A Proof of Blackwell's Theorem*

Eduardo Perez-Richet†

November 15, 2017

Abstract

This note gives a new proof of Blackwell's celebrated result. The result is a bit stronger than the classical version since the action set and the prior are fixed, and only the utility of the decision maker varies. I show directly that a decision maker has access to a larger set of joint distributions over actions and states of the world if and only if her information improves in the garbling order.

1 Introduction

This note provides a proof of Blackwell's theorem (Blackwell, 1951, 1953). If a decision maker is identified with a prior on the states of the world, an action set, and a utility function over actions and states of the world, Blackwell's theorem says that an experiment π, that provides information about the state of the world, is preferred by every decision maker to an experiment π' if and only if π' is a garble of π. The proof I provide is relatively simple, and has the merit of making the intuitive point that the choice set of the decision maker is enlarged by moving from π' to π if and only if π' is a garble of π, which is absent from other proofs (Blackwell, 1951, 1953; Ponssard, 1975; Cremer, 1982; Leshno and Spector, 1992). Another advantage of this proof is that it varies only the utility of the decision maker, and not the prior or the action

*I thank Olivier Gossner and Shuo Liu, who pointed out some mistakes in the earlier versions of this proof, and Jeanne Hagenbach, for comments.

†École Polytechnique, e-mail: eduardo.perez@polytechnique.edu

set[1], so the difficult direction of the result (π more useful than π' implies that π' is a garble of π) is slightly stronger here than in Blackwell's original formulation.

2 Setup and Preliminary Results

There is a finite action set A, with $|A| \geq 2$, and a finite set of states of the world Ω. The prior is a probability distribution $p(\omega)$ in $\Delta(\Omega)$. The payoff of the decision maker is given by a real valued payoff function $u(a, \omega)$. Let U be the set of such payoff functions. An experiment is given by a random variable x, with finite support X and a joint distribution function π on $X \times \Omega$ with marginal $p(\cdot)$ on Ω. When the decision maker can observe the realization of x, but not that of ω, she has access to mixed strategies $\sigma(a|x)$, with $\sum_a \sigma(a|x) = 1$ for all x. Let $\Sigma(\pi)$ be the set of strategies accessible to a decision maker endowed with experiment π. Ultimately, the decision maker only cares about the joint distributions of actions and states of the world, $\varphi(a, \omega)$. Let $\Phi(\pi)$ be the set of joint distributions she can generate when endowed with π, or *policy space*. It is restricted by her lack of knowledge in the following way:

$$\Phi(\pi) = \Big\{\varphi(a, \omega) \ : \ \exists\, \sigma \in \Sigma(\pi),\ \varphi(a, \omega) = \sum_x \sigma(a|x)\pi(x, \omega)\Big\}.$$

It is easy to show that this set is a compact, and convex subset of $[0, 1]^{|A|\times|\Omega|}$. Then the problem of decision maker u endowed with experiment π is given by the following linear program

$$V(\pi, u) = \max_{\varphi \in \Phi(\pi)} \sum_{a,\omega} \varphi(a, \omega)u(a, \omega).$$

Definition 1 (Usefulness Order)**.** *I say that an experiment π is more useful than another experiment π', and write $\pi \succeq \pi'$, if all decision makers get a higher value when endowed with π than when they are endowed with π', that is,*

$$\pi \succeq \pi' \ \Leftrightarrow\ V(\pi, u) \geq V(\pi', u),\ \forall u \in U$$

[1] Leshno and Spector (1992) also fix the action set, but use a different proof technique based on matrices.

Definition 2 (Garbling Order). *I say that π' is a garble of π, and write $\pi' \trianglelefteq \pi$ if there exists a function $f : X \times X' \to [0,1]$ such that $\pi'(x',\omega) = \sum_x f(x,x')\pi(x,\omega)$ and $\sum_{x'} f(x,x') = 1$ for all x'. Two experiments π and π' are equivalent , denoted by $\pi \sim \pi'$, if $\pi \trianglelefteq \pi'$ and $\pi' \trianglelefteq \pi$.*

Note that this definition provides a different interpretation of $\Phi(\pi)$ as the set of garbles π' of π such that $|X'| \leq |A|$. For each function f satisfying the conditions of the definition, I will denote by $f \circ \pi$ the corresponding garble of π.

It is useful to prove a few basic results about garbles. The first of these results shows that, if one can observe two experiments $f_1 \circ \pi$ with realization space X_1, and $f_2 \circ \pi$ with realization space X_2, which are both garbles of π, then the experiment $f_1 f_2 \circ \pi$, with realization space $X_1 \times X_2$, is also a garble of π. This easily extends to a finite number of garbles.

Lemma 1. *Let $f_1 \circ \pi, \cdots, f_K \circ \pi$ be garbles of π. Then*

$$f_1 \cdots f_k \circ \pi \;\trianglelefteq\; \pi.$$

Proof. Let $g(x, x_1, \cdots, x_K) = f_1(x, x_1) \cdots f_2(x, x_K)$. Then

$$\sum_{x_1,\cdots,x_K} g(x, x_1, \cdots, x_K) = \sum_{x_1} f_1(x, x_1) \cdots \sum_{x_K} f_K(x, x_K) = 1.$$

so that $f_1 \cdots f_K \circ \pi$ is indeed a garble of π. □

The second result shows that if one is allowed to combine experiments from the set of binary garbles of π, i.e. garbles with support of size 2, then one can reconstitute all the information in π. The idea is simple: for each possible realization x of π, define the binary garble of π that returns 1 if x is realized and 0 otherwise; then combining all these garbles gives exactly the same information as π. Without loss of generality, I can fix the set $B = \{0, 1\}$, and denote the set of binary garbles of π by

$$\Gamma_b(\pi) = \Big\{\pi'(x',\omega) \;:\; \exists\, f : X \times B \to \mathbb{R}^+,\ \pi'(x',\omega) = \sum_x f(x,x')\pi(x,\omega) \text{ and } \sum_{x' \in B} f(x,x') = 1\Big\}.$$

Lemma 2 (Reconstitution from Binary Garbles). *Consider an experiment π with support X. Then there exists $|X|$ binary garbles $f_1 \circ \pi, \cdots, f_{|X|} \circ \pi \in \Gamma_b(\pi)$ such that*

$$f_1 \cdots f_{|X|} \circ \pi \sim \pi.$$

Proof. Let $x_1, \cdots, x_{|X|}$ be the elements of X. Then let $f_k(x, 1) = \mathbb{1}_{x=x_k}$ and $f_k(x, 0) = 1 - f_k(x, 1)$. The $|X|$ functions thus defined satisfy the conditions of Definition 2, so they generate $|X|$ binary garbles $f_1 \circ \pi, \cdots, f_{|X|} \circ \pi$. It is easy to see that $(f_k \circ \pi)\,(1, \omega) = \pi(x_k, \omega)$, therefore observing the combined outcomes of all the experiments $f_1 \circ \pi, \cdots, f_{|X|} \circ \pi$ intuitively allows to reconstitute the π. To show this formally consider the experiment $f_1 \cdots f_{|X|} \circ \pi$. Its realization space is $\{0, 1\}^{|X|}$, but in fact the only vectors that occur with positive probability are the vectors with 0 on every dimension except one. Let e^k be the vector with a 1 on the k-th dimension and zeros elsewhere. Then for every $k = 1, \cdots, |X|$

$$\left(f_1 \cdots f_{|X|} \circ \pi\right)\,(e^k, \omega) = \pi(x_k, \omega),$$

which proves that $f_1 \cdots f_{|X|} \circ \pi \sim \pi$. In fact, $f_1 \cdots f_{|X|} \circ \pi$ is exactly π, up to a recoding of the set X. □

3 Blackwell's Theorem

Blackwell's theorem says that the usefulness order and the garbling order are the same. I decompose the proof in two steps. First, I show by classical arguments that an experiment is more informative than another if and only if it generates a larger policy space in the set containment order. Second, I show that an experiment generates a larger policy space than another one if and only if the latter is a garbling of the former. The latter part is the novel one and it relies on the binary decomposition result. The idea for the difficult implication is to show that, if π leads to a larger policy space than π', then the binary reconstitution of π', which is informationally equivalent to π', is a garble of π.

Theorem 1 (Blackwell)**.** $\pi \succeq \pi' \;\Leftrightarrow\; \Phi(\pi) \supseteq \Phi(\pi') \;\Leftrightarrow\; \pi \trianglerighteq \pi'$.

Proof. I write one lemma for each step.

Lemma 3. $\pi \succeq \pi' \;\Leftrightarrow\; \Phi(\pi) \supseteq \Phi(\pi')$

Proof. $\Leftarrow$ is due to the fact that maximizing a function over a larger set always yields a higher value. $\Rightarrow$ is due to a separation theorem. Indeed, suppose that there exists a policy $\varphi(a,\omega)$ in $\Phi(\pi') \setminus \Phi(\pi)$. Then because $\Phi(\pi)$ is a closed convex set, the hyperplane separation theorem implies the existence of a vector $u \in U$ such that $\sum_{a,\omega} u(a,\omega)\varphi(a,\omega) > V(\pi, u)$. □

Lemma 4. $\Phi(\pi) \supseteq \Phi(\pi') \;\Leftrightarrow\; \pi \trianglerighteq \pi'$

Proof. $\Leftarrow$ is the more natural sense. Suppose that π' is a garble of π, and let $f(\cdot)$ be the associated garbling function. Let $\varphi \in \Phi(\pi')$ be the policy generated by the associated strategy $\sigma \in \Sigma(\pi')$. Consider the strategy

$$\hat{\sigma}(a|x) \equiv \sum_{x'} f(x,x')\sigma(a|x').$$

It is an element of $\Sigma(\pi)$ since

$$\sum_a \hat{\sigma}(a|x) = \sum_{x'} f(x,x') \sum_a \sigma(a|x') = 1.$$

And I can write

$$\begin{aligned}
\varphi(a,\omega) &= \sum_{x'} \sigma(a|x')\pi'(x',\omega) \\
&= \sum_{x'} \sigma(a|x') \sum_{x} f(x,x')\pi(x,\omega) \\
&= \sum_{x}\sum_{x'} f(x,x')\sigma(a|x')\pi(x,\omega) \\
&= \sum_{x} \hat{\sigma}(a|x)\pi(x,\omega),
\end{aligned}$$

which shows that $\varphi \in \Phi(\pi)$.

For $\Rightarrow$, suppose $\Phi(\pi') \subseteq \Phi(\pi)$. Then, since $|A| \geq 2$, I have $\Gamma_b(\pi') \subseteq \Phi(\pi') \subseteq \Phi(\pi)$. Then, by Lemma 2, I can pick $|X'|$ binary garbles $f_1 \circ \pi', \cdots f_{|X'|} \circ \pi'$ in $\Phi(\pi')$ that reconstitute π', so that $f_1 \cdots f_{|X'|} \circ \pi' \sim \pi'$.

Since $f_k \circ \pi' \in \Phi(\pi)$, $f_k \circ \pi'$ is a garble of π, so it is possible to find a function $g_k : X \times B \to [0,1]$ such that $g_k(x,0) + g_k(x,1) = 1$, and $f_k \circ \pi' = g_k \circ \pi$.

Consider the function $g : X \times B \to \mathbb{R}^+$ defined by $g(x,b) = \sum_k g_k(x,b)$. I can write

$$\boxed{\sum_x g(x,1)\pi(x,\omega)} = \sum_x \sum_k g_k(x,1)\pi(x,\omega) = \sum_k \sum_x g_k(x,1)\pi(x,\omega)$$
$$= \sum_k (\pi \circ g_k)\ (1,\omega) = \sum_k (\pi' \circ f_k)\ (1,\omega)$$
$$= \sum_k \pi'(x_k,\omega) = p(\omega)$$
$$= \boxed{\sum_x \pi(x,\omega)}$$

This can be seen as a system of $|\Omega|$ equations in $|X|$ unknowns, the $\big(g(x,1)\big)_{x\in X}$. It has at least one solution which is $g(x,1) = 1$, for all x. If $|X| \leq |\Omega|$, this is the unique solution, and therefore the functions $g_k(\cdot)$ must be such that $g(x,1) = 1$, for all x.

Suppose instead that $|X| > |\Omega|$. Then I show that the $g_k(\cdot)$ functions can be chosen so that $g(x,1) = 1$ for all x. To see this note first that because $g_k(x,1) + g_k(x,0) = 1$, the problem of finding the $g_k(\cdot)$ functions can be reduced to solving the system of $|X'| \times |\Omega|$ equations in $|X'| \times |X|$ unknowns given by $\sum_x \in X g_k(x,1)\pi(x,\omega) = \pi'(x,\omega)$, for each $k = 1,\ldots,|X'|$, and each $\omega \in \Omega$. Because $|X| > |\Omega|$ the system has multiple solutions. Adding the $|X|$ equations $g(x,1) = \sum_k g_k(x,1) = 1$ to the system leaves the number of equation below the number of unknowns, so we can indeed choose the $g_k(\cdots)$ functions so that $g(x,1) = 1$, for all x.

Knowing this, I show that $f_1 \cdots f_{|X'|} \circ \pi'$ is a garble of π as follows. For every $e \in \{0,1\}^{|X'|}$, I have

$$\begin{aligned}
\left(\pi' \circ f_1 \cdots f_{|X'|}\right)(e,\omega) &= \sum_k \mathbb{1}_{e=e^k} (\pi' \circ f_k)\ (1,\omega) \\
&= \sum_k \mathbb{1}_{e=e^k} (\pi \circ g_k)\ (1,\omega) \\
&= \sum_k \mathbb{1}_{e=e^k} \sum_x g_k(x,1)\pi(x,\omega) \\
&= \sum_x \underbrace{\sum_k \mathbb{1}_{e=e^k} g_k(x,1)}_{\equiv h(x,e)} \pi(x,\omega)
\end{aligned}$$

Hence, to prove that $f_1 \cdots f_{|X'|} \circ \pi'$ is a garble of π, I just need to show that $\sum_e h(x,e) = 1$. To see this note that $h(x,e) = 0$ if e is not one of the e^k vectors, and $h(x,e^k) = g_k(x,1)$. Therefore,

$$\sum_e h(x,e) = \sum_k g_k(x,1) = g(x,1) = 1.$$

□

□

References

Blackwell, D. (1951): "The Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman, University of California Press, Berkeley, 93–102.

——— (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Cremer, J. (1982): "A Simple Proof of Blackwell's "Comparison of Experiments" Theorem," *Journal of Economic Theory*, 27, 439–443.

Leshno, M. and Y. Spector (1992): "An Elementary Proof of Blackwell's Theorem," *Mathematical Social Sciences*, 25, 95–98.

Ponssard, J.-P. (1975): "A note on information value theory for experiments defined in extensive A Note on Information Value Theory for Experiments Defined in Extensive Form," *Management Science*, 22, 449–454.