1 - \frac{R(b) - R(a)}{R(a') - R(a)}\right] \left[\frac{S(b') - S(b)}{R(b') - R(b)}\right] \\
&\geq \frac{R(b) - R(a)}{R(b') - R(a)} \left[\frac{S(b) - S(a)}{R(b) - R(a)}\right] + \left[1 - \frac{R(b) - R(a)}{R(b') - R(a)}\right] \left[\frac{S(b') - S(b)}{R(b') - R(b)}\right] \\
&= \frac{S(b') - S(a)}{R(b') - R(a)}, \qquad \text{(a.10)}
\end{aligned}$$

where the first inequality uses (a.9), and the second inequality uses Step 1 along with the fact that $\frac{R(b)-R(a)}{R(b')-R(a)} \leqslant \frac{R(b)-R(a)}{R(a')-R(a)}$. At the same time, given $a' \succ a$,

$$\frac{S(x)-S(a)}{R(x)-R(a)} \geq \frac{S(a')-S(a)}{R(a')-R(a)} \quad \text{for every } x \in (a, a']. \tag{a.11}$$

Together, (a.10) and (a.11) imply Step 2.

Step 3: Inequality (a.4) holds for all $x \in (a', b']$.

Proof. Because $b' \succ b$ and $x > a' > b$,

$$\frac{S(x)-S(b)}{R(x)-R(b)} \geq \frac{S(b')-S(b)}{R(b')-R(b)} \tag{a.12}$$

Using (a.12) in an elementary accounting identity, we therefore see that

$$\begin{aligned}
\frac{S(x)-S(a)}{R(x)-R(a)} &= \frac{R(b)-R(a)}{R(x)-R(a)}\left[\frac{S(b)-S(a)}{R(b)-R(a)}\right] + \frac{R(x)-R(b)}{R(x)-R(a)}\left[\frac{S(x)-S(b)}{R(x)-R(b)}\right] \\
&\geq \frac{R(b)-R(a)}{R(x)-R(a)}\left[\frac{S(b)-S(a)}{R(b)-R(a)}\right] + \frac{R(x)-R(b)}{R(x)-R(a)}\left[\frac{S(b')-S(b)}{R(b')-R(b)}\right] \\
&\geq \frac{R(b)-R(a)}{R(b')-R(a)}\left[\frac{S(b)-S(a)}{R(b)-R(a)}\right] + \frac{R(b')-R(b)}{R(b')-R(a)}\left[\frac{S(b')-S(b)}{R(b')-R(b)}\right] \\
&= \frac{S(b')-S(a)}{R(b')-R(a)},
\end{aligned} \tag{a.13}$$

where the first equality is just accounting, the first inequality invokes (a.12), and the second inequality uses both Step 1 and $\frac{R(b)-R(a)}{R(x)-R(a)} \geq \frac{R(b)-R(a)}{R(b')-R(a)}$. (a.13) establishes Step 3.

Steps 2 and 3 together imply (a.4) for all $x \in (a, b']$, so $b' \succ a$. ■

Lemma A.1. *For any maximal interval with edges p and p', define the corresponding percentiles $q = R(p)$ and $q' = R(p')$. Then for all $z \in [q, q']$, recalling that $H = S \circ R^{-1}$,*

$$H(z) \geqslant \Phi(z, A^*) = H(q) + (z-q)\left[\frac{H(q')-H(q)}{q'-q}\right]. \tag{a.14}$$

Proof. By (3), we know that for every $x \in [p, p']$,

$$\Psi(x, A^*) = S(p) + [R(x)-R(p)]\left[\frac{S(p')-S(p)}{R(p')-R(p)}\right]. \tag{a.15}$$

Because $p' \succ p$, we have $\frac{S(x)-S(p)}{R(x)-R(p)} \geqslant \frac{S(p')-S(p)}{R(p')-R(p)}$, or

$$S(x) \geqslant S(p) + [R(x) - R(p)] \left[\frac{S(p') - S(p)}{R(p') - R(p)} \right]. \tag{a.16}$$

Convert to percentiles using $z = R(x)$ for any generic element $x \in [p, p']$, and combine (a.15) and (a.16) to obtain (a.14), as desired. ■

Proof of Theorem A.1. Let A^* be as described in the statement of the theorem. Observation A.1 assures us that every maximal interval is well-defined, and therefore so is A^*. By Theorem 1, it will suffice to prove that $\Phi(z, A^*) = \breve{H}(z)$ for all $z \in [0, 1]$. This is obviously true for $z = 0$ and $z = 1$. Suppose, on the contrary, that $\Phi(\hat{z}, A^*) \neq \breve{H}(\hat{z})$ for some $z \in (0, 1)$. Then, by Observation 1(ii),

$$\Phi(\hat{z}, A^*) > \breve{H}(\hat{z}). \tag{a.17}$$

We observe that $H(\hat{z}) > \breve{H}(\hat{z})$. If not, then by (a.17), we have $\Phi(\hat{z}, A^*) > H(\hat{z})$, which contradicts Lemma A.1. So $H(\hat{z}) > \breve{H}(\hat{z})$, and in particular, $\breve{H}(\hat{z})$ must lie on a line segment that connects two points on the graph of H and lies below H in between. Specifically, there is w and w', both in $[0, 1]$ with $w < \hat{z} < w'$, $H(w) = \breve{H}(w)$, $\breve{H}(w') = w'$, and for every $z \in (w, w')$, including $\hat{z}$,

$$\frac{H(z) - H(w)}{z - w} \geqslant \frac{\breve{H}(z) - H(w)}{z - w} = \frac{H(w') - H(w)}{w' - w}. \tag{a.18}$$

Define $t \equiv R^{-1}(w)$, $t' \equiv R^{-1}(w')$, $\hat{x} \equiv R^{-1}(\hat{z})$, and $x \equiv R^{-1}(z)$ for generic $z \in (w, w')$. Then $\hat{x} \in (t, t')$, and recalling that $H = S \circ R^{-1}$, (a.18) implies that

$$\frac{S(x) - S(t)}{R(x) - R(t)} \geqslant \frac{S(t') - S(t)}{R(t') - R(t)}$$

for all $x \in (t, t')$, which proves that $t' \succ t$, so that $[t, t')$ must be a weak subset of some maximal interval. Consequently, by (3) and the construction of A^*,

$$\Psi(x, A^*) = S(p) + [R(x) - R(p)] \left[\frac{S(p') - S(p)}{R(p') - R(p)} \right] \tag{a.19}$$

for some maximal interval $[p, p')$. Because $\hat{x} \in [t, t')$ and $\hat{x} \in [p, p')$, these intervals are not disjoint, but because $[p, p')$ is maximal, Observation A.1 (iii) implies that

$$[t, t') \subseteq [p, p'). \tag{a.20}$$

In particular, because $p' \succ p$, this means that

$$\frac{S(t)-S(p)}{R(t)-R(p)} \geqslant \frac{S(p')-S(p)}{R(p')-R(p)} \leqslant \frac{S(t')-S(p)}{R(t')-R(p)}. \tag{a.21}$$

Now we rewrite these relationships in percentile format. To this end, let $q = R(p)$ and $q' = R(p')$, then (a.19) becomes

$$\Phi(z, A^*) = H(q) + (z-q)\left[\frac{H(q')-H(q)}{q'-q}\right] = H(q) + (z-q)\alpha, \tag{a.22}$$

for all $z \in [q, q')$, where we define $\alpha \equiv \frac{H(q')-H(q)}{q'-q}$. Inequality (a.21) then becomes $\frac{H(w)-H(q)}{w-q} \geqslant \alpha \leqslant \frac{H(w')-H(q)}{w'-q}$, or equivalently

$$H(w) \geqslant H(q) + (w-q)\alpha \text{ and } H(w') \geqslant H(q) + (w'-q)\alpha, \tag{a.23}$$

Using (a.18), we must conclude that

$$\breve{H}(\hat{z}) \geqslant H(q) + (\hat{z}-q)\alpha = \Phi(z, A^*), \tag{a.24}$$

but (a.24) contradicts (a.17), and so completes the proof.

Proof of Proposition 5. Suppose by contradiction that $[a, b)$ and $[b, c)$ with $a < b < c$ are two adjacent *maximal* pooling intervals, with no separating interval in between.

Claim 1. $\dfrac{S(x)-S(b)}{R(x)-R(b)} \geqslant \dfrac{S(c)-S(b)}{R(c)-R(b)} > \dfrac{S(b)-S(a)}{R(b)-R(a)}$ for every $x \in (b, c)$.

Proof. The weak inequality is true because $[b, c)$ is pooling and so $c \succ b$. We show the strict inequality. If it is false, then

$$\frac{S(b)-S(a)}{R(b)-R(a)} \geqslant \frac{S(c)-S(b)}{R(c)-R(b)} \geqslant \frac{S(c)-S(a)}{R(c)-R(a)}, \tag{a.25}$$

where the second inequality is a consequence of the first. Combining (a.25) with the fact that $b \succ a$, we see that for any $x \in (a, b]$,

$$\frac{S(x)-S(a)}{R(x)-R(a)} \geqslant \frac{S(c)-S(a)}{R(c)-R(a)}. \tag{a.26}$$

Moreover, for any $x \in (b, c)$,

$$\text{(a.27)}\quad \begin{aligned}\frac{S(x)-S(a)}{R(x)-R(a)} &= \frac{R(x)-R(b)}{R(x)-R(a)}\frac{S(x)-S(b)}{R(x)-R(b)} + \frac{R(b)-R(a)}{R(x)-R(a)}\frac{S(b)-S(a)}{R(b)-R(a)} \\ &\geqslant \frac{R(x)-R(b)}{R(x)-R(a)}\frac{S(c)-S(b)}{R(c)-R(b)} + \frac{R(b)-R(a)}{R(x)-R(a)}\frac{S(b)-S(a)}{R(b)-R(a)} \\ &\geqslant \frac{R(c)-R(b)}{R(c)-R(a)}\frac{S(c)-S(b)}{R(c)-R(b)} + \frac{R(b)-R(a)}{R(c)-R(a)}\frac{S(b)-S(a)}{R(b)-R(a)} = \frac{S(c)-S(a)}{R(c)-R(a)},\end{aligned}$$

where the first inequality follows from $c \succ b$ and the second one from (a.25). But (a.26) and (a.27) together imply that $c \succ a$ is a maximal pair, which contradicts the fact $[a, b)$ is maximal. This contradiction establishes the Claim. ■

Claim 2. $\dfrac{S(b)-S(x)}{R(b)-R(x)} \leqslant \dfrac{S(b)-S(a)}{R(b)-R(a)}$ for every $x \in (a, b)$.

Proof. Because $[a, b)$ is maximal,

$$\text{(a.28)}\qquad \frac{S(x)-S(a)}{R(x)-R(a)} \geqslant \frac{S(b)-S(a)}{R(b)-R(a)}$$

for every $x \in (a, b)$. Moreover,

$$\text{(a.29)}\qquad \frac{S(b)-S(a)}{R(b)-R(a)} = \frac{R(x)-R(a)}{R(b)-R(a)}\frac{S(x)-S(a)}{R(x)-R(a)} + \frac{R(b)-R(x)}{R(b)-R(a)}\frac{S(b)-S(x)}{R(b)-R(x)}$$

(a.28) and (a.29) together imply the claim. ■

Let r and s be the densities of R and S respectively. Because Claim 1 establishes a uniform lower bound on $\frac{S(x)-S(b)}{R(x)-R(b)}$, we must conclude that

$$\text{(a.30)}\qquad \lim_{x\to b}\frac{S(x)-S(b)}{R(x)-R(b)} = \frac{s(b)}{r(b)} > \frac{S(b)-S(a)}{R(b)-R(a)}.$$

But Claim 2 implies that

$$\lim_{x\to b}\frac{S(b)-S(x)}{R(b)-R(x)} = \frac{s(b)}{r(b)} \leqslant \frac{S(b)-S(a)}{R(b)-R(a)},$$

which contradicts (a.30). ■

2. EXISTENCE IN GENERAL NONLINEAR MODEL

Consider the following generalization of our model, in which the receiver takes an action predicated on his conditional expectation (after learning the category membership of the object), and the sender obtains a payoff from that action. If quality x is revealed, the sender's payoff is $U(x)$. Otherwise, if x is pooled on $[a, b)$, the receiver's action is based on her posterior $A(x) = \mathbb{E}_R(y|[a, b))$, with resulting sender payoff $U(A(x))$.

The sender picks categorization $A \in \mathcal{A}_R$ to

$$\text{maximize} \int_{\underline{a}}^{\bar{a}} U(A(x))dS(x), \tag{a.31}$$

where the receiver's payoff $A(x) = x$ in any separating region, and $A(x) = \mathbb{E}_R(y|[a, b))$ in any pooling region $[a, b)$ of A.

Recall that a categorization is *minimal* if no pooling interval in it can be further refined without a drop in sender value. We can now state

Proposition A.2. *There is a minimal categorization A^* which maximizes sender value.*

Proof. We will make use of the following lemmas.

Lemma A.2. *For each $x < y$ in $[\underline{a}, \bar{a}]$, there is $\epsilon > 0$ such that if x and y belong to distinct categories under $A \in \mathcal{A}_R$, then $A(y) \geq A(x) + \epsilon$.*

Proof. For $z \in [x, y]$, let $\epsilon(z) \equiv \mathbb{E}_R(w|[z, y]) - \mathbb{E}_R(w|[x, z])$. It is easy to verify that $\epsilon(z)$ is strictly positive and continuous on $[x, y]$. So $\epsilon \equiv \min_{z\in[x,y]} \epsilon(z)$ is well defined and strictly positive. If x and y belong to distinct categories under A, then $A(y) \geq \mathbb{E}_R(w|[z, y])$ and $A(x) \leq \mathbb{E}_R(w|[x, z])$ for some $z \in [x, y]$. The claim follows. ■

View $\mathcal{A}_R$ as a subset of $\mathcal{A}$, the space of all nondecreasing, right-continuous functions from $[\underline{a}, \bar{a}]$ to itself, with $A(\bar{a}) = \bar{a}$, endowed with the weak topology, induced here by pointwise convergence of functions at all continuity points of the limit (Billingsley, 1968, p.22).

Lemma A.3. *$\mathcal{A}_R$ is compact.*

Proof. By Helly's Selection Theorem theorem (Billingsley 1968, p. 227), $\mathcal{A}$ is compact, so it suffices to prove that $\mathcal{A}_R$ is closed. To this end, let $A^n \in \mathcal{A}_R$ converges weakly to $A \in \mathcal{A}$. We must show that $A \in \mathcal{A}_R$. Because $A \in \mathcal{A}$, we already know A to be nondecreasing and continuous with $A^n(\bar{a}) = \bar{a}$.

Case 1. Suppose $A(x) = A(y) = c$ for some $x < y$. Because A is nondecreasing and right continuous, there is a maximal interval $[a, b)$ such that $A(z) = c$ for all $z \in [a, b)$. We claim that $c = \mathbb{E}_R(w|[a, b))$. Because (a, b) are all continuity points of A, $A^n(z) \to c$ for all $z \in (a, b)$. Note that for all large n, there is a pooling interval $[a^n, b^n)$ containing (a, b). For if not, there are two points z and z' in (a, b), with $z < z'$, and a subsequence of n along which z_1 and z_2 do *not* belong to the same pooling interval. But then, by Lemma A.2, there is $\epsilon > 0$ such that $A^n(z_2) \geq A^n(z_1) + \epsilon$ along that subsequence, which contradicts the fact that $A^n(z_1)$ and $A^n(z_2)$ both converge to c.

Next, we observe that if $z < a$ or $z > b$, then $z \notin [a^n, b^n)$ along any subsequence of n. For if there were some such z, then it must *also* be that $A^n(z) \to c$, which contradicts the maximality of (a, b) as defined.

From these arguments. it follows that $a^n \to a$ and $b^n \to b$. Moreover, for every $z \in (a, b)$, $A^n(z) = \mathbb{E}_R(w|[a^n, b^n))$ while $A^n(z) \to c$. Combining these points with the continuity of R, we must conclude that $c = \mathbb{E}_R(w|[a, b))$.

Case 2. Suppose that A is strictly increasing at $x < \bar{a}$, so that $A(x) \neq A(y)$ for all $x \neq y$. Because A is right continuous and $x < \bar{a}$, there is $y > x$ such that A is strictly increasing and continuous on (x, y). By weak convergence, $A^n(z) \to A^n(z)$ for all $z \in (x, y)$. We claim that for each such z, *either* z is separating along a subsequence of n, *or* if z is in a pooling interval $[a^n, b^n)$ for all large n, then $a^n \to z$ and $b^n \to b$. Certainly, if the "either" part of the claim fails for some z, then z is in a pooling interval $[a^n, b^n)$ for all large n. So if the "or" part is false, there is a subsequence of n — retain original notation — such that $a^n \to a$, and $b^n \to b$, where $a \leq z \leq b$, with at least one strict inequality. Notice that $A^n(z')$ *converges to the same number* $\mathbb{E}(w|[a, b))$ *for every* $z' \in [a, b)$, and also — because $z \in (x, y)$ — that $(x, y) \cap (a, b)$ is a nonempty and nondegenerate interval. That is a contradiction because $A^n(z')$ must converge to $A(z')$ for every $z' \in (x, y)$, and A is strictly increasing on that interval. ∎

To complete the proof of the proposition, view the integral in (a.31) as a functional defined on $\mathcal{A}_R$. Because each $A \in \mathcal{A}_R$ has countably many discontinuities and S is continuous, weak convergence of A^n to A implies convergence S-a.e. Moreover, the integrand is bounded. So by the dominated convergence theorem, this functional is continuous on $\mathcal{A}_R$. Because $\mathcal{A}_R$ is compact by Lemma A.3, a maximum exists.

Let $\mathcal{A}_R^*$ be the set of categorizations that maximize sender value. Just for what follows, view each separating element as a singleton set. For A and A' in this set, say that $A \succ A'$ if A has a finer partition than A'. Consider any totally ordered subset $\mathcal{B}$ of $\mathcal{A}_R^*$ ordered by $\succ$. This has a maximal element in the larger set $\mathcal{A}_R$, found by simply taking the intersection over all partitions in $\mathcal{B}$, and corresponds to some categorization $\bar{A}$, which must be the weak limit of some sequence drawn from $\mathcal{B}$. Using the continuity of the integral in (a.31), we must conclude that $\bar{A}$ also belongs to $\mathcal{A}_R^*$. Consequently, by Zorn's Lemma, $\mathcal{A}_R^*$ has a maximal element, and the proof is complete. ∎

References

Billingsley, P. (1968), "Convergence of Probability Measures," New York: Wiley.

Rayo, L. (2013), "Monopolistic Signal Provision," *BE Journal of Theoretical Economics* **13**, 27–58.