

С.П. Вовк

О ПОСТРОЕНИИ ФУНКЦИИ ПОЛЕЗНОСТИ ИГРОКА

При принятии рациональных решений лицо, принимающее решение (ЛПР), стремится к максимизации ожидаемой полезности. ЛПР каждому возможному исходу приписывает полезность из диапазона [0,1] или [0,50]. Функция полезности отражает его предпочтения к тем или др. исходам в зависимости от отношения к риску.

Если ЛПР безразличен к риску, решение принимается на основании ожидаемого выигрыша. Функция полезности $U(v)$ в этом случае является прямой с положительным наклоном, где v – прибыль, получаемая при различных исходах. Для принятия решения в случае небезразличия ЛПР к риску нужно обнаружить индивидуальное отношение к риску и построить индивидуальную функцию, воспользовавшись двухшаговой процедурой Дж. Неймана и О. Моргенштерна:

1. Присваиваются произвольные значения полезностей выигрышам для лучшего S и худшего исходов s . Например, $U(S)=1$ и $U(s)=0$. Тогда полезности промежуточных выигрышей будут находиться в [0,1].
2. Игроку предлагается на выбор: получить некоторую гарантированную сумму v или принять участие в игре: получить с вероятностью p S , а с вероятностью $(1-p) - s$. Вероятность p изменяется до тех пор, пока ЛПР станет безразличным в отношении к выбору между v и игрой. Тогда полезность гарантированной суммы определяется как $U(v)=p_0U(S)+(1-p_0)U(s)$.

Методология рационального принятия решений в условиях неопределенности, основанная на функции полезности индивида, опирается на пять аксиом (сравнимости, транзитивности, сильной независимости, измеримости, ранжирования), которые отражают минимальный набор необходимых условий непротиворечивого и рационального поведения игрока. Особо важное значение имеет аксиома ранжирования: чем больше вероятность крупного выигрыша, тем

больше игра “стоит”, т.е. тем большая плата потребуется за приобретение права участвовать в этой игре.

В качестве сферы применения рассмотрим игровое моделирование взаимодействия “педагог - студент”. В условиях отсутствия информации о вероятностях исходов для определения лучшей альтернативы ЛПР может воспользоваться нечетким отношением доминирования по полезности. Для попарного сравнения исходов необходимы нечеткие оценки их полезности, получаемые в результате отображения множества исходов на функцию полезности другого игрока. Таким образом, для непротиворечивого и рационального выбора каждому из игроков нужна функция полезности противника.

Рассмотрим построение функции полезности для обучаемого. При отсутствии информации о вероятностях исходов ранее представленная процедура напрямую применена быть не может, но педагогу достаточно просто представить оценки полезности исходов, зависящие от сложности задания и попытки его выполнения (см. табл. 1).

Таблица 1

Стратегия педагога по отношению к студенту с “низкими” способностями	Оценка исхода	Средняя полезность при низком ее первоначальном уровне	Низкая оценка полезности при высоком ее первоначальном уровне
Не применять К ₁	0	0	0
K ₁	1	0,72	0,6
K ₁ K ₁	0,8	(0,72+0,48)/2=0,6	(0,6+0,5)/2=0,55
K ₁ K ₁ K ₁	0,6	(0,72+0,48+0,26)/3=0,487	(0,6+0,5+0,4)/3=0,5
K ₁ K ₁ K ₂	[0,6;0,8] ∨ [0,8;1]	(0,72+0,48+0,4)/3=0,533	(0,6+0,5+0,5)/3=0,533
K ₁ K ₁ K ₃	[0,8;1]	(0,72+0,48+0,72)/3=0,64	(0,6+0,5+0,6)/3=0,566
K ₁ K ₂	1,33	(0,72+0,7)/2=0,71	(0,6+0,66)/2=0,63
K ₁ K ₂ K ₂	0,8	(0,72+0,7+0,4)/3=0,606	(0,6+0,66+0,5)/3=0,587
K ₁ K ₂ K ₃	[1;1,33]	(0,72+0,7+0,72)/3=0,713	(0,6+0,66+0,7)/3=0,653
K ₁ K ₃	[1,33;1,66]	(0,72+0,89)/2=0,805	(0,6+0,7)/2=0,65
K ₁ K ₃ K ₃	1,33	(0,72+0,89+0,5)/3=0,703	(0,6+0,7+0,6)/3=0,633

На основании табл. 1 определяем $U(S)=0,805$ и $U(s)=0$. Тогда $U(v)=0,805p_0=0,805$ и $p_0=1$.

По величине средней полезности отдельных исходов строим функцию полезности игрока как верхнюю границу области оценок (см. табл. 2).

При некотором исходе функция полезности перестает изменяться. Этую величину предлагается рассматривать как $U(v)$. Численное значение v определяется на основе табл. 1 как исход, имеющий $U(v)$.

Таблица 2

Исход	Полезность при небезразличном отношении к риску	Полезность при безразличном отношении к риску
0	0	0
0,6	0,533	0,25
0,8	0,64	0,4
1	0,72	0,6
1,33	0,805	0,8
1,66	0,805	1

Анализ данных табл.2 показывает, что, начиная с $v=1,33$, игрок рассматриваемого класса предпочитает гарантированную сумму и отказывается от проведения игры. Поскольку в рассматриваемом классе игр всегда $U(s)=0$, то $p_0=U(v)/U(S)$, а в рассматриваемой ситуации, представленной в виде табл. 1, $p_0=1$.

Исследование многообразия типовых позиционных деревьев, моделирующих взаимодействие “педагог-студент”, зависит от способностей обучаемого и размерности шкалы оценки исхода. Следовательно, для обоснования предложенного способа построения функции полезности требуется исследовать:

1. влияние величины полезности, назначенной заданию с учетом попытки выполнения;
2. влияние способностей обучаемого на величину v .

В табл. 1 представлена игра с простейшим (K_1) уровнем сложности во время первой попытки общения между педагогом и студентом. Во время второй и третьей попыток в рассматриваемом дереве используются средний (K_2) и высокий (K_3) уровни сложности заданий.

Для построения функции полезности педагога при малом количестве “состояний природы” (состояний студента) проводилось анкетирование среди представителей класса с одинаковой лояльностью к ошибкам студента. При обработке данные должны были объединяться по классам индивидов, ведущих обучение по разным дисциплинам у экспериментального потока, но обладающих одинаковой лояльностью. Для обозначения класса использовался крите-

рий, который оказался определяющим для группы и был определен при анонимном опросе потока.

Представителям класса с одинаковой лояльностью к ошибкам, было предложено оценить полезность каждого из уровней сложности во время первой, второй, третьей ошибок общения. Поскольку выставляемые оценки различались, то в расчетах использовались оценки полезности, названные не менее 70% опрошенных. Фактически была использована процедура аналогичная определению дидактических характеристик учебных заданий.

Оценки полезности педагога, назвавшего низкое первоначальное значение полезности при 1-ой попытке, в случае 3-х уровней сложности заданий представлены в табл. 3.

Функция полезности строилась с использованием информации о предпочтениях по каждой из имеющихся тактик. Поскольку каждая из тактик представляет определенную последовательность уровней сложности, то полезность тактики может быть оценена как сумма полезности последовательно выполняемых заданий определенных уровней сложности во время соответствующих попыток. Поскольку каждая тактика связана с интервалом исходов на множестве исходов, то существуют полностью или частично пересекающиеся интервалы различающихся оценок. Поэтому полезность на интервале перекрытия оценок должна быть, по крайней мере, такой же, как у самой слабой (по предпочтению) оценки на этом интервале, т.е. $U(\omega) = \inf_{\omega \in \Omega} \{u_f\}$, где u_f - предпочтительность тактики a_f .

Таблица 3

Низкая оценка полезности при 1-ой попытке				Высокая оценка полезности при 1-ой попытке			
Уровень сложности	1-я попытка	2-я попытка	3-я попытка	Уровень сложности	1-я попытка	2-я попытка	3-я попытка
K ₁	0,72	0,48	0,36	K ₁	0,6	0,5	0,4
K ₂	-	0,72	0,4	K ₂	-	0,66	0,5
K ₃	-	0,89	0,5	K ₂	-	0,7	0,6

Анализ данных табл. 3 показывает, что низкий первоначальный уровень оценки полезности не влияет на величину v , за которую обучаемый готов отказаться от игры.

Проанализируем влияние способностей к обучению на величину гарантированной суммы. Оценки средней полезности исходов, зависящие от сложности задания и попытки его выполнения, для студента с “высокими”, “средними”, “слабыми” способностями к обучению представлены в табл. 1, 4, 5.

Таблица 4

Стратегия ЛПР по отношению к студенту со “средними” способностями	Оценка исхода	Средняя полезность
Не применять K_2	0	0
K_2	1,33	0,64
$K_2 K_2$	1	(0,64+0,49)/2=0,565
$K_2 K_2 K_2$	0,8	(0,64+0,49+0,18)/3=0,437
$K_2 K_2 K_3$	[0,8;1]	(0,64+0,49)/2=0,506
$K_2 K_3$	[1,33;1,66]	(0,64+0,56)/2=0,6
$K_2 K_3 K_3$	[1;1,33]	(0,64+0,49+0,46)/3=0,53

На основании данных табл. 4 определяем $U(v)=0,64p_0=0,6$ и $p_0=0,9375$.

Видно, что при выигрыше 1,33 наступает безразличие между v и игрой.

Таблица 5

Стратегия ЛПР по отношению к студенту со “высокими” способностями	Оценка исхода	Средняя полезность
Не применять K_3	0	0
K_3	1,66	0,72
$K_3 K_3$	1,33	(0,72+0,89)/2=0,805
$K_3 K_3 K_3$	1	(0,72+0,89+0,5)/3=0,703
$K_2 K_2 K_3$	[0,8;1]	(0,64+0,49)/2=0,506
$K_2 K_3$	[1,33;1,66]	(0,64+0,56)/2=0,6
$K_2 K_3 K_3$	[1;1,33]	(0,64+0,49+0,46)/3=0,53

Согласно табл. 5 не наступает безразличие между гарантированным выигрышем и игрой.

Анализ данных табл. 1, 4, 5 подтверждает, что чем выше “обучаемость” студента, тем выше размер гарантированной суммы, за которую игрок готов отказаться от игры. Вывод соответствует рациональному поведению игрока.

Следовательно, основным соображением при выборе типового позиционного дерева выступает степень обучаемости студента, что в принципе соответствует интуитивным соображениям.

Предлагаемая процедура построения функции полезности может использоваться для определения нечетких оценок полезности исходов игроков, используемых в играх с нечеткими исходами и предпочтениями игроков.