الفصل

الرابع

فلسفةُ

الذكاءِ الاصطناعي

ويشتمل هذا الفصل على النقاط التالية:

- ✓ ما هي فلسفة الذكاء الاصطناعي؟
- ✓ هل يستطيع للكمبيوتر أن يفكر؟

الفصل الرابع

فلسفةُ الذكاءِ الاصطناعيّ

ما هي فلسفة الذكاء الاصطناعي؟

قد تتساءلون: لماذا يبدو أن العديد من القضايا مثل الذكاء الاصطناعي يجري حولها الكثير الجدل؟ في حين أن التخصصات الهندسية الأخرى، مثل الهندسة المدنية أو الميكانيكية أو الكهربائية لا تمثلُ هدفًا للنقد الصاخب مثل مختلف الفروع من العلوم الإنسانية.

قد يكون ذلك بسبب السذاجة أو محاولة جذب الانتباه والتمويل، وقد جعل الجمهور غايته محدودةً وذلك بمطالبات واسعة لعموم نتائجها وتحقيق توقعات للمسار المستقبلي للآلة، مما يقال أيضا أن منظمة العفو الدولية لا تشكل تحديات حقيقية للوثيقة الفلسفية والدينية التي تألقتْ حول تفرُّدِ الإنسان ومكانه في الكون.

إن الآلات الذكية تجعلنا نسلط الضوء على الهدف فيما يخص الأسئلة الأساسية حول طبيعة عقولنا، ووجود الإرادة الحرة، وعما إذا كانت العوامل غير البيولوجية يمكن أن يقال بها إنه على قيد الحياة، فاحتمال تسوية العديد من المناقشات التاريخية العميقة هي عملية مثيرة ومخيفة قليلا للعقل الذي يتأملُ مثل هذه القضايا. وفي النهاية، نجد العديد أن من هذه القضايا ينزل إلى المعتقدات الأساسية لدينا عن أنفسنا، وهي أن بعض منها هو

مقاومة التفسير العلمي (مثل وجود هذه الروح البشرية، أو الفكرة الديكارتيكية بأن الأحداث العقلية هي بطريقة أو بأخرى متميزة ومستقلة عن العالم المادي).

باختصار، إن فلسفة الذكاء الاصطناعي تطرح علينا السؤال عما إذا كانت أجهزة الكمبيوتر، أو الأجهزة بشكل عام لها عقل أو تفكر؟، أو أن يقال إن أي شئ ليس من أصل طبيعيّ لا يمكن أن يكون له عقل أو يفكر. الجواب ببساطة، يعتمد على ما تعنيه عن "العقل" و"التفكير". فقد يحتدمُ النقاش في مختلف أشكاله، دون هوادة ودون حل، ويستمر كذلك لعقود، بلا نهاية واضحة.

وإليك بعض من التاريخ والحجج المتعددة التي طُرحت من قِبلِ المؤيدين والنقاد لفكرة أن الآلات يمكن أن تفعل أو تمتلك عقول يمكنها التفكير.

هل يستطيع للكمبيوتر أن يفكر؟

طرح عالمُ الرياضيات الإنجليزي الشهير **"آلان تورينغ"** سؤالًا في هذا الأمر في مقال عام 1950 بعنوان **"آلات الحوسبة والذكاء"**، وفي ذلك، يقترح وضع هذه القضية للتصويت، بناءً على ما يُسمِّيه **"لعبة التقليد"**، وهو يتخيل المحقق في غرفة منفصلة، ويتواصل مع رجل وامرأة فقط من خلال الاتصال الكتابي، في محاولة لتخمين أي المحاورين هو الرجل ومن يكون المرأة، حيث أن الرجل يحاول خداع المحقق في الاعتقاد بأنه هو المرأة، وترك امرأة تعلن عن نفسها في محاولة لمساعدة المحقق على تحديد الهوية الصحيحة.

ثم يدعو **"تورينغ"** القارئ إلى تصور استبدال آلة بالرجل، ورجل بالمرأة، وهذه اللعبة الآن تسمى على نطاق واسع **"اختبار تورينغ "**، وترك جانب المفارقةَ النفسية الرائعة لهذا العالم الشهير الذي يكلف الرجل بإقناع المحقق أنه امرأة، ناهيك عن وضْعِهِ الرجلَ في دور المخادع والمرأةَ كصراف للحقيقة، إن هذا الأمر يذهب إلى السؤال عما إذا كان من المعقول أن الجهاز يمكنه على الإطلاق الفوز بهذه اللعبة ضد الرجل، وهذا الأمر يعني أن الجهاز قد يكلف هو المحقق بالاعتقاد بأنه الرجل، في حين أن الرجل يقول الحقيقة حول من هو.

على النقيض من هذا، ننصرف إلى الاعتقاد الشائع بأن **"تورينغ"** كان يقترح **"امتحان دخول"** لتحديد ما إذا كانت الآلات قد بلغت سن الرشد وأصبحت ذكيةً، إنه في الواقع كان يتكهن بأن لدينا الاستخدام الشائع للمصطلح الذي يفكر في أن يمتد في نهاية المطاف بشكل مناسب ليتم تطبيقه أيضًا على بعض الأجهزة أو البرامج ذات القدرة المناسبة ويمكن تقديره من خلال تمكُّنِهِ من تطبيق هذه المهام.

حدث في نهاية القرن العشرين، تخمين دقيق بشكل ملحوظ وهو أننا نشير بشكل روتيني إلى أجهزة الكمبيوتر على أنها "تفكر"، وعندما نتكلم ننتظر بفارغ الصبر الرد عليهم، السؤال الأصلي، هل يمكن الاعتقاد بأن تلك آلات التي لا معنى لها تستحق كل ذلك النقاش؟. ومع ذلك، نعتقد أنه في نهاية القرن الحالي أن استخدام الكلمات والرأي العام المتعلم سوف يغير كثيرًا لدرجة أنه سيكون بمقدور المرء التحدث عن الآلات التي تفكر دون توقع أن تتناقض.

هل **"تورينغ"** على حق؟ هل هذا السؤال لا معنى له، أو يستحق النقاش؟ وبالتالي، فإن هذا النقاش هو مضيعة للوقت، ومن الواضح أن ذلك يعتمد على ما نعنيه بـ "التفكير" فقد نعتبر التفكير هو القدرة على التلاعب برموز المنطق من الافتراضات الأولية إلى الاستنتاجات.

من هذا المنظور، ينبغي أن تكون برامج الكمبيوتر كما نفهمها حاليًا، قادرة على ذلك التلاعب، وبالتالي قادرة على التفكير، لكن من المؤكد أن مجرد تحريك مخزون من الرموز ليس طفيفًا، فهو كذلك يعني شيئًا أو يفعل شيئًا.

من أهم فروع الفلسفة واللغويات التي تتعامل مع هذا الأسئلة "السيميائية" وهي خاصة بدراسة استخدام الرموز للتفكير والتواصل، ويتم التمييز عادة بين بناء الجملة، مثل قواعد ترتيب ومعالجة الرموز، والدلالات، ومعنى الرموز والقواعد، بينما بناء الجملة من السهل فهمه، فالمناهج ليست كذلك، حتى الخبراء لا يوافقون على معنى **"المعنى"**، حيث تقترح معظم النظريات أن المعنى يتطلب بعض الطرق لربط الرموز لأنفسهم للأشياء التي يشيرون إليها في العالم الحقيقي.

فكل شيء له معنى، فنحن جميعًا نعرف ماذا يعني كل من 1، 2، 3 وماذا يعني 4، يعني في الواقع أن أي شيء أكثر أو أقل من (1) يستمد معناه من كيفية توصيلها بمفاهيم أخرى أو كائنات في العالم الحقيقي، فإذا ربطنا $ بأي مجموعة من أربعة أشياء، فإنها تكون مفيدة للغاية لحل مشاكل معينة من الدلالة العملية الكبيرة، يمكنك التلاعب بالرموز طوال اليوم وهذا لا يعني شيئًا.

فالذكاء الاصطناعي والقيام بشيء ما يتطلب الاتصال بين الممثل في التلاعب بنظام الرمز وشيء خارجي لهذا الممثل في حالة برامج الكمبيوتر، وفي سياق آخر يمكن أنت تقول أن معالجات الرمز لها معنى على سبيل المثال أن تشير كم أنت مدين لشركة الهاتف هذا الشهر، حركة قطعة شطرنج (جسديا أو عمليا) أو روبوت يختار قلم رصاص.

يري **"جون سيرل"**، أستاذ الفلسفة في جامعة كاليفورنيا،أنه يلاحظ وبشكل كبير أن أجهزة الكمبيوترفي حد ذاتها، لا يمكنها "التفكير" بهذا المعنى على الإطلاق، لأنها في أحسن الأحوال لا تعني أو تفعل في الواقع أي شيء.

لكن **"سيرل"** يذهب أبعد من ذلك، ويشير إلى أنه حتى القول بأن أجهزة الكمبيوتر قادرة هي أيضًا على التلاعب بالرموز على امتداد مناهجها، فقد تكون الإلكترونيات **متفاعلة** حولها في الدوائر، ولكن نحن الذين نفسر هذا النشاط كرمز التلاعب، ومن الجدير بالذكر حجة أكثر دهاء طرحها بعض المفكرين البارزين في حين أن تلك الرموز نفسها قد تكون خالية من أي دلالات، **ربما يكون لها** معنى ينشأ من علاقاتها مع رموز أخرى، تمامًا كما هو التعبير عن **(definition)** من كلمة في قاموس مصطلحات أخرى.

بينما نحن نعتبر هذا على أنه مهمة وخطوة للأمام، ويبدو غير محسوس، ولكن يمكن للأجانب الذين يقرؤون قاموسًا أن يستخلصوا الكثير من المعلومات حول اللغة، ولكن ذلك لن يمنحهم فهمًا مرضيًا لما يعني الحب، على سبيل المثال، تعاني خوارزميات تعلُّم الآلة من نفس المفاهيم (على الرغم من أنه ليس الفشل في عوائق الذكاء الاصطناعي)، فهي تعيد تعقيد العالم الواقعي، ولكن من دون اتصال مع هذا العالم، فهي مجرد هيكل غير مستقر.

حجج **"سيرل"** مقنعة وذات صلة من قبل الآخرين، كلها تعطي حالة حسنة بالشعور بالبديهية حتى تقوم بتطبيقها على الناس، ونحن نعتبر أنه من المُسَلَّمِ به أن يفكر الناس، لكن ما الفرق بين الأفكار التي تدور حولك في الدماغ وبين **"البايتات"** المتدفقة في جهاز الكمبيوتر؟ في كلتا الحالتين، تكون المعلومات مستمرة، ممثلة في شكل ما، والتي يمكن أن تسمى بشكل رمزي "إشارات عصبية منفصلة"، يتم معالجتها، وتعود للخارج إشارات عصبية إلى يدك لمفاتيح الضغط على لوحة المفاتيح.

إن برنامج الكمبيوتر لا يمكنه أبداً أن يقوم بأية مهمة معينة، سواء كان ذلك لطلاء لوحات جميلة، أو اكتشاف قوانين الطبيعة، أو عزاءك على فقدان شخص عزيز. لكن البرنامج يحاكي التفكير، ولا يكرر العملية التي تحدث في عقول البشر عندما ينخرطون في هذه الأنشطة.

إن العقل البشري في حد ذاته لا يمكنه التفكير أيضًا، إذا تم فصله عن العالم الخارجي، وهو في هذه الحالة سوف يفعل ما يفعله الكمبيوتر، لكن هذا لا يتناسب مع حدسنا المنطقي حتى لو كنا نجلس في غرفة مظلمة وهادئة، محرومين من كل المدخلات والإنتاج، لا يزال بإمكاننا الجلوس هناك والتفكير.

فإذا كان التلاعب بالرمز هو أساس الذكاء، فإنه يمكن لكل من الناس والآلات التفكير من حيث المبدأ، إن لم يكن في الممارسة اليومية.

ولكن إذا كنت تفضل الحفاظ على الشعور بالراحة فالبشر مميزون، ومختلفون عن الآلات في بعض الأشياء الأساسية.

إن مفهوم التفكير مرتبط بالإنسان بشكل فريد، والآلات هي مجرد ادّعاء لقدراتنا المعرفية، لكن قبل ذلك ضع في عقلك واعتبارك أن هناك مجموعة متراكمة من الأدلة تقلل من حدسنا الواضح فيما يبدو حول قدراتنا البشرية الأساسية، من أجل تمييز الآلة، على سبيل المثال، إننا لدينا بالفعل الإرادة الحرة، هل يمكن للكمبيوتر أن يكون لديه إرادة حرة؟

يعتقد الجميع تقريبًا أن البشر وربما بعض الحيوانات لديها إرادة حرة، ولكن هل يمكن للآلة أو جهاز الكمبيوتر أيضًا أن يكون لديه إرادة حرة؟ للإجابة على هذا السؤال، من الضروري أن يكون لدنيا مفهوم لما نعنيه بالإرادة الحرة.

هناك فلسفة طويلة للذكاء الاصطناعي في التاريخ الفكري والديني للجدل حول الطبيعة ووجود الإرادة الحرة، وعادة ما نعني أنه لدينا القدرة على ذلك بجعل الخيارات مدروسة، وربما تمايلت ولكن لم تحدد من قِبلِ قوىً خارج أنفسنا، الشيء الوحيد الذي يجب مراعاته هو أننا نميز بين الداخل والخارج في فهم الإرادة الحرة، فلدينا التفاف حول ما "نحن" لفصله عن ما "ليس نحن"، لكن هذا وحده ليس كافيًا.

يجب أن نكون أحرارًا في النظر في خياراتنا دون أي تفكير لا مبرر له حتى نتمكن من التفكير مليًا في الاختيار، دون وجود استنتاج معين مسبقًا أو مفروض علينا، إحدى النتائج الهامة لهذا المفهوم هي أن قراراتنا من حيث المبدأ لا يجب أن تكون قابلة للتنبؤ بها، ولن نكون حقًا بصدد اتخاذ خيار حر.

الآن، قد تفترض أن أجهزة الكمبيوتر لا يمكن أن تكون حرة لأنها تختلف عنا من ناحيتين أساسيتين.

- **أولاً**: إنها تعمل وفقًا للهندسة المفهومة جيدًا، وهكذا يمكن دائما التنبؤ بها.
- **ثانيًا**: لا يمكنها ذلك حتى تنظر في الخيارات بنفس الشعور كما يفعل الناس.

المشكلة هي أن كلًّا من هذه التأكيدات مشكوك فيها، فلنبدأ بالتدقيق في مفهوم القدرة على التنبؤ، وأما عن أغراض هذه المناقشة سوف نفترض مثل معظم الناس أن هذا العالم المادي يعمل وفقًا لبعض القوانين الطبيعة، وسواء كنا نعرف أو لا نعرف ما هي تلك القوانين فهذا لا يعني أن كل شيء محدد سلفًا في الواقع، فقد تكون العشوائية في الواقع جزءًا أساسيًا من الطبيعة. لكن العشوائية هي فقط عشوائية، وليس تمريرة حرة للأشياء أن تحدث وفقًا لخطة أكبر، أو مبدأ يكون بطريقة ما خارج قوانين الطبيعة، وإلا فإن تلك الخطط ستكون ببساطة جزءًا من القوانين.

هناك طريقة أخرى لهذا للقول، وهي أنه في حالة معينة للعقل، ستكون هناك حالة متميزة من الدماغ لا يمكن أن تنشأ عنهما أفكار أو معتقدات مختلفة غير متوافقة مع ترتيب واحد من المادة والطاقة في منطقة الدماغ، لكن هذا لا يعني بالتأكيد أن هذه الافتراضات صحيحة في الواقع، فإن الكثير من الجدل التاريخي حول الإرادة الحرة يركز على هذه الافتراضات بدقة.

الآن، تخيل أننا وضعناك في غرفة على طريقة استجواب الشرطة، مع مرآة أحادية الاتجاه على الحائط، فقد أصبح بإمكان علماء المستقبل الأذكياء أن يلاحظوا كل شيء يخصك، بما في ذلك حالة وسلوك كل خلية عصبية في جهازك الدماغي، ثم طلب منك أن تقول بصوت مرتفع إما "أحمر" أو "أزرق"، ولكن قبل أن تفعل ذلك، فإننا نتحدى العلماء أن نتنبأ بما أنت ذاهب لاختياره، قبل تشغيل اختباراتهم، ونماذج المحاكاة، وأي شيء آخر يريدونه، يثبتون أنه يمكنهم توقع ذلك بشكل صحيح.

من هذا المنطلق، يعلنون بفخر أن ليس لديك الإرادة الحرة، بعد كل شيء، بغض النظر عن مدى صعوبة المحاولة، فلا يمكنك بذلك خداعهم، ولكنك تتوسل الاختلاف، وتطلب فرصة لإثبات أنك في الواقع لست عملة، وممكن التنبؤ بها، أولًا، تحاول أن تقرر ما ستقوم باختياره، فتقوم بتغيير رأيك لكن هذا لا ينفع، لأنه العلماء يستطيعون التنبؤ بأنك ستفعل ذلك، لكنك تحصل على فكرة هي أنه إذا جلست بهدوء شديد، يمكنك سماع العلماء يناقشون تنبؤاتهم، لذلك في المرة القادمة عندما يطلب منك اختيار لون، يمكنك الاستماع على المداولات ومعرفة ما تنبأ به، ثُم تقوم ببساطة بتحديد اختيار اللون الآخر،وأنت غير محبط من قبل ابتكارك، فهم يدمجون هذا في نماذجهم، وهذا لا يقتصر عليك فقط، ولكن أيضًا لديك إمكانية الوصول إلى التنبؤ الخاصة بهم قبل أن تفعل ذلك.

الآن، لا يوجد شيء في هذه الحجة لا يمكن بالنسبة لك تطبيقه جيدًا للآلة، فيمكننا بناء روبوت يفعل ذلك الذي فعلته بالضبط، بغض النظر عن كيفية برمجة هذا الروبوت لاتخاذ القرارات، وبغض النظر عن مدى إمكانية التنبؤ بأن الروبوت هو كما هو، طالما أن لديه إمكانية الوصول إلى توقعات خارجية لإجراءاتها الخاصة، لا يمكن أن يكون هذا

التوقع صحيحًا دائمًا، فيمكن للإنسان ببساطة الانتظار لتلك التوقعات، ثم القيام بعكس ذلك، لذلك فإنه قادر بامتياز بعدم التنبؤ بما يقوم به بالإنسان الآلي، حيث "يمكن أن يكون قادرًا حقًا" يعني أن لديها إمكانية الوصول إلى محاولة التنبؤ بما ستفعله.

هذا مثال على ما يسميه علماء الكمبيوتر مشكلة لا يمكن تحمّلُها، لا توجد خوارزمية فعالة يمكنها حل تلك المشكلة تمامًا، بمعنى أن تعطي الإجابة الصواب في جميع الحالات.

هل يمكنك كتابة برنامج A الذي سيدرس أي برنامج آخر B مع مدخلاته، ويخبرك ما إذا كان B سوف يتوقف في نهاية المطاف أم لا ؟ وبعبارة أخرى، يمكن معرفة ما إذا كان B أفضل من أي وقت مضى من إصدارات البرامج ؟

أظهر **تورينج** بأنه لا يمكن أن يوجد مثل هذا البرنامج، لكن، ما الذي يحدث بالفعل؟ البرنامج لا يرتكب الخطأ، لكنه قد أعطاك إجابة خاطئة، لكن ذلك ببساطة لا يتوقف أبدًا.

وفي حالة علمائنا المستقبليين، بغض النظر عن مدى ذكائهم في عملية التنبؤ الخاصة بهم، ففي بعض الحالات لن تصل أبدًا إلى نتيجة ما إذا كنت ذاهب لاختيار أحمر أو أزرق. هذا لا يعني أنك لن تصل إليه، فاختر إجابتك، فقط لأنهم لا يستطيعون دائمًا معرفة ذلك مسبقًا ماذا سوف تختار. قد يرتكب العلماء خطأ، مشيرًا إلى أنهم ليسوا مخطئين أبدًا، وهذا صحيح، لكنك تقابل أن الخطأ لا يكون أبداً هو الشيء نفسه لتوقع سلوكك بشكل موثوق به، لذلك، ليس الحال أن الآلة الحتمية، تلك التي تحمل سلوكًا محددًا ومفهومً تمامًا، يمكن دائما التنبؤ بها. فأي حالة معينة من برنامج كمبيوتر قد ينتقل إلى حالته القادمة بطريقة يمكن التنبؤ بها تمامًا، لكن من المدهش أننا لا نستطيع ببساطة أن نقرأ المعرفة بهذه الأمور المجموعة معًا للحصول على صورة كاملة عن البرنامج وما سوف يفعله في النهاية.

ونفس الشيء بالطبع ينطبق عليك على وجه الخصوص، فلا يمكنك أبدًا التنبؤ بدقة بسلوكك الخاص، من الممكن أن يكون هذا هو السبب في أننا نمتلك الحدس القوي بأن لدينا إرادة حرة، لكن هذا ببساطة أمر مثير للاهتمام لكنه ليس حقيقة مؤكدة،

فالاحتمالات الأخرى هي أننا لدينا إحساس ناشيء ذاتي من الإرادة الحرة لخدمة بعضنا بعضًا. حتى الآن، يتم تحديد الهدف والأغراض التطورية، مثل الرغبة في الحلويات أو الانجذاب إلى الجنس الآخر.

لننتقل الآن إلى السؤال عما يعنيه لك اتخاذ قرار من الإرادة الخاصة بك، فقط لأنك تستطيع ذلك الاختيار فلا يعني أن لديك إرادة حرة.

واحدة من أوضح وأكثر الانتقادات انتقائية للاعتماد على فرصة لتوفير مساحة للمناورة المطلوبة لشرح الإرادة الحرة من قِبلِ المفكر المعاصر **"سام هاريس"**، حيث يقول: إن الفكرة كلها هي أنه يمكنك إجراء اختيار مدروس ذي معنى بغض النظر عن الخارج أو قبله، ببساطة لا معنى له، يطلب منك أن تتخيل عالمين كلاهما هو بالضبط نفس الحق، حتى اتخاذ قرار من الفلسفة في الذكاء الاصطناعي فيما يخص إرادتك الحرة الخاصة بك، ثم تتباعد بحكم اختيارك، في واحدة قد اخترت اللون الأحمر وتختار باللون الأزرق، الآن، في أي معنى قمت عن قصد باختيار واحد بدلًا من آخر؟ كان تفكيرك هو نفسه بالضبط حتى تلك الدقة لحظة، ولكن بطريقة ما قد قمت بعمل اختيار مختلف، لكنك قد تصادف، لأنك قمت بتكوين عقلك.

"هاريس" سوف يقوم بالرد، لقد أدى شيء ما إلى قرارك، ويفترض أن تكون المداولات العقلية الداخلية، وإلا فإن قرارك قد تم تحديده ببساطة من خلال بعض العمليات التي اعتمد عليها، على الرغم من ذلك ربما يكون عشوائيًا، فلا يعيد أي شيء يشبه ما نعني بقصد التداول، ولكن هذا يعني أن "أحمر"، في حين قد اختفت العوالم "الزرقاء" بالفعل قبل أن تقرر؛ لذا دعنا نعيد خط البداية إلى البداية عندما بدأت التفكير حول المشكلة ربما يكون ذلك عندما تمارس الحرية، لكن عند هذه النقطة لم تكن قد قرّرت أي شيء على الإطلاق، لم تكن قد بدأت في التفكير في الأمر، حين انتهى **"هاريس"**، بشكل معقول بما فيه الكفاية، تلك الإرادة الحرة بمعنى المقصود بالاختيار، غير مقيد وغير محدد من قِبلِ الأحداث السابقة، فهو لا شيء بل أكثر من وهم.

الآن دعونا ننظر إلى مسألة كيفية اتخاذ أجهزة الكمبيوتر القرارات، بخلاف الناس، لدينا فكرة جيدة عن كيفية ذلك العمل، ومع ذلك، يمكنهم الاختيار دون الاعتماد على العشوائية، ويمكنهم أن يقدموا الأدلة، ويقوموا بتطبيق المعرفة والخبرة، واتخاذ القرارات، وتجنب المخاطر، وتعديل خططهم بناء على معلومات جديدة، ومراقبة نتائج أفعالهم الخاصة.

والسبب في حالة رمزية المعالجة، أو استخدام ما يمكن أن يسمى **"الحدس"**، على سبيل المثال، عن طريق استخدام التعلم الآلي لإعلامهم بالإجراءات في غياب أي فهم أعمق من السببية في العلاقات، ففكرة أننا يمكن أن الوصول إليها كانت عمليات إعادة التنظيم الداخلية الخاصة بنا مجرد أضغاث أحلام، لذا كان بإمكان الفلاسفة أن يفترضوا أنه قد يكون هناك شيء سحري أو غامض أو غير مادي حول عقليتنا فيما يخص تلك العمليات.

لكن علماء النفس التجريبي قد اكتشفوا أدلة جديدة ومزعجة تزعم أن أدمغتنا تستطيع اتخاذ القرارات قبل أن تدرك بالوعي، وتستطيع تنظيم ضغط دمنا دون تدخلنا، على سبيل المثال، في عام 2008 طلبت مجموعة من الباحثين اختبار المواضيع لاختيار بحرية ما إذا كان الضغط على زر مع يمين، باستخدام ماسح الدماغ بالرنين المغناطيسي الوظيفي، كانوا قادرين على التنبؤ بأي اليد التي ستستخدمها في الموضوعات حتى عشرة قبل ثوان من اتخاذ الموضوعات بوعي القرار، إذن ما الذي يقوله هذا عن الصندوق الذي نحتاج إليه "نحن" في مقابل العالم الخارجي؟ ونحن نتعلم أكثر وأكثر حول كيفية عمل أدمغتنا في مقابل عقولنا.

في الواقع يبدو أن عالمنا الذهني الخاص يتقلص في الخفاء؛ لذلك إذا لم يكن هناك شيء ما كالإرادة حرة، فلماذا ينبغي عليك ذلك؟ حاول أن تفعل أي شيء، على سبيل المثال، لإنقاص الوزن؟ **"سام هاريس"** يذهب إلى جعْل الملاحظة المثيرة للاهتمام التي قد لا يكون لديك أي خيار ذي مغزى لمعرفة ما إذا كان النظام الغذائي صحي أم لا، ولكن شيء واحد مؤكد هو أنك إذا لم تحاول، فلن تنجح. وبالتالي حتى إذا كانت الإرادة الحرة غير موجودة، فإنها لا تفلت منك للمحاولات التي تسير جنبًا إلى جنب مع القيام به في الواقع.

لتلخيص ذلك، ليس من الواضح ما إذا كان هذا يعني أن يكون لديك إرادة حرة فالكثير من الأشخاص الأذكياء يبدو معقولاً لهم أن حسك في الاختيار ليس أكثر من وهم، ويفترض أن عقلك ككائن مادي يلعب بنفس طريقة القواعد مثل بقية العالم المادي، وهكذا قد تكون خاضعة للتفتيش والتحليل، وإذا كان عقلك ينشأ من حسابك، على مستوى ما يجب أن تعمل أيضًا وفقًا لبعض قوانين الطبيعة، سواء كنا نفهم تلك القوانين بعد أم لا.

إن إدخال العشوائية إلى الصورة لا يُحدث هذه المشكلة، ولكن الحقيقة الغريبة أن الكثير من العمليات الحتمية لا تخضع للتنبؤ، حتى في فلسفة استخبارات العواصم، من حيث المبدأ.

أخيرًا، ليس هناك سبب آخر غير تمني التفكير في اقتراح أن الآلات في هذا الصدد مختلفة عنا. فهذا لا يعني أن الناس والآلات متساوية من جميع النواحي، فمن الواضح أنها ليست كذلك، لكن عندما يتعلق الأمر بالاختيارات، حتى الآن، على الأقل، ليس هناك ما هو جيد عن أسباب للاعتقاد أنها تعمل وفقًا لمختلف الطبيعية أو مبادئ علمية.

لقد تركنا النتيجة التالية: إما يمكن أن يكون لدى الأشخاص والحواسيب إرادة حرة، أو لا يستطيع أي منهما على الأقل حتى اكتشاف بعض الأدلة، على العكس، يأخذ الاختيار الخاص بك، فهل يمكن للكمبيوتر أن يكون واعيًا؟ كما هو الحال مع الإرادة الحرة.

"ديفينتاديس" يقول بأن الوعي مريب جدًا كلما بدا لنا أن نتعلم عن الدماغ، نجد العلم أكثر إشكالية من أن يصبح فكرة مجردة من الوعي، وبعض الباحثين يربطون بين الوعي والحالات العاطفية والتجسيد المادي، وآخرون لديهم أدلة وضعت أن حجب الاتصالات عبر أجزاء مختلفة من الدماغ سوف تتسبب في توقف الوعي.

تشير الدراسات التي أجريت على المرضى في الحالات الخضرية إلى أن الوعي ليس له شكل واضح ومحدد بالكامل، ولكن يمكن أن يكون في مكان ما بين أعصاب الدماغ، مما أدى إلى محدودية الوعي والقدرة على الاستجابة للأحداث الخارجية.

وقد طور **"أنطونيو داماسيو"** عالم الأعصاب المعرفي في جامعة جنوب كاليفورنيا نظرية داخلية يقترح في جزء منها أن الروابط واسعة عبر أدمغتنا، والهيئات هي أساس من الإحساس، أما **"ماديسون"** الذي يحمل الكرسي المتميز في العلوم العصبية في جامعة **"ويسكونسن"**، فيعتقد أن الوعي ينشأ من التكامل الواسع للمعلومات داخل الدماغ حتى نحصل على طريقة موضوعية للدفاع عن الإنسان واختباره الواعي بخلاف مجرد مراقبة الآخرين.

فهناك لا يوجد أساس عقلاني للاعتقاد بأن الناس لديهم وعي، ولكن الذكاء الاصطناعي في الآلات لا يمكن أن يكون كذلك، لكن من غير المبرر كذلك تأكيد ذلك بأن الآلات يمكن أن تكون واعية.

في الوقت الحالي لا يوجد طريقة موثوقة لتحديد ما إذا كانت أجهزة الكمبيوتر والحيوانات أو الناس الآخرين يمكن أن تكون واعية، لهذه المسألة لا يمكن تجربة الوعي بنفس الطريقة التي نشعر بها، وهذه مشكلة خطيرة، ومعظمنا يوافق على ذلك.

إن إيذاء أو قتل كائن واعٍ ضد إرادته أمر خطأ أخلاقيًا، ولكن ماذا لو لم يكن واعيًا؟ يمكنني بناء آلة تعترض بقوة على إيقاف تشغيلها، ولكنها تفعل ذلك خطأ؟

ومع ذلك مفهوم الوعي، هو تجربة ذاتية بشكل أكثر عمومية، ببساطة لا ينطبق على الآلات، فمن المحتمل أن الآلات، على الأقل، تتصرف كما لو كانت واعية، وترك لنا مع بعض الخيارات العنيفة حول العواقب.

وأطفالنا، الذين من المحتمل أن يتم ترقيتهم من قبل المريض، أو الأجهزة، أو الأجهزة الثاقبة، قد يجيبون بشكل جيد على هذا السؤال بشكل مختلف عما نحن عليه اليوم، هل يمكن أن يكون جهاز الكمبيوتر "يشعر"؟ ربما تكون قد لاحظت خيطًا مشتركًا حتى الآن: لطرح أسئلتنا يتوقف إلى حد كبير على ما إذا كنت تقوم بكلمات مثل الذكاء، والتفكير، والشعور كأنها تدل على شيء مقدس حول البشر أو على الأقل من الناحية البيولوجية.