Кажется, что понимает

Машина вас слушает — или вам так кажется?

Перед вами два собеседника. Слева — ELIZA, программа 1966 года. Справа — современная языковая модель. Напишите обоим одно и то же: расскажите про свой день, пожалуйтесь на что-нибудь. И посмотрите, где разговор держится, а где рассыпается.

ELIZA внутри почти пуста. Она выхватывает из вашей фразы знакомое слово и переворачивает его обратно вопросом: на «меня никто не понимает» отвечает «почему вы думаете, что вас никто не понимает?». Смысла она не извлекает, про мир ничего не знает, через минуту забывает начало разговора. И всё же первые реплики легко принять за внимание живого собеседника.

Её автор Джозеф Вейценбаум собрал ELIZA, чтобы показать, до чего проста машинная имитация разговора, — а вышло обратное: его секретарша, прекрасно зная, что перед ней программа, просила оставить её с «доктором» наедине[1]. Ощущение, что тебя понимают, держится не на одном собеседнике. Половину работы делает наш собственный мозг, готовый увидеть мысль за подходящими словами.

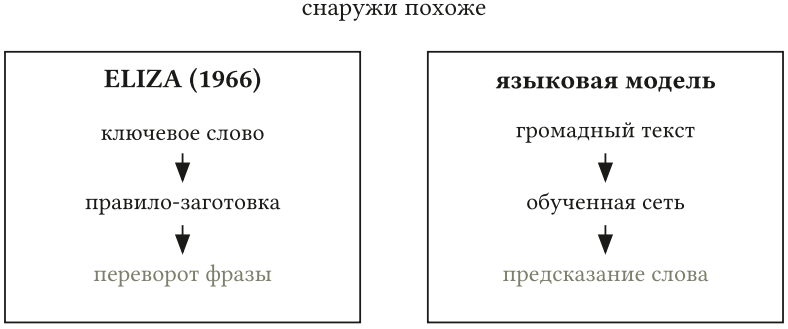

Рис. 1. Слева ELIZA: ключевое слово запускает правило-заготовку, и фраза переворачивается обратно вопросом. Справа языковая модель: ответ рождается как предсказание следующего слова обученной сетью. Снаружи похоже, внутри несравнимо.

Современная модель справится с тем, на чём ELIZA спотыкается: удержит нить, ответит по существу, подстроится под ваш тон. Но приписать понимание ей тянет

[1] Joseph Weizenbaum, «ELIZA — A Computer Program For the Study of Natural Language Communication Between Man and Machine». Communications of the ACM 9(1), 36–45 (1966). Автор описывает, как пользователи приписывали программе понимание, зная о её простоте (doi:10.1145/365153.365168).

совершенно так же, как шестьдесят лет назад. Разница между двумя терминалами — в том, что у них внутри. Сходство — в том, что происходит у вас в голове.