

Turing i la intel·ligència

Carlo Sala Gancho
Història de les Matemàtiques
Grau de Matemàtiques
Universitat Autònoma de Barcelona

Juny 2023

El geni incomprès

Alan Turing [1] (Londres, 1912–1954) estudià a fons la computació des d'un punt de vista matemàtic, fins i tot abans que els primers ordinadors tant analògics com digitals fossin creats (l'ENIAC, considerat el primer ordinador com a tal, fou construït al 1946 i aquest text fou publicat al 1950). En aquest text Turing prova d'esclarir si els ordenadors són, o no, intel·ligents. És una pregunta curiosa, i encara vàlida avui en dia. A més, amb l'aparició de ChatGPT i l'ús cada cop més intensiu de les IAs pel gran públic que hem viscut aquests últims anys, sembla que ha revifat i ha omplert els titulars dels mitjans de comunicació.

Cal comprendre la vida de Turing per entendre les idees que va tenir. Des de ben petit va mostrar senyals de tenir un talent especial per les ciències i la tecnologia, però l'educació formal britànica mai no va fer per ell. Obtenint qualificacions mediocres, no va poder accedir al *Trinity College*, tot i que seguia excel·lint en les seves assignatures preferides. Va treballar pel servei d'intel·ligència britànic, ajudant a descryptar la màquina ENIGMA [3]. Posteriorment, durant l'inici dels 50, va ser perseguit per la seva orientació sexual, arribant a sotmetre's a un tractament (conegut popularment com a *castració química*) molt estès en aquella època per evitar la presó. Va acabar suïcidant-se l'any 1954.

El joc de la imitació

Per poder descobrir si les màquines pensen, primer cal definir les paraules *màquina* i *pensar*, que sembla del tot trivial de definir, però no ho és en absolut. Definirem (tal i com fa Turing) una màquina com un ordinador digital, és a dir, una màquina amb memòria, unitat de processament amb una taula d'instruccions (sumar, llegir en memòria, escriure en memòria, etc) i una unitat de control, que comprovi que les instruccions demanades s'executen correctament. Tenint en compte aquesta definició, podem considerar els ordinadors les màquines universals, ja que adequadament programats, poden dur a terme qualsevol procés computacional que pogués executar una màquina no universal.

Turing proposa el següent joc:

- Siguin A una màquina, B un humà i C un interrogador.
- L'interrogador no veu ni sent els subjectes.
- L'interrogador pot formular preguntes que seran respostes en paper i escrites a màquina per evitar que es pugui extreure informació de la cal·ligrafia dels subjectes, i no només del contingut de les respostes.

El que Turing vol veure, és si l'interrogador serà (o no) capaç de discernir quin dels dos subjectes és la màquina.

Sembla un procés senzill i simple, i en certa manera ho és. Tanmateix, en aquest joc només s'estudia la capacitat d'un subjecte d'imitar-ne un altre. Una de les preguntes que queden obertes és si amb un conjunt donat (i finit, a ser possible) de preguntes es podria discernir de forma clara i unívoca si el subjecte és o no intel·ligent.

Els problemes associats al pensament de les màquines

Podria ser que suposés algun problema fonamental, el simple fet que les màquines pensessin? És curiós com, en efecte, hi ha certes premisses en diferents àmbits de la nostra societat que fan del tot incompatible assumir que les màquines pensen (prenent com a certes aquestes premisses, és clar). Turing en llista fins a 9, de les quals en comentarem algunes de les més rellevants.

La raó teològica

Simplificarem molt el problema, i només ens centrarem en la fe cristiana. La Bíblia diu que Déu ha donat ànima a tots els homes i dones, però no als animals i altres cossos inerts (i, per extensió, tampoc a les màquines). Com que es considera que el pensament (i, per tant, la intel·ligència) són una funció exclusiva de l'ànima, cal concloure que les màquines no poden pensar.

Quedar-se amb aquesta idea seria, de ben segur, conformista. Des d'un punt de vista exclusivament logicomatemàtic, no hi ha cap demostració formal que digui que Déu existeix i que el que diu la Bíblia és cert, caldria prendre-ho com a axioma per acceptar-ho com a veritat. Turing (que com a dada d'interès, era ateu) no accepta cap d'aquestes premisses, i intenta desmuntar-les amb arguments intrínsecament teològics. Mostra com la Bíblia es va fer servir per refutar les hipòtesis copernicanes, arguments que avui en dia ens semblen del tot falsos i que, en essència, podria estar passant el mateix amb el pensament de les màquines.

La raó logicomatemàtica

La pròpia lògica matemàtica ens porta a pensar que, segons com quedi definit el terme *intel·ligència*, es pot provar que les màquines no poden pensar. Tenint en compte que hem definit el concepte *màquina* com a ordinador, és a dir, una màquina discreta. Encara que considerem un ordinador digital de capacitat infinita, jugant al joc de la imitació hi haurà respostes que mai serà capaç de respondre en un temps finit, o bé que respondrà incorrectament. Això és una conseqüència del teorema d'incompletesa de Gödel [2], tot substituint els sistemes lògics per l'ordinador en qüestió. Altres autors també arriben a altres resultats similars (i equivalents en aquest context), tals com Church, Kleene o Rosser.

Aquest contraargument té un problema essencial i intrínsec del qual serà molt difícil desfer-se'n. Considerant als humans com a éssers intel·ligents, mai no s'ha arribat a demostrar (ni rebatre) que nosaltres no tinguem les limitacions anteriorment mencionades. Per tant, no seria necessari demostrar res relacionat amb la capacitat discreta de les màquines per considerar-les éssers intel·ligents. Ara bé, considerar aquest argument com a fals tampoc tindria cap mena de consistència matemàtica. Caldria demostrar, primerament, que els humans no estiguessin limitats de la mateixa manera i, posteriorment, veure que és una característica necessària pel desenvolupament de la intel·ligència.

El sentit dels contraarguments

Tants els arguments a favor com en contra de si les màquines pensen queden curts, i s'auto-boicoteegen contradient-se a si mateixos, i posant-se pals a les rodes. Sense anar més lluny, l'argument matemàtic parteix de la base que hi ha coses sobre la intel·ligència

humana que no s'han demostrat, i que per tant no s'han d'exigir a les màquines. En resum, molts s'aturen entre banalitats sense arribar a demostrar res, conformant-se amb el "*no en podem estar segurs*". Crec que cal prendre un camí molt més pautat, i amb objectius més petits i realistes. El fet de si pensen (o no) és una pregunta massa gran com per intentar resoldre-la tota de cop, i amb un únic enfocament.

Pensen?

El propi Turing acaba conclouent que no hi ha una visió clara de sí pensen o no ho fan. Més de setanta anys després, seguim exactament igual. Hi ha molts àmbits en els que les màquines superen amb escreix les capacitats humanes, i cada vegada hi haurà més distància. Ara bé, de qui és mèrit? El fet que una màquina sigui capaç de treure un deu a la selectivitat és mèrit de la màquina, o dels humans que la van dissenyar, construir, programar i posar en marxa? D'alguna manera, penso que els ordinadors no deixen de ser extensions que creem per a la nostra ment per millorar-la cada vegada més, i portar-la el més lluny possible. No sabem calcular, amb 10 decimals precisos i menys de 10ms $\sqrt{947}$, però hem estat capaços de construir un enginy fent servir les eines que sí que tenim perquè ho faci per nosaltres.

Aquesta és, en essència, la veritable intel·ligència que ens ha distanciat de la resta d'espècies, i que Turing deixa de banda en el seu text. Centrar-se en, única i exclusivament, les capacitats bàsiques dels humans és quedar-se absolutament curt. La creativitat real (no els grafs pesats i amb certa component aleatòria del *DALL-E*), la feina en equip, el companyonia, les ganes de progressar i l'expressió de les emocions són allò que ens defineix com a humans, i que encara sembla que queda un temps fins que les màquines ens atrapin. Potser sí que pensen, però encara falta que sentin. I ningú nega que no puguin arribar a fer-ho. La ciència, fins ara, no ha tingut límits.

Referències

- [1] *Alan Turing*. Consultat 03/06/2023. URL: https://en.wikipedia.org/wiki/Alan_Turing.
- [2] *Gödel's incompleteness theorems*. Consultat 17/06/2023. URL: https://en.wikipedia.org/wiki/G%C3%B6del%27s_incompleteness_theorems.
- [3] *Màquina ENIGMA*. Consultat 03/06/2023. URL: https://en.wikipedia.org/wiki/Enigma_machine.