

datamaskin, lastet med et veldig stort antall programmer, å svare på spørsmål på en slik måte at en uvitende samtalepartner kun lures til å tro den er et menneske?» (Turing, 1950/2009, vår oversettelse).

Turingtesten bygger på en forutsetning om at behandling av informasjon handler om å følge formale regler. Denne oppfatningen har å gjøre med logikken som historisk ligger til grunn for feltet kunstig intelligens. Logikk handler ikke om (menneskelig) betydning av ting, men heller om å relatere meningsløse symboler ved å følge visse regler korrekt. Dette innebærer en bevegelse bort fra semantikk (som handler om mening) mot syntaks (et system av regler som symboler kan kobles sammen med, på en logisk riktig måte). Det betyr ganske enkelt at du ikke trenger å forstå meningen med ting så lenge du vet hvordan du følger regler om hvordan de kombineres (Clark, 2000, s. 8–10).

Dette regelbaserte synet på intelligens kan kanskje illustreres med et eksempel fra hverdagen. Forstill deg konstruksjonen av et tremøbel, for eksempel en hylle. Man trenger ikke være trespesialist eller vite nøyaktig hva de ulike delene i hylleboksen er, hvis man er i stand til å følge et sett med regler. Man kan (forhåpentligvis) bygge hyllen ved å gjøre trinn for trinn hva som står i monteringsanvisningen.

Det samme gjelder en datamaskin. Så lenge den gjør noe som ser ut som intelligens, kan det ifølge Turing være tilstrekkelig for å betraktes som intelligent. Oppsettet for Turingtesten er et såkalt imitasjonsspill hvor målet er å se hvor godt en datamaskin kan etterligne menneskelige ferdigheter. Spillet ser slik ut: Det er tre «aktører». Aktør A, en datamaskin, aktør B, et menneske og aktør C, et annet

menneske (testpersonen). Alle tre befinner seg i forskjellige rom og kommuniserer med hverandre via tekstdata. Spørsmålet er om C klarer å skille mellom datamaskinen A og det andre mennesket B etter tilstrekkelig lang kommunikasjon med begge. I sin tid antok Turing at ved slutten av hans århundre (omkring år 2000) ville datamaskiner ha så stor datakraft at de ville klare seg godt på denne testen. Så bra at en normal person ikke ville være i stand til å skille med 70 % sikkerhet om han for øyeblikket kommuniserte med en datamaskin eller et menneske. Han antok at på det tidspunktet ville vi mennesker også finne det helt akseptabelt hvis noen sa at maskiner kan tenke (Turing, 1950/2009).

Hvordan ser situasjonen ut i dag? Er datamaskinene våre nå sterke nok til å bestå Turingtesten? Inntil nylig så det ikke slik ut. Til den årlige konkurransen om «Loebnerprisen» sendes dataprogrammer inn som kandidater til Turing-testen. Foreløpig har ingen lyktes. Men vi er kanskje ikke så langt fra å lage en datamaskin som kan føre en troverdig samtale med oss mennesker. GPT-3 (Generative Pre-Trained Transformer) er en språkprogramvare som er i stand til å generere alle slags menneskelige tekster, som poesi og historier (Oppy & Dowe, 2021). Kvaliteten på selve tekstene kan vanskelig skilles fra dem som er skrevet av mennesker (Johnson & Izhev, 2022).

Dette er spesielt tilfelle med den siste utviklingen som chatGPT. Nyere studier viser at tekstene skrevet av denne AI, ikke kan skilles fra tekster skrevet av mennesker, heller ikke av eksperter (Holly, 2023). I studien fikk forskere chatGPT til å skrive sammendrag for medisinske artikler, som deretter skulle kontrolleres for ekthet av både plagiatprogramvare og fageksperter. Plagiatprogramvaren fant