

ДАНИЕЛЬ ДЕННЕТТ

## Условия присутствия личности

**Я** личность (person), и вы тоже личность. В этом пункте сомнений нет. Я человек, и вы, возможно, тоже человек. Если вы посчитали себя оскорбленным в отношении слова «возможно», вы виновны в своего рода расизме, ибо важным относительно нас с вами является не наша принадлежность к одному биологическому виду, а то, что мы оба являемся личностями, а вот это сомнению не подвергалось. Достоинство не зависит от происхождения — тут неважно, были ли вы рождены от женщины, были ли вообще рождены. Обычно мы не обращаем внимания на подобные тонкости, считая принадлежность к человеческому роду критерием обладания личностью — потому, без сомнения, что эти понятия имеют одинаковый или почти одинаковый локальный объем. В настоящее время человеческие существа являются единственными признаваемыми нами личностями; кроме того, мы признаем личностями почти все известные нам человеческие существа. И все же, с одной стороны легко вообразить личности, биологически весьма отличные от нас — они, может статься, населяют другие планеты — с другой, мы признаем условия, при которых человеку отказывают либо целиком в личности, либо в некоторых важных ее составляющих. К примеру, умственно неполноценные человеческие существа, дети и существа, признанные невменяемыми лицензированным психиатром, личностями не считаются, или же, по крайней мере, считаются лишенными ее существенных элементов.

Можно было бы ожидать, что у настолько важного понятия — мы признаем и отрицаем его с такой уверенностью — окажутся ясные необходимые и достаточные условия применения, однако, если это и так, мы их до сих пор не обнаружили. Не исключено, что мы в конце концов придем к выводу, что понятие личности непоследовательно и отслужило свой срок. Скиннер, среди прочих, утверждал именно это, но его доктрина не прижилась — частично оттого, без сомнения, что нам трудно и даже невозможно вообразить, к чему привел бы отказ от понятия личности<sup>1</sup>. Идея, что мы можем пере-

<sup>1</sup> Скиннер (Burthus Fredric Skinner, 1904–1990) является основателем теории бихевиоризма.

Эта радикально-эмпирическая теория, претендуя на научность в понимании человеческого поведения, отказывается от ментальных предикатов на основании их ненаблюдаемости, объявляя все понятия, связанные с состояниями сознания, нереальными, мифологическими фикциями. Поведение как животных так и людей, настаивает бихевиоризм, следует видеть и объяснять как набор рефлексов, как определенную зависимость между внешними стимулами и поведенческими реакциями. Деннетт является противником бихевиоризма

стать рассматривать других и *себя* в качестве личностей (если тут не подразумевается наше полное самоуничтожение и, следовательно, прекращения рассмотрения чего бы то ни было в качестве чего бы то ни было) — такая идея, согласно существующим аргументам, внутренне противоречива (См. гл.12) Поэтому, даже если мы оставим в стороне сильные и слабые стороны рассуждений, послуживших основанием доктрины Скиннера, трудно представить себе как они смогли бы одержать победу в борьбе с таким интуитивно неуязвимым понятием как личность. И все же, если это понятие и является некой неустранимой частью нашего концептуального аппарата, оно, похоже, находится не в столь хорошей форме, как можно было бы ожидать. Может статься, к примеру, что личность есть нечто вроде висящего в воздухе почетного звания, которое под настроение все мы с удовольствием присваиваем себе и окружающим, руководствуясь при этом эмоциями, эстетическими пристрастиями, политическими соображениями и т.п. — подобно тому как светским львом является тот и только тот, кому удастся заставить считать себя светским львом тех, кто считает себя светскими львами. Личность есть безусловно *нечто сходное*, и если она только это и не более того, нам следует пересмотреть значимость, которой мы наделяем сейчас это понятие.

Даже предположив, что личность *действительно* есть нечто большее, охотник за необходимыми и достаточными условиями все равно может наткнуться на препятствия — в случае, если имеется несколько различных понятий личности. Есть основания полагать, что дела именно так и обстоят. Грубо говоря, мы, похоже, имеем дело с двумя переплетенными понятиями, одно из которых можно назвать моральным, другое — метафизическим. Локк утверждает, что «личность» (person)

есть юридический термин, связанный с действиями и их сущностью; личностью, поэтому, обладают только разумные агенты, способные руководствоваться законом и испытывать счастье и несчастье. Личность, посредством сознания, продолжается за пределы сиюминутного существования по направлению к прошлым событиям — и таким образом получает возможность призывать к ответу и нести ответственность. (Опыт, Книга II, Глава XXVII).

*Совпадает ли* метафизическое понятие личности — понятие, в самом общем смысле, разумного, обладающего сознанием и чувствами агента — с моральным понятием агента ответственного, обладающего правами и обязанностями? Или же личность в метафизическом смысле есть лишь необходимое, но не достаточное условие личности в моральном смысле? *Одно ли и то же* быть существом, которому приписывается сознание или самосознание, и быть целью-в-себе — или же первое есть лишь предварительное условие второго? В теории справедливости Роулза, должен ли вывод из исходного положения рассматриваться в качестве демонстрации того, как метафизические личности *могут стать* моральными личностями, или же как показ того, почему метафизические личности *должны являться* моральными личнос-

и в книге *Brainstorms*, из которой взята настоящая статья, в главе «Skinner Skinned» («Скиннер скидывает кожу») приводит детальную критику идей и методов Скиннера. — *Прим. перев.*

<sup>2</sup> В статье «Justice as Reciprocity (Справедливость как взаимность)», новой редакции предыдущей статьи «Justice as Fairness (Справедливость как честная игра)», напечатанной в книге S. Gorovitz, ed., *Utilitarianism* (Indianapolis: Bobbs Merrill, 1971), Роулз допускает, что личности

тями?<sup>2</sup> В менее абстрактных контекстах различие выступает также ясно: когда мы объявляем человека душевнобольным, мы перестаем считать его ответственным за свои действия и отказываем ему в большинстве прав; при этом, если он не слишком уж далеко зашел по пути безумия, наше отношение к нему практически неотлично от нормального личностного отношения. В некотором смысле слова «личность», мы, похоже, продолжаем рассматривать его как личность и вести себя с ним соответственно. В самом начале я заявил, что мы — вы и я — являемся личностями. Я не мог рационально рассчитывать — и уж точно не имел достаточных оснований утверждать — что все читатели этой статьи легально вменяемы и способны нести моральную ответственность. Если что-то и было вне сомнений, так это то, что любой, к кому адекватно применимы использованные в первом предложении личные местоимения «я» и «вы», является личностью в метафизическом смысле. Если только это не подвергается сомнению, то метафизическое и моральное понятия личности различны. Однако даже если мы согласимся, что это так, имеются все основания полагать, что присутствие личности в метафизическом смысле есть необходимое условие присутствия личности в смысле моральном<sup>3</sup>.

Здесь я хотел бы остановиться и рассмотреть шесть известных тем — каждая из них является попыткой определить необходимое условие присутствия личности; каждая, в рамках определенной интерпретации, кажется мне верной. Вопрос будет стоять так: во-первых, как (в рамках моей интерпретации) они логически зависят одна от другой; во-вторых, почему они являются необходимыми условиями присутствия личности в моральном смысле; и, в-третьих, почему так трудно определить, являются ли они совместно достаточными условиями присутствия моральной личности.

*Первая*, самая очевидная, тема связана с утверждением, что личности суть *рациональные существа*. Она присутствует, например, в этических теориях Канта и Роулза, а также в «метафизических» теориях Аристотеля и Хинтикки.<sup>4</sup>

*Вторая* тема утверждает, что личности есть существа, которым либо атрибутируются состояния сознания, либо приписываются психологические, ментальные или *интенциональные предикаты*. Строусон, в частности, определяет концепт личности как «концепт такого типа сущности, что к ней при-

(persons) в исходном положении могут представлять собой «нации, провинции, деловые фирмы, церкви, команды, и так далее. Принципы справедливости применимы к конфликтным притязаниям, исходящим от лиц всех этих различных типов. Случай человеческих индивидуумов, по-видимому, обладает неким логическим приоритетом.» (Стр.245) В *Теории справедливости* (A Theory of Justice, Cambridge, Mass.: Harvard University Press, 1971) Роулз признает, что участники в исходном положении могут быть ассоциациями и другими образованиями, а не человеческими индивидуумами. (e.g. p. 146), и очевидная взаимозаменяемость фраз «участники в исходном положении» и «личности в исходном положении» приводит к идее, что Роулз считает, что для определенного морального понятия личности, моральная личность может быть *составлена* из метафизических личностей, которые сами по себе могут быть, а могут и не быть моральными личностями.

<sup>3</sup> Оставляя в стороне возможные составные моральные личности (лица) Роулза. О составных лица см. Amelie Rorty, «Persons, Policies, and Bodies», *International Philosophical Quarterly*, XIII, 1 (March, 1973)

<sup>4</sup> J. Hintikka, *Knowledge and Belief* (Ithaca: Cornell University Press, 1962).

менимы как предикаты, приписывающие состояния сознания, так и предикаты, приписывающие телесные, физические свойства.»<sup>5</sup>

Третья тема связана с тем, что вопрос о присутствии у чего-либо личности зависит в определенном смысле от занимаемой по отношению к этому чему-либо установки (*stance*). Из этой темы вытекает, что дела обстоят не так, что, определив сначала объективный факт наличия у чего-либо личности, мы затем относимся к этому существу соответственно, но наоборот, само это наше специфическое отношение к чему-либо или кому-либо конституирует его (будь то он, она или оно) как личность — определенным, естественно, образом и до известного предела. Вариации на эту тему высказывали Маккей, Струосон, Рорти, Путнам, Селларс, Флу, Нагель, Ван де Вэйт, и я сам.<sup>6</sup>

Четвертая тема говорит о том, что объект, по отношению к которому занимается эта установка, должен быть в состоянии каким-то образом *ответить взаимностью*. Крайне разнообразные версии такой идеи высказывались или намечались Роулзом, Маккеем, Струосоном, Грайсом, и другими. Эта взаимность иногда выражается следующим, весьма, впрочем, малоинформативным, лозунгом: быть личностью значит относиться к другим как к личностям — и такую формулировку обычно сопровождает утверждение, что относиться к другому как к личности означает относиться к нему морально — возможно, следуя Золотому Правилу — однако это утверждение смешивает различные типы взаимности. По Нагелю «крайне враждебное отношение к другому человеку совместимо с отношением к нему как к личности» (стр. 134); Ван де Вэйт замечает, что разница между некоторыми формами преднамеренного и непреднамеренного убийства заключается в том, что в первом случае убийца относится к жертве как к личности.

Пятая тема связана с тем, что личности должны быть способны осуществлять *вербальную коммуникацию*. Это условие эффективно отрицает у животных полноту личности и связанную с ней моральную ответственность; явно или неявно оно присутствует во всех этических теориях типа социального контракта. Это условие также подчеркивается или предполагается

<sup>5</sup> См. P. F. Strawson, *Individuals* (London: Methuen, 1959), стр. 101–102. Часто отмечалось, что определение Струосона откровенно широко, ибо оно покрывает все активные существа обладающие сознанием. См. например Н. Frankfurt, «Freedom of the Will and the Concept of a Person», *Journal of Philosophy*, LXVIII (January 14, 1971). Можно также утверждать (и я с этим согласен), что состояния сознания являются лишь подклассом психологических или интенциональных состояний, но мне кажется ясным, что Струосон в данном случае как раз и хотел забросить свои сети так широко, чтобы в них попали и психологические состояния в самом общем смысле.

<sup>6</sup> D. M. MacKay, «The use of behavioral language to refer to mechanical processes», *British Journal of Philosophy of Science* (1962): 89–103; P. F. Strawson, «Freedom and Resentment», *Proceedings of the British Academy* (1962), перепечатано в Strawson, ed., *Studies in the Philosophy of Thought and Action* (Oxford, 1968); A. Rorty, «Slaves and machines», *Analysis* (1962); H. Putnam, «Robots: Machines or Artificially Created Life?», *Journal of Philosophy*, LXI (November 12, 1964); W. Sellars, «Fatalism and Determinism», в книге K. Lehrer, ed., *Freedom and Determinism* (New York: Random House, 1966); A. Flew, «A Rational Animal», в книге J. R. Smythies, ed., *Brain and Mind* (London: Routledge & Kegan Paul, 1968); T. Nagel, «War and Massacre», *Philosophy and Public Affairs* (Winter 1972); D. Van de Vate, «The Problem of Robot Consciousness», *Philosophy and Phenomenological Research* (December 1971; Глава 1 настоящего издания).

многими авторами (включая меня самого), работающими в тех разделах философии сознания, где моральное измерение личности не обсуждается.

*Шестая тема* заключается в том, что личности отличаются от других существ наличием той или иной формы сознания: мы обладаем сознанием такого типа, какого нет ни у каких других биологических видов. Иногда этот тип сознания определяется как определенный вид *самосознания*. Трое философов с разных сторон утверждают, что наличие сознания есть предварительное условие возникновения морального агента — Анскомб, в книге *Интенция*, Сартр, в работе *Трансценденция Эго*, и Франкфурт, в недавней статье «Свобода воли и концепция личности».<sup>7</sup>

Ниже я попытаюсь продемонстрировать, что порядок, в котором я привел эти шесть тем — с одной оговоркой — является порядком их логического следования. Оговорка следующая: первые три темы взаимосвязаны; агент рационален, когда он интенционален, когда он является объектом определенной установки (*stance*). Все вместе эти три темы являются необходимым, но не достаточным условием наличия той формы взаимности, которая, в свою очередь, является необходимым, но не достаточным условием способности к языковой коммуникации, которая является необходимым<sup>8</sup> условием обладания особой формой сознания, которая, согласно тому, что разными способами доказывают Анскомб и Франкфурт, есть необходимое условие присутствия моральной личности. (Идеи Сартра здесь обсуждаться не будут.)

В прошлом я уже обращался к первым трем темам — рациональности, интенциональности, и установки — с целью определить не личность как таковую, но более широкий класс, который я назвал классом *интенциональных систем*. Ниже я собираюсь строить свои рассуждения на основе этого понятия, поэтому стоит привести его краткий обзор. Интенциональной является система, поведение которой может (по крайней мере иногда) объясняться и предсказываться на основе приписывания ей *представлений* (*beliefs*) и *желаний* (*desires*) — а также других интенциональных состояний, которые далее я буду называть просто *интенциями*, имея в виду надежды, страхи, намерения, восприятия, ожидания, и т.д. В случае *любой* системы могут найтись и другие способы объяснения и предсказания ее поведения — например, механический или физический — однако интенциональная установка может оказаться либо самой удобной или эффективной, либо просто приводящей к успеху — и этого достаточно, чтобы определить ту или иную систему как интенциональную.

При таком определении очевидно, что не все интенциональные системы являются личностями. Мы приписываем представления и желания рыбам и собакам и таким образом предсказываем их поведение; мы даже можем употребить этот же метод для предсказания поведения некоторых машин. Подобная стратегия удачна в отношении, к примеру, шахматного компьютера (это, кстати, единственная удачная стратегия в данном случае). *Предпола-*

<sup>7</sup> Н. Frankfurt, «Freedom of the Will and the Concept of a Person», *Journal of Philosophy*, LXVIII (January 14, 1971).

<sup>8</sup> И достаточным, но я здесь утверждать этого не буду. Я пытаюсь доказать это утверждение в книге *Content and Consciousness (Содержание и Сознание)*, а также, в более свежем и явном виде, в главах 2 и 9 настоящего издания.

гая, что в связи с разворачивающейся партией у компьютера есть определенные представления (информация) и желания (функции предпочтения), я могу рассчитать — в подходящих обстоятельствах — его наиболее вероятный следующий ход, *полагая, что компьютер рационален по отношению к этим представлениям и желаниям*. Компьютер в этом примере является интенциональной системой не потому, что он обладает какими-то специфическими внутренними свойствами, а также не потому, что у него *на самом деле* есть представления и желания (как бы мы их не определяли), но просто потому, что по отношению к нему оказывается успешной некая *установка*, а именно интенциональная установка, согласно обычным правилам приписывающая ему интенциональные предикаты — установка, рассматривающая компьютер в качестве рационального представителя практического разума.

Важно осознать, насколько мягким является такое определение, и, соответственно, насколько широким оказывается класс интенциональных систем. Если, к примеру, мне удастся предсказать, что некое растение — скажем, плющ в горшке — обогнет угол и будет расти далее по направлению к свету, потому что оно «ищет» света и «хочет» выбраться из тени, а также, потому что оно «ожидает» или «надеется», что свет есть за углом, то я занимаю по отношению к этому растению определенную установку, и — приятный сюрприз — в достаточно узких пределах она работает. И в той мере, в какой эта установка эффективна, некоторые растения являются интенциональными системами очень низкого уровня.

Реальная польза принятия интенциональной установки по отношению к растениям стала мне окончательно ясна в разговорах с лесорубами штата Мэйн. Лесорубы неизменно называют дерево не «оно», а «он», и говорят о молодой ели так: «Он хочет раскинуть лапы в стороны, но делать это ему давать нельзя; тогда, чтобы выбраться к свету, ему придется подрости, « или же так: «Сосны не любят мочить ноги, в этом они не похожи на кедры». «Можно «обмануть» яблоню, и «заставить ее поверить, что настало лето», поздней осенью разжигая под ней небольшой костер; в результате она зацветает. Такой язык не сводится целиком к живописности, и отнюдь не является суеверием; это просто эффективный способ понимания, контроля, прогнозирования и объяснения поведения растений, способ, который элегантно справляется с недостатком знания их управляющих механизмов. Изошренный биолог может говорить о передаче информации от периферии дерева к другим его частям, и это опять же будет пусть менее живописный, но еще интенциональный язык. Полное воздержание от интенциональных понятий может превратиться в настолько же героическое, неуклюжее и бессмысленное предприятие, каким стало аналогичное бихевиористское табу в отношении крыс и голубей. И даже когда эвристическая ценность интенционального подхода к поведению, скажем, дерева, сокращается до сходящего к нулю минимума, мне все же кажется более разумным принять, что дерево является вырожденной, неинтересной, исчезающе-интенциональной системой, нежели проводить черту, выше которой интенциональные интерпретации «объективно, на самом деле соответствуют истине».

Из вышеизложенного очевидно, что интенциональность системы не является достаточным условием присутствия личности, но является необходи-

мым условием. Если мы не можем по отношению к чему-либо принять интенциональную установку (частью которой является предположение о рациональности), то это что-то не может считаться личностью.

Нельзя ли, как следствие, определить личность как особый подкласс интенциональных систем? С первого взгляда кажется плодотворной гипотеза, что личностями является подкласс таких интенциональных систем, которые действительно обладают представлениями, желаниями и так далее, а не просто предполагаются имеющими таковые в целях резкого увеличения предсказательной силы. Похоже, однако, что попытки определить действительное представление (*belief*), исключающие возможность появления оно-го у собаки, дерева или компьютера, заканчиваются тем, что мы накладываем на это реальное представление такие условия, что (1) их жесткость противоречит нашим интуициям, и (2) они оказываются связаны с критериями личности, которые находятся дальше в нашем списке. Например, возможно утверждать, что действительные представления (*beliefs*) обязательно должны быть *вербально выразимы* их носителем,<sup>9</sup> или же, что носитель должен *осознавать*, что он их имеет — однако у нас, судя по всему, есть множество представлений, которых мы не можем выразить словами, а также множество неосознаваемых представлений — кроме того, я надеюсь продемонстрировать, что способности с одной стороны к вербальному выражению, а с другой к осознанию, занимают свое особое место в наборе необходимых условий присутствия личности.

Гораздо дальше, мне кажется, продвигает нас обращение к четвертой теме, взаимности, и попытка рассмотреть как взаимность может быть определена в терминах интенциональных систем. Эта тема указывает, что личность должна отвечать взаимностью на определенную по отношению к ней установку, откуда понятно, что под такой критерий подпадает интенциональная система, сама занимающая интенциональную позицию по отношению к другим объектам. Определим как *интенциональную систему второго порядка* такую систему, которой мы приписываем не только простые представления, желания и другие интенции, но и представления, желания и другие интенции *по поводу* представлений, желаний и других интенций. Интенциональная система *S* будет системой второго порядка, если мы приписываем ей, например, следующее: *S* думает, что *T* желает, чтобы *p*; *S* надеется, что *T* боится, что *q*, а также возвратные случаи — *S* думает, что *S* хочет, чтобы *p*. (Важность возвратных случаев значительно вырастет (что неудивительно), когда мы обратимся к интерпретации нашего шестого условия, *само-сознания*. Некоторые считают, что возвратность автоматически превращает все интенциональные системы в системы второго, и даже *n*-ного порядка, на том основании, что, когда я думаю, что *p*, это предполагает, что я думаю, что я думаю, что *p*, и так далее — однако это фундаментальная ошибка; итерация представлений и других интенций никогда не бывает формально-избыточной; поэтому, несмотря на естественность и ожидаемость некоторых итераций, в общем случае они никогда не происходят тривиально и автоматически.)

<sup>9</sup> Ср. Bernard Williams, «Deciding to Believe», в книге H. E. Kiefer and M. K. Munitz, eds., *Language, Belief and Metaphysics* (New York: New York University Press, 1970).

Зададимся теперь вопросом, являются ли человеческие существа единственными интенциональными системами второго порядка. Я рассматриваю этот вопрос как эмпирический. Мы приписываем представления и желания собакам, кошкам, львам, птицам и дельфинам, и таким образом, если все идет нормально, прогнозируем их поведение. Трудно при этом измыслить случай настолько изоциренного поведения животного, что нам пришлось бы, в целях предсказательных или объяснительных, приписать ему интенции второго порядка. Конечно, если истинна какая-либо из версий механического физикализма (во что я сам верю), нам никогда *не понадобится* окончательно приписывать интенции чему бы то ни было, однако если в эвристических и прагматических целях мы все же решаем признать интенции у животных — возникнут ли у нас в этом случае прагматические мотивы предположить наличие у них также и интенций второго порядка? Психологи часто взывают к принципу, известному как Канон Экономности Ллойда Моргана — его можно рассматривать как один из вариантов Бритвы Оккама, ибо он утверждает, что, пытаясь адекватно описать поведение, организму следует приписывать настолько мало разумности и сознания, насколько это возможно. Этот принцип может интерпретироваться (зачастую именно так и происходит), как требование, доходящее до радикального бихевиоризма,<sup>10</sup> но мне кажется, что это ошибка, и что его можно понимать как требование при принятии интенциональной установки ограничиваться предположениями о простейших, наименее высокоорганизованных представлениях, желаниях и т.д., достаточных для адекватного описания. Учитывая это, мы согласимся, к примеру, что Бобик *хочет* ужинать, и *думает*, что хозяин накормит его, видя как он просит, однако, нам при этом не требуется приписывать Бобику более развитое *представление* о том, что его умоляющий вид порождает у хозяина *идею*, что он, Бобик, *хочет* ужинать. Подобным же образом, когда я опускаю десятицентовик в автомат, продающий конфеты, мои *ожидания* не основаны на *представлении* более высокого порядка, а именно, что уходящая в автомат монета вызывает у него *представление*, что я *хочу* конфету. Иными словами, несмотря на значительное сходство бобикова умоляющего вида с коммуникацией второго порядка (в которой пес относится к хозяину как к интенциональной системе), если мы предположим, что Бобик считает хозяина просто выдающей ужин машиной, запускаемой умоляющим видом, мы придем к столь же адекватной предсказательной теории, более скромной в своих посылках, но при этом без сомнения все еще интенциональной.

Должны ли мы сделать отсюда вывод, что собаки, шимпанзе и другие «высшие» животные, неспособны подняться до уровня интенциональных систем второго порядка — и если да, то почему? Раньше я такой вывод делал, видя причину в том, что в отличие от человека животные не обладают языком, который, как мне казалось, необходим для интенций второго порядка. Иными словами, я думал, что четвертое условие присутствия личности может вытекать из пятого. Меня искушала следующая гипотеза: животные не могут иметь, скажем, представлений второго порядка — представлений

<sup>10</sup> Напр. в статье B. F. Skinner, «Behaviorism at Fifty», в книге T.W. Wann, ed., *Behaviorism and Phenomenology* (Chicago: University of Chicago Press, 1964).



о представлениях — по той же причине, по какой они не могут иметь представлений о пятнице или о поэзии. Некоторые представления могут приобретаться, и, следовательно, ментально репрезентироваться только при помощи языка.<sup>11</sup> Но даже если некоторые представления действительно нельзя получить в обход языка, неверно, что все представления второго порядка принадлежат к их числу, и неверно также, что ничто, кроме человеческих существ, не может быть интенциональной системой второго порядка. Питер Ашли в письме описывает такой случай:

Однажды вечером я сидел у себя дома в кресле — в *единственном* кресле, где разрешалось спать моей собаке. Собака лежала передо мной на полу и скулила. Ей никак не удавалось «убедить» меня освободить кресло. Ее следующий ход является самым интересным, более того, *единственным* интересным местом всей истории. Собака встала и пошла к двери, причем так, чтобы я мог ее ясно видеть. Там она стала скрестись, наводя меня на мысль, что, отказавшись от попыток заполучить кресло, она решила вместо этого выйти во двор. Но как только я подошел к двери, чтобы ее выпустить, она стремглав бросилась назад через комнату и залезла в кресло, которое «вынудила» меня оставить.

Здесь вроде бы нам следует приписать собаке *намерение* заставить хозяина *поверить*, что она *хочет* во двор — намерение не просто второго, но третьего порядка. Ключ к этому примеру — элемент, делающий его примером интенции более высокого уровня — связан с тем, что представление, которое собака пытается вызвать у хозяина, является ложным. Если мы хотим обнаружить другие примеры животных, ведущих себя, как интенциональные системы второго порядка, стоит рассмотреть случаи обмана, когда животное, считая, что *р*, пытается вынудить другую интенциональную систему посчитать, что *не-р*. Там, где животное пытается вызвать в другом поведении, отличное от того, что вызвали бы *истинные* представления, в объяснении нельзя «отделить мясо от костей» и описать все в терминах интенций только первого порядка. Даже не приводя к этому причин, можно обобщить: когда *X* пытается вызвать в *У* поведение, неадекватное реальному окружению или нуждам *У*, но адекватное тому, что *У* *принимает за свою реальность или свои нужды*, нам приходится приписывать *X* интенции второго порядка. В такой форме это утверждение нам знакомо, ибо его часто использовали критики бихевиоризма: объясняя и контролируя поведение подопытных животных, можно быть бихевиористом только в отсутствии значительного рассогласования между реальной экспериментальной ситуацией и тем, что животные *воспринимают* как реальность. Известен тактический маневр против лабораторного бихевиориста — он заключается в организации эксперимента, обманывающего подопытных: если обман удастся, их поведение оказывается предсказуемо, исходя из их ложных представлений о ситуации, а не из самой ситуации. Отсюда видно, что интенциональная система первого порядка ведет себя как бихевиорист; она ничему не приписывает интенций. Поэтому, если мы хотим получить доводы в пользу того, что некая система *S* *не* является бихевио-

<sup>11</sup> По поводу идей, проливающих свет на отношения языка к представлениям и рациональности, см. Ronal de Sousa, «How to Give a Piece of Your Mind; or the Logic of Belief and Assent», *Review of Metaphysics* (September 1971).

ристом (что означает, что она является системой второго порядка) — это возможно только в тех ситуациях, где бихевиористские теории неадекватны, то есть, только тогда, когда бихевиоризм не объясняет каким образом системе S удается успешно манипулировать поведением другой системы.

Из этого вытекает, что пример Эшли не так уж и убедителен, и что его можно опровергнуть, предположив, что собака является в некотором роде бихевиористом. Ей не нужно считать, что поскребывание в дверь вызовет у Эшли идею, что она хочет во двор; собака, как хороший бихевиорист, может просто считать, что привила Эшли условный рефлекс идти к двери всякий раз, когда она в нее скребется. Она в этом случае прилагает обычный стимул, получает обычную реакцию, и это все. Пример Эшли не работает, если в описанном случае собака использует *стандартный способ* добиться открывания двери (скорее всего, так и есть), ибо тогда стоит удовлетвориться более скромной гипотезой, что собака считает, что хозяин натаскан подходить к двери, когда она в нее скребется. Если бы собака придумала *что-нибудь новенькое* с целью обмануть хозяина (скажем, подбежать к окну и подозрительно зарычать, выглядывая наружу), вот тогда мы бы согласились, что вставание с кресла было не просто рефлексом Эшли и не могло так «восприниматься» собакой — правда, подобная собачья изобретательность весьма маловероятна.

В чем же тогда разница между этим маловероятным вариантом и хорошо описанными случаями, когда строящая гнезда птица, дабы отвлечь хищника от птенцов, притворяется, что у нее сломано крыло? Достижимый ею эффект нов, ибо скорее всего она не могла привить всем хищникам вблизи своего гнезда нужные рефлексy, и поэтому нам вроде бы приходится объяснять эту ее уловку как пример настоящего обмана, где птица *намеревается* внушить хищнику ложное *представление*. Если бы нам самом деле пришлось принять такую интерпретацию, стоило бы восхититься подобной птичьей изворотливостью; однако хорошо известно, что действия птицы являются просто следствием инстинкта. Почему выражение «следствие инстинкта» звучит тут пренебрежительно? Утверждение, что мы имеем дело с инстинктом, означает, что все птицы этого вида к ней прибегают; они прибегают к ней даже в не вполне подходящих обстоятельствах; даже когда есть более веские резоны оставаться в гнезде; рисунок поведения зафиксирован раз и навсегда, является неким тропизмом<sup>12</sup>, управляющие воздействия которого встроены в организм генетически, а не приобретены в результате обучения или собственного изобретения.

Осторожность подсказывает не заходить слишком далеко в подобном пренебрежении. Ошибочно утверждать, что птица осуществляет свою уловку «бездумно» — она, конечно, ни в каком смысле не пробегает про себя ни по каким рассуждениям или схемам («Итак, если я буду махать крылом так, словно оно сломано, лиса подумает...»), но человек ведь тоже может делать нечто сходное по сложности, действительной разумности, новизне и соответствию ситуации, и при этом не иметь никаких «сознательных мыслей». *Думание мыс-*

<sup>12</sup> Тропизмом называется произвольная моторная реакция на раздражитель или внешние обстоятельства. — *Прим. перев.*

лей, как бы мы его ни определили, не является тем, что делает разумное поведение разумным. Анскомб говорит: «В самом общем смысле абсурдно считать, что [такая внешняя артикуляция рассуждения] описывает реальные ментальные процессы. Интересы теории требуют, чтобы эта артикуляция соответствовала некоему порядку, имеющему место всякий раз, когда действие совершается с намерением.»<sup>13</sup> Но «порядок имеет место» как в случае птицы, так и в случае человека. Иными словами, когда мы задаем вопрос, почему у птиц в результате эволюции развился этот тропизм, мы в ответ даем объяснения, отмечая полезность такого инструмента для *обмана* хищников, то есть, для вызывания у них ложных представлений; в объяснении нуждается только источник происхождения у птицы интенций второго порядка. Я буду последним, кто станет отрицать или списывать со счетов огромную разницу между инстинктивным или тропистским поведением с одной стороны и более разносторонним, разумным поведением человеческих и прочих существ с другой — однако в данный момент я настаиваю на том, что, если с целью предсказания поведения безоговорочно принять интенциональную установку, то птица оказывается в той же степени интенциональной системой второго порядка, что и человек. Ввиду этого следует особенно подозрительно отнестись к соблазнительному аргументу, что *репрезентации* интенций второго порядка каким-то образом основаны на языке.<sup>14</sup> Далеко не ясно, необходимо ли для эффективного предсказания поведения системы путем *притисывания* ей интенций, чтобы все или даже какие-либо из этих интенций *репрезентировались* «внутри» системы. В ситуации, которую мы пытаемся прояснить, приводя намерение птицы вызвать ложные представления в хищнике, нет, судя по всему, ни возможности, ни необходимости репрезентации такого рода сложной интенции в «мыслях» или в «сознании» какого-либо существа — обоснованность наших объяснений не зависит от того, думает ли необходимые мысли сама птица, эволюционная история, или Матушка-природа.

Взаимность, таким образом, при условии, что мы понимаем ее просто как способность интенциональных систем иметь интенции более высокого порядка, зависит от первых трех условий, и не зависит от пятого и шестого. Соответствует ли такое понимание взаимности тому, что обсуждают другие авторы, станет ясно, только когда мы рассмотрим, как взаимность переплетается с последними двумя условиями.

В отношении нашего пятого условия, способности к вербальной коммуникации, нам поможет теория смысла Грайса. Грайс пытается определить нечто, что он называет «намеренный (nonnatural)» смысл — смысл, связанный с *намерениями* говорящего, смысл, который говорящий имеет в виду, совершая высказывание. Начальное определение таково<sup>15</sup>:

<sup>13</sup> G. E. M. Anscombe, *Intention* (Oxford: Blackwell, 1957): стр. 80.

<sup>14</sup> Ср. Ronald de Sousa, «Self-Deception», *Inquiry*, XIII (1970), особенно стр. 317.

<sup>15</sup> Ключевые статьи: «Meaning», *Philosophical Review* (July, 1957), и «Utterer's Meaning and Intentions», *Philosophical Review* (April, 1969). Начальная формулировка, приводимая в первой статье, подвергается серии новых редакций во второй статье, из которой и взято настоящее определение (стр. 151).

- «*U* имел нечто в виду, высказывая *x*» истинно, если *U* высказывал *x* для некого адресата *A*, желая чтобы
- (1) — у *A* возникла определенная реакция *r*.
  - (2) — *A* понял (признал), что *U* намеревался, чтобы (1).
  - (3) — (1) произошло в результате (2)...

Заметим, что (2) приписывает *U* интенцию не только второго, но и третьего порядка: *U* должен *хотеть*, чтобы *A* *понял*, что *U* *хочет*, чтобы у *A* возникла реакция *r*. В результате серии контрпримеров Грайсу пришлось перейти от изначального определения к гораздо более сложным его версиям, но это совершенно не важно, ибо все они включают интенцию третьего порядка типа (2). Из грайсовского анализа намеренного смысла можно сделать два важных вывода. Во-первых, поскольку намеренный смысл — смысл, который мы сообщаем путем высказываний — должен быть частью любой вербальной коммуникации, и поскольку он зависит от интенций третьего уровня, наше пятое условие зависит от выполнения четвертого, а не наоборот. Во-вторых, Грайс демонстрирует, что простых интенций *второго порядка* недостаточно для обеспечения действительной взаимности; для нее нужны интенции *третьего* порядка. Грайс вводит условие (2) с целью исключить случаи следующего типа: я оставляю на полу разбитый дочерью фарфор, чтобы его увидела жена. Этот случай отличен от попытки сообщить смысл путем высказывания и намерения, ибо, несмотря на то, что я пытаюсь вызвать у моей жены определенные представления относительно нашей дочери (моя интенция второго порядка), успех здесь не зависит от ее осознания этого моего намерения, а также от ее осознания, что я вообще как-то здесь замешан. Пользуясь удачным термином Эрвина Гофмана, можно сказать, что в этом случае между нами нет реального *контакта*, нет *взаимопризнания*. Чтобы можно было говорить, что произошел акт передачи и усмотрения смысла, акт имения в виду и усмотрения оногo, между говорящим и адресатом должен произойти контакт — но с одной стороны контакт может происходить и в отсутствии намеренного смысла (собака Эшли), а с другой, зависящие от интенций третьего порядка действия не обязательно подразумевают контакт (напр. *A* может хотеть, чтобы *B* посчитал, что *C* желает, чтобы *p*). Таким образом, интенции третьего порядка являются необходимым, но не достаточным условием наличия намеренного смысла, осуществления вербальной коммуникации.

Ситуации намеренного смысла Грайса не случайно попадают в класс, другими членами которого являются обман и манипуляция. Вспомним ловкий контрпример, выдвинутый Серлем против одной из грайсовских формулировок: американец, схваченный во время второй мировой войны за линией фронта в Италии, пытается заставить поймавших его итальянцев принять его за немецкого офицера — и с этой целью произносит единственное известное ему немецкое предложение: «*Kennst du das Land, wo die Zitronen blühen?*»<sup>16</sup> [«Знаешь ли ты край, где цветут лимонные деревья?»] Грайс отмечает, что такие случаи объединяет с ситуациями намеренного смысла элемент экспли-

<sup>16</sup> John Searle, «What is a Speech Act?» в книге Max Black, ed., *Philosophy in America* (London: Allen & Unwin, 1965); это также обсуждает Грайс в статье «Utterer's Meaning and Intentions», стр. 160.

атации рациональности жертвы. Успех тут зависит от того, удастся ли вынудить жертву начать цепь рассуждений, для которой прямо или косвенно выдвигаются посылки. В случае обмана эти посылки сам говорящий истинными не считает; в случае нормальной коммуникации считает. В грайсовской форме, коммуникация есть кооперативная манипуляция адресата говорящим; она зависит не только от рациональности адресата, который должен усмотреть намерения говорящего, но и на *доверии* адресата к говорящему. В качестве особого рода манипуляции, даже при условии необходимой рациональности адресата, коммуникация не достигала бы успеха, не будь доверие адресата к говорящему *обоснованным*, иными словами, разумным. Нормой высказываний, таким образом, является искренность; высказывания не могли бы выполнять свои основные функции, если бы в нормальном случае не вызывали доверия.

Ложь как форма обмана, может осуществляться только на фоне правдивости, однако другие формы обмана от доверия жертвы не зависят. В этих случаях успех связан с тем, что жертва *достаточно* умна, но все же не вполне. Глупые игроки в покер — бич умных игроков, ибо они не видят рассчитанных на них блефов и уловок. Сложные обманы не обязательно основаны на прямом контакте. Существует книга о том, как распознавать поддельный антиквариат (неизбежно являющаяся руководством его изготовления) — и эта книга дает следующий хитрый совет как провести покупателя-«эксперта»: завершив изготовление столика или любой другой вещи и воспользовавшись всеми средствами подделки ее возраста и степени износа, нужно взять обычное сверло и просверлить дырку в каком-нибудь заметном, но странном месте. Покупатель будет рассуждать так: никто без причины не стал бы сверлить тут такую кошмарную дырку (ее никоим образом нельзя заставить выглядеть «аутентично»); следовательно, дырка эта была зачем-то нужна; это означает, что столиком где-то пользовались, а если им где-то пользовались, это значит, что он не был изготовлен специально для продажи в этой антикварной лавке... следовательно, он подлинный. Даже если подобное «заклучение» не до конца уничтожит сомнения, покупатель будет настолько занят продумыванием возможных способов использования дырки, что потребует несколько месяцев, чтобы у него снова возникли сомнения.

Важным в подобном обмане, как и в случае с притворяющейся птицей, является то, что успех не зависит от наличия у жертвы *осознаваемой* цепи рассуждений. Все будет также, если покупатель просто заметит дырку и у него «возникнет ощущение», что вещь подлинная. Позже, он, *возможно*, примет те аргументы, которые придут ему в голову в качестве «обоснования» решения о подлинности, но может случиться, что он, обманывая себя, будет отрицать эти аргументы, несмотря даже на то, что *мысли эти* вообще никогда не приходили ему в голову. Цепь рассуждений объясняет, почему дырка срабатывает (если она срабатывает), но, как утверждает Анскомб, рассуждения не обязаны «описывать действительные ментальные процессы», если под этим понимать сознательные процессы или события. То же самое, конечно, справедливо и для коммуникаций Грайса; ни говорящий, ни адресат не обязаны сознательно иметь им описываемых сложных интенций — удивительно, что никто до сих пор не использовал этого в качестве возражения про-

тив его теории. Условия Грайса часто критиковались в отношении их недостаточности, однако можно выдвинуть еще не нигде не использовавшийся аргумент, что они не являются даже необходимыми. Крайне редкие люди сознательно продумывали эти сложные намерения до того, как их описал Грайс, человеческая коммуникация тем не менее имеет место уже довольно давно. До Грайса, столкнувшись с вопросом: «Хотели ли вы, чтобы ваш собеседник понял ваши намерения, дабы вызвать у него эту реакцию?», ответ, скорее всего, был бы: «Ничего такого замысловатого у меня на уме не было. Я просто хотел сообщить, что меня не будет дома к ужину» (или что-то в этом роде). Поэтому, если эти сложные намерения и являются постоянной основой наших коммуникаций, они, судя по всему, должны быть бессознательными. Действительно, на статьи Грайса наиболее естественно отреагировать, замечая, что *человек никогда не осознает*, что осуществляет все в них описанное в ходе коммуникации. Согласно мощным аргументам Анскомб, такая реакция показывает, что действие (в подобной дескрипции) не является *намеренным*.<sup>17</sup> И, поскольку человек, порождая высказывания, не *осознает* этих намерений, нельзя считать, что он *говорит с этими намерениями*.

Почему же никто не использовал подобный аргумент против Грайса? Потому, мне кажется, что Грайс явно наткнулся здесь на нечто глубокое — его теория действительно дает нам необходимые условия намеренного смысла. Его анализ проливает свет на столько проблем! Вступаем ли мы в коммуникацию с компьютером на языке Фортран? Фортран выглядит как язык; у него есть грамматика, словарь, семантика. Взаимодействия на Фортране между оператором и компьютером часто рассматриваются как случай *человека, вступающего в коммуникацию с машиной*, однако такие взаимодействия являются лишь бледными копиями человеческих речевых коммуникаций — именно потому, что грайсовские условия намеренного смысла в этом случае обходятся. Их просто не к чему приложить. Достижение цели при передаче порции Фортрана машине не основано на признании машиной наших *намерений*. Это не означает, что в будущем подобный недостаток (или преимущество, в зависимости от задач) будет характерен для всех наших взаимодействий с компьютерами; это просто означает что в настоящее время мы в сильном (грайсовском) смысле с компьютерами в коммуникацию не вступаем.<sup>18</sup>

Если мы не хотим отказаться от модели Грайса, и при этом в обычных разговорах не наблюдаем никаких подобного рода *намерений*, следует отправить эти *намерения* в подполье и назвать их бессознательными или предсоз-

<sup>17</sup> См. G. E. M. Anscombe, *Intention* (Oxford: Blackwells, 1957), стр.11. Здесь, а также в нескольких следующих абзацах, я использую слова «intentions» и «intentional» в их *обычном* смысле — и чтобы подчеркнуть это, даю их курсивом.

<sup>18</sup> Ховард Фридман заметил мне, что большое количество современных фортрановских компьютеров, «поправляют» оператора, вставляя знак «плюс», скобки, и т. д., дабы сформировать правильные выражения — может показаться, что они удовлетворяют критериям Грайса, ибо в узких пределах они диагностируют *намерения* «говорящего» и действуют на основе этого диагноза. На это можно ответить, что, во-первых, современные машины могут диагностировать только так называемые синтаксические намерения оператора, а во-вторых, эти машины судя по всему не удовлетворяют следующим, более детальным определениям Грайса; при этом мне не хотелось бы создать впечатление, что я утверждаю, что никакой компьютер этого сделать не может.

нательными. Эти *намерения* содержат «порядок, имеющий место», когда происходит коммуникация, — они являются ее предварительным условием, и в нормальном случае мы их не осознаем.

Заметим, что до настоящего пункта нам был вообще не нужен никакой тип сознания, и поэтому, если и существует некая концептуальная зависимость между сознанием или самосознанием с одной стороны и нашими остальными условиями с другой, то именно сознание должно вытекать из остальных условий, а не наоборот. Но чтобы показать это, мне нужно прежде продемонстрировать, как первые пять условий сами по себе играют определенную роль в этике, что следует из теории справедливости Роулза. Центральным в этой теории является описание некой идеализированной ситуации, «исходного положения (*original position*)», в котором находятся идеализированные личности; из этой идеализации выводятся первые принципы справедливости, проливающие затем свет на остальную часть теории. В данный момент меня не интересует ни содержание этих принципов, ни обоснованность их вывода; меня интересует лишь тактическая природа роулзовского подхода. Роулз воображает группу идеальных субъектов, определяемых им как рациональные, заинтересованные в собственной выгоде сущности; каждая из них в определенных рамках производит расчет эффектов взаимодействия своих индивидуальных интересов с антагонистическими им интересами других (что требует образования интенций высокого порядка, например, представлений о желаниях других, представлений о представлениях других по поводу их собственных желаний, и так далее). Роулз утверждает, что подобный расчет имеет оптимальное «решение», которое для всех заинтересованных в собственной выгоде субъектов является рациональной альтернативой гоббсовскому природному состоянию. Это решение заключается в подчинении, вместе с членами группы, очерченному Роулзом принципам справедливости. Какого же рода доказательство этих принципов предлагает автор? Принятие их, утверждает Роулз, можно рассматривать в качестве решения «игры самого высокого порядка» или «проблемы переговоров». Это подобно выводам теории игр и доказательствам в эпистемической логике Хинтикки<sup>19</sup>; это сходно также с «демонстрацией» того, как шахматная программа делает тот или иной ход, который является наиболее рациональным при наличии информации, доступной программе по ходу игры. Все основано на предположении об *идеально* рациональных агентах расчета; решения таких агентов являются существенно нормативными. Вывод из исходного положения Роулза видится мне, поэтому, как непрерывно перетекающий в те выводы и экстраполяции, с которыми мы имеем дело в более элементарных случаях обращения к интенциональной установке с целью объяснения и управления поведением простых сущностей. Для представлений нормативны истинность и последовательность, для высказываний нормативна искренность; подобным же образом, если Роулз прав, для межличностных отношений нормативна справедливость — в том виде в каком он ее определяет. Частью нашего взгляда на нечто как на обладающее представлениями и другими интенциями является возможность считать это нечто рацио-

<sup>19</sup> См. Hintikka, *Knowledge and Belief*, стр. 38.

нальным — совершенно также частью нашего отношения к чему-то как к личности является возможность рассматривать это нечто как подчиняющееся принципам справедливости. Одним из способов лучше осознать обсуждаемую здесь своеобразную концепцию личности будет сказать, что, хотя Роулз и не пытается доказать, что справедливость есть неизбежный результат взаимодействия между *человеческими существами*, он в результате доказывает, что справедливость есть неизбежный результат взаимодействия между *личностями*. Иными словами, понятие личности необходимо нормативно и идеально; в той мере, в какой справедливость не проявляется в отношениях и взаимодействиях существ, эти существа не являются личностями. Тут мы снова видим, что «порядок, имеющий место» в справедливо устроенном обществе, не зависит от любых реальных эпизодов сознательной мысли. Существование справедливых практик и скрытое в них «признание» не основано на чем бы то ни было сознательном или намеренном расчете вокруг идеального исходного положения, не связано ни с какими взаимными договорами и сознательными установками в отношении других участников.

Признание другого как личности основано на реакции и поведении по отношению к нему определенного типа; этот тип связан с различными *prima facie* обязательствами. Признание до определенной степени этих обязательств, и, таким образом, элемент моральности, не есть вопрос ни выбора, ни интуитивного усмотрения моральных качеств, ни выражения чувств или отношений... это просто следование определенной форме поведения, в которой выражается признание других в качестве личностей.<sup>20</sup>

Важность попытки Роулза вывести принципы справедливости из «исходного положения» проистекает, безусловно, из того, что, несмотря на узнаваемо *моральную* нормативность ее результата, *сам вывод* не есть вывод моральной нормы. Моральность за участниками исходного положения не предполагается. Это в свою очередь означает, что сам по себе вывод нормы не дает ответа на вопрос как и когда у нас возникает право считать личность *морально ответственной* за отклонения от этой нормы. Здесь нам помогает Анскомб, одновременно вводя наше шестое условие. *Если меня следует считать ответственным за мои действия* (за какую-то часть моего поведения в определенной дескрипции), я должен осознавать это действие в данной дескрипции<sup>21</sup>. Почему? Потому что только если я осознаю действие, я могу *сказать*, в чем состояло мое намерение, и могу таким образом в качестве привилегированного участника включиться в игру вопросов и ответов, представляющую резоны моих действий. (Если у меня нет привилегированного статуса в ответах на вопросы о резонах моих действий, то нет особого повода задавать эти вопросы *мне*). Важность способности участвовать в такой игре заключается в том, что только тех, кто может давать резоны, можно рационально убедить принять какой-либо образ действий или отношений, или от-

<sup>20</sup> J. Rawls, «Justice as Reciprocity», в книге S. Gorovitz, ed., *Utilitarianism* (Indianapolis: Bobbs Merrill, 1971): 259.

<sup>21</sup> Меня можно считать ответственным за события и положения дел, которых я не осознавал, но должен был осознавать, однако это не связано с *намеренными* действиями. В подобных случаях я отвечаю за предсказуемые последствия моих осознаваемых действий — включая акты недеяния.



говорить от него, и если некто оказывается не в состоянии «внимать голосу разума» в каком-то деле, то его нельзя считать в нем ответственным. Именно поэтому способности к вербальной коммуникации и осознанию своих действий центральны для всех, кто поддается аргументам и убеждению, и такая убеждаемость, такая взаимная регуляция интересов, осуществляемая совместным использованием рациональности, есть свойство оптимального режима личностного взаимодействия.

Эта способность к участию в совместном использовании рациональности есть основание для еще одного условия личности, недавно выявленного Гарри Франкфуртом<sup>22</sup>. Франкфурт утверждает, что личности являются подклассом интенциональных систем способных к тому, что он называет «*волеиями (volitions) второго порядка*». С первого взгляда это выглядит просто как класс интенциональных систем второго порядка, но ниже мы увидим, что это не так:

Кроме желания, решения и согласия *делать* то или это, люди также могут хотеть иметь (или не иметь) определенные желания и мотивы. Они способны хотеть быть иными в своих предпочтениях и целях, отличными от того, чем они являются... Никакие другие животные кроме человека, похоже, не обладают способностью к рефлексивной самооценке, которая проявляется в формировании желаний второго порядка (с. 7).

Франкфурт указывает, что существуют случаи, где о человеке можно сказать, что он хочет иметь определенное желание, даже если он не хочет, чтобы это желание заработало, стало «его волей» (Человек, к примеру, может хотеть испытывать желание употребить героин с целью просто понять как именно субъективно переживается желание героина, совершенно не стремясь к тому, чтобы оно стало его действующим желанием). В более серьезных случаях человек хочет желать того, чего в настоящий момент не желает, и хочет также, чтобы это желание стало его волей. Эти случаи Франкфурт называет волеиями второго порядка, и утверждает, что именно они «*существенны для возникновения личности*» (с. 10). Я постараюсь как можно яснее изложить здесь его аргумент. Он начинается с анализа разницы между обладанием свободой действия и свободой воли. Согласно этому анализу, свободой воли обладает тот, кто обладает желаемой волей, тот, чьи волеия второго порядка удовлетворимы. Воля личностей не всегда свободна, и в некоторых обстоятельствах личности отвечают за действия, совершенные без свободы воли, однако при этом личности всегда должны быть «*сущностями, для которых свобода их воли проблемна*» (с. 14) — иными словами, сущностями, способными к формированию волений второго порядка, неважно удовлетворимых или нет. Франкфурт вводит замечательный термин «*распущенные (wanton)*» для тех, «*у кого есть желания первого порядка, но... нет волений второго порядка*». (Волеия второго порядка по Франкфурту являются, конечно, *рефлексивными* желаниями второго порядка).

Франкфурт утверждает, что мы интуитивно считаем всех животных, а также детей и некоторых умственно дефективных людей «*распущенными*», и тут я не могу придумать ни одного контрпримера. Сила этой теории в том,

<sup>22</sup> Н. Frankfurt, «Freedom of the will and the concept of a person», там же см. сноску 3. Франкфурт не говорит, рассматривает ли он это условие как только необходимое или же и достаточное для возникновения моральной личности.

что, согласно Франкфурту, человеческие существа (единственные признаваемые нами личности) отличаются от животных именно в этом отношении. Но чем же особенны воления второго порядка? Почему именно они среди всех высших интенций отмечают область личности? Это связано, я полагаю, с тем, что «рефлексивная самооценка», о которой говорит Франкфурт, как раз и есть настоящее самосознание, достигаемое путем занятия по отношению к *самому себе* установки не только участия в коммуникации, но задавания вопросов и убеждения. Как замечает Франкфурт, желания второго порядка остаются пустым понятием, если человек не может в соответствии с ними *действовать* – действие же в соответствии с желанием второго порядка должно логически отличаться от действия в соответствии с его компонентом первого порядка. Действие в соответствии с желанием второго порядка (операции с целью занять определенное желание первого порядка) есть действие, направленное на самого себя в том же смысле, в каком действие может быть направлено на другого: человек *воспитывает* себя, убеждает себя, выдвигает аргументы, угрожает себе, подкупает себя в надежде привести к возникновению в себе желания первого порядка<sup>23</sup>. Установка по отношению к себе и *доступ к себе* в этих случаях существенно подобны установке и доступу по отношению к другому. Тут необходимо *спрашивать себя* каковы твои истинные желания, мотивы, резоны, и только если человек в состоянии осознать свои желания и ответить на такой вопрос, он может вызвать в себе изменения.<sup>24</sup> Мне кажется, что начиная именно с этого пункта, «имеющий место порядок» невозможен, не будучи переплетен с сознательной мыслью, находящейся в диалоге с самой собой.

Нельзя ли тут утверждать, что все вместе эти необходимые условия моральной личности являются также и достаточными? Этого утверждать нельзя, и причина тут, как я попытался показать выше, в неизбежной нормативности понятия личности. Человеческие существа и другие сущности могут только стремиться приблизиться к идеалу, но нет никакого способа установить «проходной балл», который не был бы произволен. Даже если все шесть условий (в жесткой интерпретации) рассматривать как достаточные, то это все равно не обеспечит наличие у чего-либо личности, ибо тогда вообще не найдется ничего, что им бы удовлетворяло. Моральное и метафизическое понятия личности есть не отдельные несвязанные концепции, но две различные подвижные точки одного континуума. Относительность пронизывает критерии присутствия личности на всех уровнях. Объективно удовлетворимого достаточного условия того, что нечто или некто действительно обладает представлениями, не существует, и по мере того, как под

<sup>23</sup> Мое внимание было обращено на тот факт, что собаки часто мастурбируют во время течки, явно с целью *усилить желание* к совокуплению. Такого рода случаями можно пренебречь, ибо, даже предполагая, что собака работает над своим желанием, стремясь его усилить, эффект достигается неинтенциональным ('чисто психологическим') образом; собака не обращается к своей рациональности, не использует ее с целью достижения результата. (Подобно тому, как если бы для человека единственным способом работы над волением второго порядка было принятие таблетки или стояние на голове, и т. п.)

<sup>24</sup> Margaret Gilbert, в статье «Vices and Self-Knowledge», *Journal of Philosophy*, LXVIII (August 5, 1971): 452, рассматривает следствия того факта, что 'только если человек считает, что у него есть какая-либо черта характера, он может решить измениться так, чтобы ее не было.'

интенциональной интерпретацией какой-либо сущности обнаруживается иррациональность, размываются основания для приписывания этой сущности каких бы то ни было представлений, особенно если имеется неинтенциональное, полностью механическое объяснение (а такое объяснение в принципе *возможно* всегда). Сходным образом, наше предположение, что нечто есть личность, оказывается под сомнением именно в тех случаях, когда это действительно важно — когда причинен ущерб и встает вопрос об ответственности. В такой ситуации основания для утверждения, что человек виновен (свидетельства причиненного ущерба, свидетельства того, что человек осознавал, что причиняет ущерб, а также, что он делал это свободно), являются по природе своей также основаниями сомневаться, что мы имеем дело с личностью. И если задаться вопросом, что могло бы *разрешить* наши сомнения, то ответ будет — ничего. Когда встают такие вопросы, мы даже относительно самих себя не можем сказать, являемся ли мы личностями.

*Перев. с англ. Григория Хасина*