Алгоритмы со вторым шансом

Иван Гунич и Андрей Законов

21 апреля 2008 г.

Содержание

1. Идея алгоритмов с апелляциями

1.1. Постановка задачи

Афинная максимизация — это основной метод для разработки правдивых механизмов. Результаты, полученные этим способом, не оставляют надежды для большинства сложных проблем. Если P≠NP, тогда все механизмы получаются вырожденными и неразумными, и VCG не очень-то реализуешь.

Nisan и Ronen предложили метод, который позволил обойти эту проблему. Рассмотрим VCG-based механизм. В нём, как обсуждалось ранее, польза для агента равна суммарной пользе всех агентов. Единственный причиной, по которой агент мог соврать по поводу своего типа — это улучшение общего результата, то есть эгоистичный агент неизбежно должен оптимизировать всеобщее счастье. Поэтому, если агент не может улучшить текущий алгоритм, то у него нет причины врать. Воспользуемся этим свойством и сконструируем «почти» правдивый механизм.

Для любого алгоритма и соответствующей задачи оптимизации мы определим основанный на нем *механизм второго шанса*. Этот механизм является модификацией VCG механизма, где агентам разрешается подавать *апелляции* на используемый в механизме алгоритм распределения. Неформально, апелляция — это функция, которая говорит: «если у нас были типы $(v_1, \dots, v_N)$, то нам лучше было бы сказать $(v'_1, \dots, v'_N)$, тогда всеобщее счастье было бы больше». Агент должен суметь найти более эффективный для всеобщего счастья путь, чем нашёл субоптимальный алгоритм, который работает в механизме.

Обоснуем почему правдивость выгодна в алгоритме со вторым шансом. Механизм проверяет апелляции и удовлетворяет их, если действительно всеобщее счастье получается больше. Тогда получается, что агенту выгодно говорить правду, потому что он может попробовать соврать в апелляции, и тем самым перепроверить, будет ли выгодно врать в этом случае. То есть агент сможет и правду сказать, и перепроверить, а не соврать ли.

1.2. Формальная постановка

Напомним определения. Пусть есть N агентов, у агента i есть функция v_i, определяющая его тип. Агент i в частном порядке знает v_i.

Опр. *Ценности квазилинейные: если при результате o механизм платит агенту p_i, то общая польза агента*

$$u_i = v_i(o) + p_i.$$

Каждый агент хочет максимизировать u_i. Механизм хочет быть эффективным, то есть максимизировать $V(v, o) = \sum_{i=1}^{N} v_i(o)$.

Опр. *Прямой механизм — это механизм, где:*

- *спрашивают типы $w = (w_1, \ldots, w_N)$,*
- *а потом вычисляют исход $g(w) \in \emptyset$*
- *и платежи каждому участнику $p(w) = (p_1, \ldots, p_N)$.*

Здесь, как и раньше, агенты могут врать, т.е. априори $w_i \neq v_i$.

Опр. *Механизм правдивый, если каждый агент для каждого возможного своего типа v_i и каждого возможного множества ставок других агентов w_{-i} максимизирует свою u_i, если объявляет свою настоящую функцию ценности v_i. То есть мы будем говорить о правдивости в доминантных стратегиях.*

Опр. *Механизм (g, p) является VCG-механизмом, если:*

- *$g(w)$ максимизирует общую ценность относительно w, т.е. $g(w) \in \max_o V(w, o)$;*
- *выплаты вычисляются по VCG-формуле:*

$$p_i(w) = \sum_{j \neq i} w_j(g(w)) + h_i(w_{-i}),$$

  где h_i — некая (произвольная) функция от w_{-i}.

Зачастую функцию распределения g за разумное время не посчитать, то есть не решить задачу оптимизации. Примеры обсуждались в предыдущих лекциях.

Поэтому рассмотрим механизм основанный на произвольной функции g.

Опр. Механизм (g, p) является VCG-механизмом, основанным на функции g, если выплаты вычисляются по VCG-формуле:

$$p_i(w) = \sum_{j \neq i} w_j(g(w)) + h_i(w_{-i}),$$

где h_i — некая (произвольная) функция от w_{-i}.

То есть берется функция распределения и выплаты строятся соответственно. Выплаты здесь не равны выплатам VCG, т.к. мы подставляем не оптимальную функцию, а функцию g.

1.3. Механизм второго шанса и его основные свойства

Опр. *Обозначим через* $\Theta = \prod_i \Theta_i$ *множество типов всех агентов. Апелляция — это частичная функция* $l : \Theta \to \Theta$.

То есть, функция агента говорит о следующем: «а вот когда типы $\Theta = (\theta_1, \dots, \theta_N)$, алгоритм вашего механизма g даёт больше всеобщего счастья, если вместо Θ ему дать $l(\Theta)$». Таким образом апелляция позволяет агенту предложить свой способ увеличения всеобщего счастья. Если θ не принадлежит $dom\ l$, то это означает, что агент не может предложить вариант, который увеличит всеобщее счастье.

Механизм второго шанса является модификацией VCG механизма, в котором агентам дается возможность подавать функции апелляции.

Этот механизм с апелляциями работает следующим образом:

1. **Перед исполнением** Публикуется алгоритм g и временной лимит T на работу апелляций (иначе апелляция могла бы просто решать VCG, но мы бы никогда не смогли такую апелляцию обсчитать).

2. **Заявление** Каждый агент подаёт свой тип w_i (ставку) и функцию апелляции l_i.

3. **Распределение** Механизм подсчитывает $g(w)$, $g(l_1(w)), \dots, g(l_N(w))$ и выдаёт тот ответ, который максимизирует всеобщее счастье.

4. **Оплата** Если $\hat{o}$ — избранный исход, то платежи получаются как
$$p_i = \sum_{j \neq i} w_j(\hat{o}) + h_i(w_{-i}, l_{-i}),$$
где h_i — произвольная функция.

Агенты отсылают программы представляющие функции апелляции. Затем эти программы выполняются механизмом. Механизм имеет право прекратить вычисление любой апелляции по истечении временного лимита T. Таким образом, мы предполагаем что все апелляции должны быть завершены к определенному моменту. Проблема оптимального выбора лимита времени являются отдельной задачей.

Так как h_i не влияют на действия агентов, они не принимаются в рассмотрение, когда агенты делают свой выбор. Для удобства, будем пока предполагать, что $h_i = 0$.

Опр. *Действием агента в механизме второго шанса называется пара* (w_i, l_i), *где* w_i *— его тип и* l_i *— функция апелляции. Будем говорить, что агент правдивый, если* w_i *равно его истинной скрытой ценности* v_i.

Тогда очевидно, что для механизма с апелляциями для каждого вектора ставок w, если агенты правдивы ($w = v$) и избран исход $\hat{o}$, то
$$V(v, \hat{o}) \geq V(w, g(v)).$$

Для правдивых агентов результат всегда не хуже, чем $g(v)$. Это следует из определения механизма. Кроме того, верен факт, аналогичный соответствующему свойству механизмов VCG:

Лемма. *Прибыль агента* i *равна*
$$V((v_i, w_{-i}), \hat{o}) + h_i(w_{-i}, l_{-i}).$$

То есть агенту выгодно говорить $w_i \neq l_i$, только если это либо даст распределение лучше чем $g(v)$, либо поможет апелляции какого-нибудь другого агента.

А если агент будет врать, то это может ему повредить в двух случаях — или алгоритм g станет хуже работать, или же начнет измерять всеобщее счастье на основании неправильных типов, и поэтому выберет не ту альтернативу.

Мы хотели бы доказать, что агенту выгодно сказать правду о своем типе, а все свои сомнения в оптимальности g изложить в рамках апелляции.

1.4. Пример

Рассмотрим комбинаторный аукцион на двух вещах. Предположим, что агент оценивает пару из двух вещей в 3 миллиона, но каждую по отдельности в 1 миллион.

Тип агента состоит из трёх чисел — ценности вещи 1, ценности вещи 2 и ценности их обеих.

Пусть для агента i вещи комплементарны: за обе он готов отдать \$3, а за каждую — только по \$1. Тогда его тип будет $v_i = \{3, 1, 1\}$.

Давайте предположим, что агент заметил, что алгоритм распределения частенько работает лучше, если ему не давать вариантов, а говорить жестко $w_i = \{3, 0, 0\}$ вместо $\{3, 1, 1\}$, т.е. скрыть что он готов купить одну вещь.

В VCG-based механизме агент предпочтет говорить w_i. Могут возникнуть две проблемы.

1. Если даже остальные правдивы, может быть много разных векторов v_{-i}, для которых w_i вместо v_i даёт результаты не лучше, а хуже.

2. Но даже если каждый агент выберет w_i так, чтобы $V((v_i, w_i), g(v_i, w_{-i})) \geq V((v_i, w_i), g(w))$, может так случиться, что общий результат в итоге всех врущих агентов станет хуже: $V(v, g(w)) < V(v, g(w))$.

Механизм второго шанса дает агенту возможность проверить, приведет ли объявление неправильного типа к лучшему результату, т.е. станет ли лучше от вранья *в данном конкретном случае.* Он подаст $w_i = \{3, 1, 1\}$ и апелляцию, которая заменяет w_i на $w_i' = \{3, 0, 0\}$. Механизм подставит $\{3, 0, 0\}$ вместо w_i *с данными конкретными типами других агентов.* То есть ситуация для i беспроигрышная. Заметим, что механизм позволяет апелляции агента повлиять не только на свой тип, но и на остальные типы. Это дает сильный аргумент в пользу правдивости.

2. Доказательство правдивости

2.1. О правдивости в механизмах с апелляциями

В данном разделе докажем рациональность правдивости в механизмах второго шанса.

Нам придётся немножко по-другому определять правдивость, потому что тут более сложная конструкция. Придётся изменить и понятие доминантной стратегии. Неформально, стратегия будет доминантной, если агент не знает о стратегии, которая лучше.

Введем следующие понятия:

Опр. *Обозначим через A_i множество действий агента i, которые теперь не совпадают с типами. Функция ревизии агента i — это частичная функция $b_i : A_{-i} \to A_{-i}$. Семантика $b_i(a_{-i})$ фактически означает: «если бы я знал, что другие сделают a_{-i}, я бы выбрал $b_i(a_{-i})$ вместо a_i».*

Пусть i — агент, b_i — его функция ревизии, a_{-i} — вектор действий других. Тогда действие a_i удовлетворяет условию целесообразности отсутствия сомнений, если либо a_{-i} не лежит в области определения $dom\ b_i$, либо $u_i((b_i(a_{-i}), a_{-i})) \leq u_i(a)$. Т.о. агент не знает ничего лучшего, чем сделать a_i.

Теперь введем понятие *целесообразно доминантных действий*, которые учитывают тот факт, что возможности агентов ограничены. Затем продемонстрируем, что при разумных предположениях об агентах, правдивые, получаемые за полиномиальное время, целесообразно доминантные действия существуют.

Опр. *Действие a_i называется целесообразно доминантным, если для каждого вектора a_{-i} действий других агентов, a_i удовлетворяет условию целесообразности отсутствия сомнений относительно a_i и b_i.*

Обычно делается предположение, что оптимальная функция вычислима, но в большинстве ситуаций пространство действий слишком большое и эту функцию слишком сложно вычислить, даже приближенно за разумное время. В таких ситуациях предположение сделанное выше более не является корректным.

То есть агент ни для каких a^{-i} не знает лучше, чем сделать a_i.

Опр. *Действие называется целесообразно правдивым, если оно правдиво и целесообразно доминантно.*

2.2. Функции ревизий и целесообразно правдивые действия

Мы показали, что агенты правдивы, когда счастье не меньше $V(v, g(v))$. Также мы обсудили, что целесообразно правдивое действие возможно, когда у агента есть стимул выбрать его. В данном разделе будет показано, что в случае разумных предположений об агентах, вычислимые за полиномиальное время целесообразно правдивые действие существуют.

Опр. *Обозначим $\perp$ пустую апелляцию, а $(w, \perp)$ — вектор действия, при котором все агенты объявляют w^i и все апелляции пустые.*

Опр. *Функция ревизии b называется апелляционно независимой, если у каждого вектора в её области определения есть только пустые апелляции, т.е. $\forall a^{-i} \in dom\ b$ существует такой w^{-i}, что $a^{-i} = (w^{-i}, \perp)$.*

Будем говорить, что алгоритм T-ограничен, если он работает за T.

Апелляционно независимая функция T-ограничена, если она сама T-ограничена и каждая апелляция в её образе тоже T-ограничена.

Теор. *Если у агента в механизме с апелляциями есть апелляционно независимая T-ограниченная функция ревизии, и алгоритм распределения механизма тоже T-ограничен, то для каждого T'-ограниченного механизма, $T' = \Omega(T)$, у этого агента есть целесообразно правдивое действие.*

Доказательство:

Пусть b_i — функция ревизии. Нужно определить апелляцию. Для каждого w_{-i} есть два исхода — $o_1 = g(w)$ и $o_2 = g(\tau_i(w))$, где $(w_i, \tau_i) = b_i(w_{-i}, \perp)$.

Мы определим $l_i(w)$ как лучший из этих исходов: $l_i(w) = argmax_{j=1,2} V((v_i, w_{-i}), o_j)$, то есть l_i проверяет, хорошо ли будет агенту сказать w_i. Тогда $a_i = (v_i, l_i)$ будет целесообразно правдивым.

Докажем это утверждение. Если a_i не правдиво, то существует вектор $a_{-i} = (w_{-i}, \perp) \in dom\ b_i$, для которого $u(a_i, a_{-i}) < u(b_i(a_{-i}), a_{-i})$.

Обозначим $b_i(a_{-i}) = (w_i, \tau_i)$. Как говорилось ранее, доход агента равен всеобщему счастью $V((v_i, w_{-i}), o)$ (с точностью до h_i, которые не зависят от действий агентов).

Теперь рассмотрим случай, когда агент выбрал $b_i(a_{-i})$. Обозначим полученный исход, как $\hat{o}$. В соответствии с определением механизма $\hat{o}$, который был выбран из множества $\{o_1, o_2\}$, при условии того что счастье измерялось относительно заявки агента w_i.

А если агент выбрал правдивое действие a_i и получил исход $\tilde{o}$, то в этом случае этот исход был выбран механизмом из $o_0 = g(v_i, w_{-i})$, а также из o_1 и o_2 (по определению l_i). Это множество, включает в себя данное множество и множество исходов, полученных в первом случае, т.е. множество, из которого алгоритм выбирал полным перебором, строго увеличилось. Более того исход был выбран в соответствии с правильным типом v_i.

Это значит, что $V((v_i, w_{-i}), \tilde{o}) \geq V((v_i, w_{-i}), \hat{o})$. А это в свою очередь является противоречием.

Заметим, что l_i действительно $\Omega(T)$-ограничена. Это ясно, так как g и τ_i являются T-ограниченными. Это завершает доказательство теоремы. т

Для данной функции ревизий агента легко сконструировать апелляцию l_i, описанную выше (то есть сконструировать программу, которая вычислит ее). Следовательно, если у агента есть такая апелляционно независимая функция, то он может гарантировать себе целесообразно доминантные действия.

2.3. d-ограниченные функции ревизии

Класс d-ограниченных функций ревизии агентов, которые помимо алгоритма представляют полиномиальное семейство потенциальных апелляций для других агентов. Этот класс является обобщением d-ограниченных апелляционно-независимых функций.

Опр. *Функция ревизии b_i d-ограниченная:*
Если она $O(n^d)$-ограничена и для множества всех апелляций в области определения b_i:
$L = \{l_i \mid \exists l_{-ij}, w_{-i} : \ (w_{-i}, (l_i, l_{-i})) \in dom\ b_i\} \bigcup \{l_i \mid \exists (w_{-i}, l_{-i}), w_i : \ (w_i, l_i) = b_i(w_{-i}, l_{-i})\}$.
Тогда $|L| = O(n^d)$ и для некоторой константы с каждая апелляция из L cn^d-ограничена.

2.4. О хороших механизмах с апелляциями

Теор. *Если у агента в механизме с апелляциями есть d-bounded функция ревизии, и алгоритм распределения механизма тоже $O(n^d)$-ограничен, то для каждого T'-ограниченного механизма, $T' = \Omega(n^{2d})$, у этого агента есть целесообразно правдивое действие.*

Доказательство:
Пусть i — агент, а b_i функции ревизии. Для каждого вектора w_{-i} вычислим следующие исходы:

1. $o_0 = g(w)$.
2. Из теоремы, доказанной ранее, для всех $\tau_j \in L$ определим $o_j = g(\tau_j(w))$.
3. Определим $l(w) = argmax_{0 \le j \le |L|} V((v_i, w_{-i}), o_j)$, как результат с максимальной прибылью, в соответствии с (v_i, w_{-i}), среди всех результатов, описанных выше.

Покажем, что l_i будет n^{2d}-ограничена.
В соответствии с определением, апелляция l_i совершает $n^d + 1$ операции, каждая из которых требует не более чем cn^d единиц времени. Следовательно, суммарное вычисление занимает не более чем $O(n^{2d})$.
Теперь покажем, что $a_i = (v_i, l_i)$ будет целесообразно правдивым.
Докажем это утверждение от противного:
Пусть существует вектор a_i в образе b_i, такой что $u(a_i, a_{-i}) < u(b_i(a_{-i}), a_{-i})$.
Рассмотрим случай, когда агент выбирает $b_i(a_{-i}) = (w_i, \tau_i)$. Механизм получает результат $\hat{o}$, который максимизирует всеобщее счастье (в соответствии с w) из следующего множества исходов S:

1. $o_0 = g(w)$.
2. $o_j = g(l_j(w))$ для каждого $i \neq j$, то есть результат апелляций других агентов.
3. $o_i = g(\tau_i(w))$

Когда агент выбирает a_i, результаты измеряются в соответствии с вектором «правильных» типов (v_i, w_{-i}). Более того того они выбраны из множества, которое включает в себя данное множество и множество S:

1. $o'_0 = g((v_i, w_{-i}))$ (из определения механизма).
2. $o'_j = g(l_j((v_i, w_{-i})))$ для каждого $i \neq j$, то есть результат апелляций других агентов. (также из определения механизма)
3. $o'_j = g(\tau(w))$ для любого $\tau \in L$. Так как a_{-i} находится в образе b_i, то множество включает все результаты формы $g(l_j(w))$ в том случае, когда i выбирает $b_i(a_{-i})$. Также оно содержит результат своей собственной апелляции $\tau_i(w)$.

4. $g(w)$ из определения l_i.

Пусть $\tilde{o}$ является выбранным результатом в этом случае, так как множество результатов во втором случае является множеством, включающем в себя данное множество и множество исходов первого случая, $g((v_i, w_{-i}), \tilde{o}) \geq g((v_i, w_{-i}), \hat{o})$. В соответствии упомянутой ранее леммой, прибыль агента при выборе a_i получается больше, чем при выборе $b_i(a_{-i})$. Таким образом мы получили противоречие.

Следовательно, теорема доказана.

Как и в случае с апелляционно-независимыми функциями, теорема дает предписание для создания апелляций, котрые гарантируют агенту целесообразно доминантное действие.