Multi Agent Systems - Lecture 10

Victor Dweck

2007-01-10

Considerations

In this lecture Prof. Moshe Tennenholtz presented the work published on the paper On the Foundations of Qualitative Decision Theory (Brafman and Tennenholtz, 1996).Their aim in this paper is to initiate a work on the foundations of *qualitative decision making* similar to the one initiated by Leonard J. Savage [1] and continued by others on *quantitative decision making*.

The Basic Model

Let's define the agent's model and the concept of a policy, which formally describes the agent's behavior:

- S - the set of possible states of the environment
- A - the set of possible actions
- $L = 2^S$ - the set of possible information states; the power set of S
- $P : L \rightarrow A$ - an explicit representation for a policy, assigning an action for every possible information state

Definition 1 (maxmin representation)**.** *We say tha a Policy may have a maxmin representation if there exists a utility function on $A \times S$ such that for every $l \in L$ we have*

$$P(l) = \arg\max_{a \in A} \min_{s \in l} \{u(a, s)\}$$

[1] The Foundations of Statistics, 1972

Consider the maxmin representation of a policy given below:

	s_1	s_2	s_3	s_4	s_5
a	**1**	*2*	⃝-1	**2**	*4*
b	**1**	*3*	**0**	⃝-7	*4*
→ **c**	**5**	*2*	**5**	⃝0	*1*

The information state considered is $l = \{s_1, s_3, s_4\}$. Pay attention that according to the *maxmin criteria*, the last row is chosen as the action taken by the agent.

Binary Policies

Definition 2 (Closed Under Unions)**.** *A policy P is closed under unions if $P(u) = P(w) \to P(u \cup w) = P(u)$ for every $u, w \subseteq S$.*

Let's take as an example the binary policy of taking a sweater together with you, when going out. There are three possible states for the environment: $S = \{cold, ok, hot\}$ and two possible actions, take the sweater or not: {YES or NO}. The **explicit representation** of the problem is exponential $O(|A|^{|S|})$:

	{cold}	{ok}	{hot}	{c,o}	{c,h}	{o,h}	{c,o,h}
sweater	YES	NO	NO	YES	NO	NO	NO

The **maxmin representation** of the same problem can be defined in $O(|A| * |S|)$:

	{cold}	{ok}	{hot}
YES	3	2	0
NO	1	3	3

Rational Attitude If you believe that a *binary policy* defines a rational attitude, then you believe in a solution using the maxmin representation table. The calculation is very simple. Unfortunately, people don't behave according to this representation and the maxmin criteria.

Maxmin representability How can we guarantee that a policy can be represented using the maxmin qualitative format shown above?

Definition 3. *Consider $u, w \subseteq S$, we say that u dominates w $(u >_P w)$, if $P(u \cup w) = P(u)$ and $P(u) \neq P(w)$.*

Claim 1 (Transitivity)**.** *Let P be a binary policy closed under unions, and $u_1 >_P u_2 >_P \cdots >_P u_k$. If $P(u_1) \neq P(u_k)$ then $u_1 >_P u_k$.*

Proof. Assume that $u >_P v >_P w >_P x$. Let $A = \{a, a'\}$. Suppose w.l.o.g. that $P(u) = a$. Since $u >_P v >_P w >_P x$, $P(v) = a'$, $P(w) = a$ and $P(x) = a'$. We need to show that $P(u \cup x) = a$. Suppose in negation that $P(u \cup x) = a'$. Since $v >_P w$, $P(v \cup w) = a'$. Closure under unions implies that $P(u \cup x \cup v \cup w) = a'$. $u >_P v$ and $w >_P x$ imply that $P(u \cup v) = a$ and $P(w \cup x) = a$, respectively. Closure under unions implies that $P(u \cup v \cup w \cup x) = a$ which is a contradiction. Therefore, $P(u \cup x) = a$. □

Algorithm: Utility Function Assignment

The algorithm presented below assigns $u(a, s)$ for every element in $A \times S$, given a policy P.

1. Set i to 0 and T to S.
2. **[highest payoff]** For each $v \in S$ let $u(P(v), v) = n = |S|$.
3. **[Dominant States]** Let $U = \{s \in T \mid \nexists v \in T \text{ such that } v >_P s\}$.
4. $\forall a \in A, s \in U$ such that $a \neq P(s)$ assign $u(a, s) = i$.
5. Let $i = i + 1$ and $T = T \setminus U$.
6. If $T = \phi$, then terminate. Otherwise, go to step 3.

Claim 2 (Preferred Action)**.** *For every information state l, there is a unique preferred action.*

Proof. Actions are equally preferred only if they have the same worst case payoff. This can be the case only if the set U at this stage contains states s and v, on which P assigns different actions. However in this case $s >_P v$ or $v >_P s$ (based on the action assigned to $\{s, v\}$), and therefore s and v cannot be in U simultaneously. □

Claim 3 (Identical Behavior)**.** *Applying maxmin in the constructed utilities results in a behavior identical to the one defined by the original policy.*

Proof. Notice that it trivially holds for every l where $P(l) = 1$ (under certainty).

Suppose that maxmin chooses a on U, then for some $w \in l$, $u(w, a')$ is the worst utility on U (and in w P selects a).

Denote by S_a the states in U where a is selected and $S_{a'}$ the states where a' is selected. Our claim is that if $v \in S'_a$, then $w >_P v$.

To see this, consider our construction process: Since $u(w, a')$ is minimal, w must have been choosen in step 3 before all other elements in U.

Hence, there is no $v \in S'_a$ such that $v >_P w$. However, since the actions taken in the elements of $S_{a'}$ and w are different, then we must have $w >_P v$.Therefore, $w >_P v$ for every $v \in S'_a$.

This implies that for every $\{v, w\}$ the action choosen is a, since $U = S_a \cup S'_a \cup w$ and therefore closure under unions implies that P selects a on U. □