Курс:

ТЕОРИЯ ИГР

Тема 10:

РАФИНИРОВАНИЕ РАВНОВЕСИЙ ДЛЯ РАЗВЕРНУТОЙ ФОРМЫ.

ТЛЕУЖАНОВА МАНАТЖАН АШИМКУЛОВНА

Взгляд вперед. Мы обсудили равновесия Нэша, центральное понятие некооперативной теории игр. Что нас ждет дальше? Мы видели, что равновесия (во всяком случае смешанные) «всегда» существуют. Однако их может быть слишком много, и среди них могут быть дурацкие. Естественно возникает вопрос – нельзя ли наложить на равновесия некоторые более жесткие требования, которые отсекали бы «плохие, шаткие» равновесия? Это тема рафинирования Нэша и совершенных равновесий.

Вторая тема – модификации игры или понятия равновесия, которые расширяют возможности решений. К этому относятся игры с сообщениями, коррелированные равновесия и повторяющиеся игры.

Наконец, нужно обсудить важное понятие игр с несовершенной информацией и соответствующее понятие Байесова равновесия.

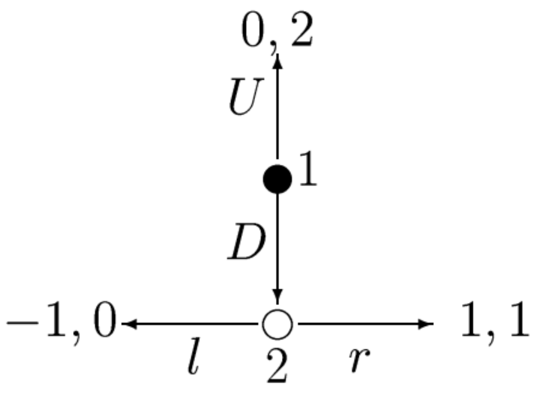
Строгие равновесия. Вспомним один пример с дурацкими равновесиями: голосование простым большинством при двух кандидатах. Все любят *A*, но голосуют за *B*. Это равновесие, но очень дурацкое. Конечно, индивидуальное переключение с *B* на *A* не улучшает исход, но и не ухудшает же! Стратегия *A* доминирует *B* (слабо, конечно). В этой частной ситуации у игрока два лучших ответа – *A* и *B*, и он выбрал не самый лучший. Подобных казусов, можно надеяться, не будет, если равновесие таково, что лучший ответ однозначен для каждого игрока.

Это хорошее понятие, единственный недостаток – что оно исключает смешанные равновесия, а поэтому не всегда и существует. Можно предложить некоторое его смягчение.

Определение. Профиль стратегий $(s_i^*) \in S_N$ называется *строгим равновесием*, если для любого игрока i стратегия s_i^* является наилучшим ответом не только на s_{-i}^* но на любой набор s_{-i}, где $s_j \in Best_j(s_{-j}^*)$.

Недостаток опять в том, что такие равновесия не всегда существуют; пример – орлянка. Ниже мы обсудим несколько более удачных понятий рафинирования.

Неправдоподобные угрозы. Рассмотрим следующую игру:



Эту игру можно понимать как простейшую форму ситуации «входа в отрасль». Фирма 1 (*агрессор*) может не входить (*U*) в некую отрасль занятую монополистом 2. Либо может войти (*D*), и тогда «наседка» 2 может ответить либо войной цен (l), либо пододвинется (r) и они поделят рынок. Что же произойдет в этой игре?

Алгоритм Цермело-Куна дает равновесие (D, r) с выигрышами (1,1). Однако есть другое равновесие, а именно (U, l) с выигрышами (–1,0). Здесь «наседка» как бы угрожает агрессору: «Если ты войдешь, то я буду сражаться, и ты получишь –1». Если агрессор воспримет эту угрозу как реальную, он воздержится от входа. Однако скорее всего он воспримет эту угрозу как неправдоподобную. Дело в том, что если первый все-таки войдет, то второму игроку будет невыгодно приводить угрозу в исполнение. Понимая это, первый игрок скорее всего не поверит в угрозу и проигнорирует ее.

В чем же дефективность стратегии l? В том, что она находится вне равновесного пути и поэтому не подвергается реальному испытанию на оптимальность. Формально она является наилучшим ответом на стратегию первого U, но фактически это не наилучший ход в состоянии 2. Это проявляется и в том, что стратегия l слабо доминируется стратегией r.

Этот пример поднимает еще один интересный вопрос. Угроза может быть сообщена только если есть сообщения; мы же предполагаем, что игра идет молча. Как же тогда 1-й может узнать, что собирается делать 2-й в позиции, которая лежит вне пути игры? Скорее это не то, что собирается делать 2-й, а то, что 1-й думает об этом. Но тогда получается, что стратегии 2-го выбирает 1-ый!

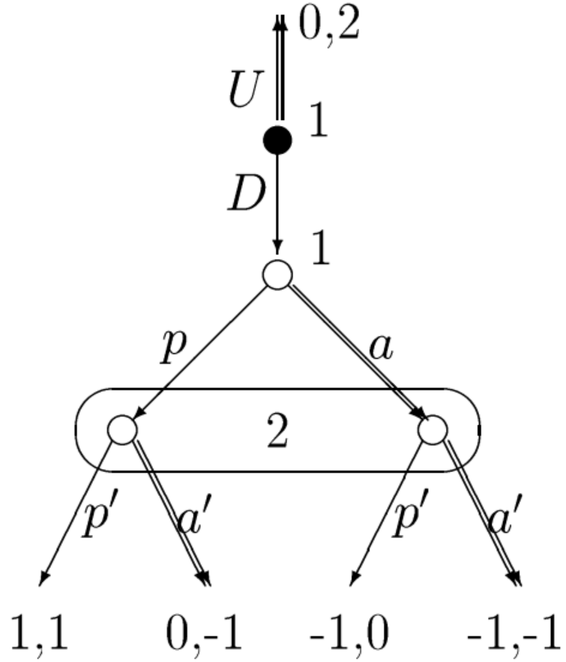
Неправдоподобные обещания. Быть может, есть похожие примеры и с неправдоподобными обещаниями, хотя мне не удалось придумать такой пример с двумя игроками. Однако подобная ситуация часто встречается в экономическом контексте. Например, фирма может пообещать покупателям, что она выпустит ограниченное число изделий и они будут редкостью. Покупатели заинтересованы купить раритет. Но где гарантии, что фирма не начнет в погоне за прибылью делать дополнительные партии изделий? Отсутствие таких гарантий (даже при честных намерениях фирмы) может привести к отказу покупателей, что плохо и тем, и другим. Но не так-то просто убедительно связать себя. Пример с киднэпингом: Джон похитил Мэри и готов отпустить ее за умеренный выкуп, но боится, что после освобождения она заложит его. Мэри чистосердечно обещает ему, что не продаст, но где гарантии? И если ей не удастся как-то убедить Джона и дать твердые гарантии, ему придется избавляться от нежеланного свидетеля.

Рафинирование равновесий Нэша. Наличие малоправдоподобных равновесий заставило теоретиков искать более сильные условия на равновесия. Эта программа известна под названием рафинирования, утоньчения, усовершенствования. Скажем сразу, что окончательного, приемлемого во всех случаях ответа получить не удалось; имеется несколько альтернативных подходов. Общее в них то, что предлагается принимать в расчет и невероятные события, которые имеют нулевую вероятность при движении по равновесному пути.

Могло бы показаться – зачем учитывать поведение в невероятных состояниях, которые не реализуются в равновесии? Дело, однако, в том, что часто нелепые равновесия возникают именно потому, что некто ведет себя иррационально вне равновесного пути, как в приведенном выше примере со входом в отрасль. Причем это состояние только потому и не встречается на равновесном пути, что там второй игрок ведет себя иррационально.

Совершенство относительно подыгр. Идея неправдоподобных угроз работает не только в случае игр с совершенной информацией (где она приводит к обратной индукции), но в некоторых других случаях. Например, когда мы обсуждали конкурентное равновесие, мы предполагали, что потребители (после выбора цен) ведут себя «конкурентно», не пытаясь своим нерациональным поведением повлиять на цены. Или, вернемся к примеру с входом в отрасль. Допустим, что если агрессор решил входить, у него тоже есть возможность вести себя пассивно или начать активную борьбу с «наседкой». Т.е. игра выглядит как на рисунке ниже.

И если второй настроен активно, то первому лучше не входить (а в случае входа - все равно как вести). А с другой стороны, в игре 2×2, начинающейся после входа, у обоих игроков есть доминирующие стратегии (p, p'), и это единственное равновесие в этой подыгре. Поэтому агрессор должен вроде понимать, что с случае его входа все закончится мирным исходом (1, 1), и смело входить.



Идея совершенства – оптимальность поведения во всех подыграх, чтобы равновесие оставалось равновесием начиная с любой позиции, независимо от того, сколь ничтожна вероятность попадания в эту позицию. Однако тут же возникает вопрос, что считать подыгрой. Казалось бы, возьмем любую вершину и включим все, что идет после нее. Но ведь имеются еще информационные множества, и они не должны выводить за пределы нашей «подыгры».

Более точно, пусть t – вершина дерева игры, и $G(t)$ – множество вершин, следующих за t. Если для любой вершины $t' \in G(t)$ все информационное множество $h(t')$ лежит внутри $G(t)$, мы говорим, что t определяет подыгру $G(t)$ с началом в t.

Профиль поведенческих стратегий называется совершенным равновесием (относительно подыгр, *subgame perfect equilibrium*), если его ограничение на любую подыгру является равновесием Нэша. Это очень интересное понятие было предложено Селтеном в 1965 г.

Свойства совершенных равновесий. Для игр с совершенной информацией совершенство совпадает с алгоритмом Цермело-Куна (или обратной индукцией). Более того, это одновременно учитывает и возможную неоднозначность алгоритма, и возможность несовершенства информации (т.е. возможность информационных множеств).

Так как исходная игра тоже подыгра, совершенное равновесие является равновесием Нэша. Но, конечно, не любым, как показывают приведенные выше примеры.

Далее, подыгра любой подыгры игры G является подыгрой игры G. Поэтому совершенное равновесие остается совершенным равновесием и в любой подыгре. И вообще, подыгры игры G сами образуют дерево. Это обстоятельство подсказывает, как искать все совершенные равновесия. Нужно, как в методе обратной индукции, начинать с конца, в каждой терминальной подыгре находить равновесия, и ставить в начало этих подыгр один из получающихся исходов. Помимо прочего это дает теорему существования совершенных (к подыграм) равновесий.