

## Онтология от первого лица и функционализм

Знаменитый мысленный эксперимент Дж. Сёрла, его аргумент «китайской комнаты» имеет серьезные последствия для функционализма и компьютерных теорий сознания и значения. В частности, аргумент нацелен на демонстрацию того, что компьютер мог бы даже пройти тест Тьюринга, но, в то же время, пользовался бы языком чисто синтаксически, не владел бы им. Этот аргумент, таким образом, демонстрирует, что компьютер не может понимать, не может «схватить» значения языковых выражений.

Кратко говоря, аргумент «китайской комнаты» заключается в следующем [1]. В некую комнату помещен человек, который не знает китайского языка, но в этой комнате имеется коробка с китайскими символами (база данных) и инструкция по манипулированию данными символами (программа). Извне этот человек получает некие послания на китайском языке (ввод данных) и корректно отвечает на них в соответствии с инструкцией (вывод данных). Как уже было указано, этот весьма странный объект — комната и заключенный в ней человек, не знающий китайского языка, но имеющий инструкцию, — проходит тест Тьюринга. Согласно Дж. Сёрлу, так как данный странный объект схож функционально с обычным компьютером, то компьютер не может обладать знанием китайского языка. По мнению Дж. Сёрла, этот аргумент имеет весьма далекие следствия. В особенности, как полагает Дж. Сёрл, его аргумент является опровержением такого понимания сознания, которое отождествляет сознание с системой обработки информации, подобной компьютеру (вариант функционализма).

Аргумент «китайский комнаты» нацелен на опровержение сильного ИИ, или «компьютерного функционализма» [2], точки зрения утверждающей, что компьютер способен понимать естественный язык и обладает ментальными способностями, сходными с человеческими при условии наличия соответствующей программы; точки зрения, что формальные манипуляции с символами производят мысль. Второй важный вывод, поддерживаемый Дж. Сёрлем, заключается в том, что наличие и даже работа программы (в случае китайской комнаты — инструкция и ее использование) не дает понимания, что знание семантики не выводится из знания синтаксиса.

Тем не менее, мысленный эксперимент Дж. Сёрла представляется весьма неоднозначным. С одной стороны, как уже указывалось, он, по мнению многих философов-аналитиков, утверждает, что понимание языка не тождественно умению механического манипулирования символами. С другой стороны, аргумент «китайской комнаты» может использоваться как доказательство необоснованности радикальной редукции значений к ментальным процессам (подобно другим мысленным экспериментам, например, Х. Патнэма с Землей-Двойником [3] и Д. Дэннета с различием между лошадьё и смошадью [4]). Последнее положение Дж. Сёрл, думается, принял бы лишь частично. В своих работах по теории речевых актов он пытается занять «примирительную» позицию между подходами к значению позднего Л. Витгенштейна [5] и П. Грайса [6], утверждая, что семантика языка есть система конститутивных правил, а речевое общение — игры [7, с. 60, 73], оставаясь при этом в рамках психологической семантики, — для Дж. Сёрла значение выражения в *языке* есть конвенциональная фиксация интенций. Но есть ещё и значение, которое пытался донести до слушателей говорящий, — его намерение. И для Дж. Сёрла, вслед за П. Грайсом именно последнее является базисом языка, а конвенции — его

надстройкой. П. Грайс отстаивает чистый психологизм значений, Дж. Сёрл же — умеренный психологизм. Основа языка у Дж. Сёрла лежит в намерениях, желаниях, говоря более широко, — в первичной интенциональности. Синтаксис опирается на «замороженную», «фиксированную» интенциональность. И то, что «понимает» компьютер, есть лишь эта вот вторичная интенциональность, а потому это «понимает» всегда в кавычках.

Пусть узник китайской комнаты и не знает китайского языка, но он может вступать в коммуникацию на китайском языке, «обращаясь за значениями» к окружающим факторам — набору китайских символов и инструкции.

Для понимания аргумента «китайской комнаты» наблюдателю должна быть доступна возможность занять позицию узника китайской комнаты и сравнить его «ментальную» жизнь с жизнью человека, обладающего знанием китайского языка. Строго говоря, для того, чтобы понимать (даже чтобы понимать сам аргумент Дж. Сёрла), необходимо обладать первичной интенциональностью. Того, что можно назвать «количественными» отличиями между человеком, обладающим пониманием, и компьютером, этим пониманием не обладающим, весьма много. Но аргумент «китайской комнаты» претендует на то, чтобы показать «качественное» отличие. В действительности, аргумент «китайской комнаты» предполагает немой вопрос: «Видите разницу?» Разница между человеком, понимающим китайский язык, и компьютером, оперирующим чисто синтаксически, безусловно, существует. Но аргумент Дж. Сёрла не демонстрирует сути этого различия. В соответствии с аргументом «китайской комнаты» понимание должно сопровождаться первичной интенциональностью. Она не свойственна компьютеру. Но из этого вовсе не следует, что первичная интенциональность свойственна человеку.

Наблюдатель со стороны никогда не сможет однозначно решить вопрос о том, понимает ли узник китайской комнаты китайский язык или же нет. Этот странный объект — «китайская комната» — запросто проходит тест Тьюринга. Мысленный эксперимент Дж. Сёрла не доказывает существование принципиального различия между человеком и компьютером. Аргумент «китайской комнаты» должен был продемонстрировать «силу» онтологии сознания от первого лица. Предвкушалось, что, коль скоро сознание человека не будет для нас больше «черным ящиком», мы неизбежно обнаружим его «сущностные свойства» [8, с. 21], убедимся, что компьютер не может мыслить, и если человек *понимает* язык, то компьютер способен лишь *оперировать* символами. Думается, то, что действительно эксплицитно демонстрирует аргумент «китайской комнаты», так это, что дело не в бихевиористской установке. Интенциональность остается неуловимой даже для онтологии сознания от первого лица. Вспомним, например, что Джонс из мысленного эксперимента Д. Дэннета не может, находясь на ипподроме, понять, где лошадь, а где — смошадь.

В действительности, аргумент «китайской комнаты» лишь допускает то, что следовало бы доказать: что качественное отличие между понимающим человеком и оперирующим компьютером существует. Ведь то, что узник китайской комнаты не обладает интенциональностью, не доказывает, что все остальные люди (или даже какая-то часть всех людей, например, только китайцы) этой интенциональностью обладают. Фактически данным аргументом доказывается отсутствие понимания китайского языка и интенциональности у узника китайской комнаты, а, следовательно, и у компьютера, а также методологическая ущербность бихевиоризма. Однако ошибочно полагать, что этот аргумент имеет силу против сторонников функционализма, так как последние не только могут обойтись без бихевиористского основания, но и существенно обогащают свою позицию доводами «китайской комнаты», получая возможность утверждать: «Да, узник китайской комнаты не обладает интенциональностью, компьютер не обладает интенциональностью, но

и остальные объекты (люди, животные, боги) тоже не обладают интенциональностью». Фактически перед функционалистами, для последовательного разворачивания их программы, требовалось доказать вовсе не наличие интенциональности у компьютера и человека (как это могли предполагать некоторые эмерджентисты), а вовсе наоборот — её отсутствие в обоих случаях. И в этом преградой для них было бихевиористское наследие, отбросить «методологические оковы» которого им помог их самый ревностный оппонент введением онтологии сознания от первого лица.

### Литература:

1. *Searle J.* The Chinese Room // *Wilson R.A., Keil F.* The MIT Encyclopedia of Cognitive Science. Cambridge, Mass.: The MIT Press, 1999. P. 115—116.
2. *Dennett D.* Brainstorms: Philosophical Essays on Mind and Psychology. Cambridge, Mass.: The MIT Press, 1978.
3. *Putnam H.* The Meaning of «Meaning» // *Putnam H.* Mind, Language and Reality. Philosophical Papers. Vol. 2. Cambridge: Cambridge University Press, 1975. P. 215—271.
4. *Dennett D., Haugeland J.* Intentionality // *Gregory R.L.* The Oxford Companion to the Mind. Oxford: Oxford University Press, 1987. P. 383—386.
5. *Витгенштейн Л.* Философские исследования // *Витгенштейн Л.* Философские работы. Ч. I. М.: Гнозис, 1994. С. 75—319.
6. *Грайс Г.П.* Значение говорящего, значение предложения и значение слова // Философия языка. М.: Едиториал УРСС, 2004. С. 75—98.
7. *Сёрл Дж.Р.* Что такое речевой акт? // Философия языка. М.: Едиториал УРСС, 2004. С. 56—74.
8. *Сёрл Дж.* Открывая сознание заново. М.: Идея-Пресс, 2002.