

Hermeneutikai megközelítések a deliberatív demokrácia és a mesterséges intelligencia viszonyát illetően

Smrcz Ádám

A közelmúlt antropológiai és evolúcióbiológiai kutatásainak eredményeiből sokan arra következtetnek, hogy az embernek más élőlényekkel szembeni evolúciós lépéselőnye nem elsősorban a nyelvhasználatból, hanem sokkal inkább a *kumulatív kulturális evolúciónak* nevezett jelenségből fakad (Tomasello-Kruger-Ratner 1993; Tomasello 1999). A kumulatív kulturális evolúció elmélete azt mondja ki, hogy kulturális kódok tanítás és tanulás útján halmozódnak fel különböző közösségekben, ennek azonban a nyelvhasználat még csak nem is képezi szükséges feltételét (jóllehet, a folyamatot magát megkönnyítheti). Sokkal inkább elengedhetetlennek tűnik szerintük az, hogy egy közösség tagjai saját magukhoz hasonló mentális állapotokkal rendelkező egyedekként tekintsenek egymásra, vagyis az, hogy társainkat ugyanolyan örülni, félni, éhségérzetet érezni stb. képes lényekként gondoljuk el, mint saját magunkat.

De ezek közül a mentális állapotok közül is kitüntetett fontossággal bírnak az ún. intencionális állapotok a kulturális felhalmozás során. Intencionalitás alatt Arisztotelész óta az elmének azt a képességét értjük, mely révén az képes dolgokra irányulni és dolgokat megjeleníteni a saját maga számára (De Anima II.5; Sorabji 1991, 227-260). A viselkedéstudományokban azonban – részben a múlt század fordulóján felívelő pragmatizmus hatására – némiképp eltérő értelemben használatos ez a kifejezés: jelentésének először a „megjelenítéssel” kapcsolatos, korábban szükségszerű aspektusát veszítette el (James 1884, 188-205; Ratcliffe 2005, 179-202), majd ezt követően vette fel a mai köznyelvben is használatos értelmét, melynek folytán többé nem (csak) mentális aktusainknak, hanem cselekedeteinknek is a valamire való irányultságát jelöli. A kifejezés két értelme tehát nem mond ellent egymásnak, egyszerűen csak más-más aspektusait domborítja ki.

A *kumulatív kulturális evolúció* elméletében mindenesetre a második értelemben vett intencionalitás játszik kitüntetett szerepet, méghozzá azért, mert az elgondolás szerint bármilyen (kulturális, közösségi, politikai) tanulási folyamat során az embereknek azt kell feltételezniük, hogy az adott, elsajátítandó kulturális kód *valamilyen* célra irányul, vagy pedig, hogy az adott kulturális kód birtokosa

valamilyen cél érdekében birtokolja és alkalmazza azt (akkor is, hogyha ez a cél esetleg nem ismert számukra pontosan). Illetve egy másik lehetséges megfogalmazás szerint, az embereknek „meg kell érteniük az eszközhasználatnak vagy egy jelképes gyakorlatnak az *intencionális* jelentőségét – vagyis azt, hogy mire való, vagy hogy mit teszünk vele 'mi', ennek az eszköznek vagy szimbólumnak a felhasználói” (Tomasello 1999, 6 – kiemelés tőlem: S. Á.).

Mennyiben vihet mindez közelebb minket a deliberatív demokrácia és a mesterséges intelligencia viszonyának megértéséhez? A fő problémát az okozza, hogy gépeknek nem tulajdonítunk intencionális mentális állapotokat (Dennett 1987, 13-43; Searle 1980, 417-447) (mely állítás igazságértéke független attól, hogy, a gépek ténylegesen rendelkeznek-e ilyenekkel), ez pedig a kumulatív kulturális tőkefelhalmozás előbb látott alapvetéseinek értelmében megghiúsítana minden olyan, hagyományos értelemben vett deliberációt, aminek célja egy adott közösség tagjainak egymás közötti megértése (Habermas), meggyőzés általi befolyásolása, vagy a közéletben való részvételre való ösztönzése (Rawls).

Előadásomban amellet igyekszem érvelni, hogy a fent említett probléma áthidalhatatlan, ennek érzékeltetéséhez pedig olyan elméletalkotókat (Collingwood, Gadamer, Ricoeur, Skinner) hívok segítségül, akiknek a nevei elvétve sem szoktak a mesterséges intelligenciával kapcsolatos okfejtésekben felmerülni, meglátásom szerint azonban mégis segítségünkre lehetnek a probléma egy ritkán tárgyalt aspektusának megvilágításában. Az előbbieket implicit, vagy néhol explicit vitáinak homlokterében ugyanis az állt, hogy lehetséges-e, és ha igen, milyen lehetőségfeltételek fennállása esetén lehetséges szerzői intenciók adekvát rekonstruálása. Gondolatmenetem során nem a köztük feszülő (olykor egészen számottevő) nézetkülönbségekre koncentrálok majd, hanem a „legkisebb közös többszörös” megtalálására törekszem. Félreteszem ezért a szerzői intenciók adekvát rekonstruálhatóságának kérdését, és helyette azt próbálom megválaszolni, hogy mi szerintük az emberi megértés, és ezáltal tanulás legfontosabb lehetőségfeltétele.

Állításom szerint az utóbbi az intenciótulajdonításban ragadható meg: mindaddig (értelemszerűen) nem is törekedhetünk ugyanis mások intencióinak megértésére, ameddig nem feltételezzük bennük ezeknek meglétét. És mivel gépek esetében – mint azt említettük – nem feltételezünk ilyesmit, a gépi döntéshozatal episztémikus értéke, és ezáltal politikai szerepvállalása összeegyeztethetetlennek tűnik a deliberatív demokrácia kívánalmaival.