

# 人工智慧是否「具有思想」？ 一個胡塞爾現象學式的分析\*

楊穎茜

Husserl-Archive, University of Cologne  
E-mail: yepyoung@gmail.com

## 摘要

人工智慧技術日趨成熟，擬真的互動和智慧讓機器人與真人的差異越來越模糊。人工智慧究竟是否和人類一樣，「具有思想」？在這篇文章中，我將透過胡塞爾的現象學來論證，人工智慧「能思考」，並擁有系統性的知識；但還不「具有思想」。因為「具有思想」，需要一套知覺經驗的動機網絡作為必要條件。如此的知覺經驗具有肉身性，並且對世界和他人開放。人工智慧和人類智慧的差異不在於高層次的語言計算能力，而是在於意識更底層的根本作用。意識有其自身持續的內在建構，此建構既是不停地在變動，也是自主地持續朝往「成為」的過程。這也代表，真正「具有思想」，蘊含的是具有肉身的生命。

**關鍵詞：**意識、知覺、認知、人工智慧、胡塞爾現象學

---

© 中央研究院歐美研究所

投稿日期：111.3.22；接受刊登日期：111.9.22；最後修訂日期：111.10.17

責任校對：蔡旻芳、林碧美、范馨文

\* 感謝匿名評審的細心閱讀以及提出的修改建議。

## 導言

1950 年圖靈 (Alan Turing) 提出模擬遊戲來測試程式機器人是否能和人一樣地思考。此測試在英美哲學領域中，引起相當多的討論。其中著名的為瑟爾 (John Searle) (1990) 的中文房測試論證 (Chinese room argument)；趙之振 (2020) 也在〈戴維森論圖靈測試〉一文中探討戴維森 (Donald Davidson) (1990) 對思想為何的看法。當年的圖靈測試 (Turing test)，不只是在英美哲學界，在起源於歐陸哲學傳統的當代現象學學者中，也產生了相當程度的討論。有別於英美學界瑟爾以及傑佛斯 (Hubert Dreyfus) (1980) 認為人工智慧並非具有思想的立場，胡塞爾 (Edmund Husserl) 現象學學者反而各別用不同的文本，來論證機器人「能思想」在胡塞爾意向性理論內部的可能性 (Holenstein, 1988; Mensch, 1991)。

在這篇文章中，我並非提出技術性的建議，如何讓人工智慧變得具有思想；而是著重探討，在胡塞爾哲學中「具有思想」的本質為何。胡塞爾將會承許，機器人「能思想」，它可以就外在輸出的結果來看，和人一樣計算、做數學演算、回答問題，甚至下棋贏過人類。在這個脈絡下，精確的邏輯符號和數學演算的能力，與人工智慧所模擬的函數計算，其實並無兩樣。我們甚至可以進一步說，當邏輯符號只是運算技術時，如此的邏輯學家和機器人似乎並無不同。但是，「能思想」和「具有思想」是不一樣的。我認為，胡塞爾不會進一步宣稱，只是輸入和輸出的機器程式內部真的「具有思想」。

我將以胡塞爾早期《邏輯研究》(*Logische Untersuchungen*, 1900-1901) 以及晚期《形式與超驗邏輯》(*Formale und transzendente Logik*, 1929) 與《經驗與判斷》(*Erfahrung und Urteil*, 1928-1934) 的文本為主要基礎，介紹他對「思想是甚麼」、「思想本質

為何」的看法。以此為基礎，我將論證，真正「具有思想」，要以活生生 (lebendig/living) 的「意識經驗」(Erfahrung/experience)<sup>1</sup> 作為基礎。真正「具有思想」，需要具備意識在自身知覺經驗中自主生成的能力。認知並非單純的輸入與輸出 (input-output)，而是一種非常廣義的理由關係 (when . . . , so . . .)。甚至在不需專注、未察覺且習以為常就可運用的日常知識中，更是已蘊含如此的動機 (Motivation) 關係；即使這個理由關係沒有被主體意識到或彰顯開來。

我們會進一步看到，胡塞爾現象學在哲學知識論中的特殊地位：在知覺中不需要概念的作用就已有認知能力。我將介紹胡塞爾意識理論中的知覺活動。對胡塞爾來說，我們的知覺經驗都已具有某種類型化 (typisiert/typified)，在意識內在自主建構的作用下，如此的知覺已經擁有認知和某種概念程度的思維。我們的身體、情感以及周遭世界的因素，也都參與我們知覺形成的建構過程。如此的認知是具有肉身性，並且在世界中展開。

在第壹節中，我將介紹胡塞爾現象學內部，支持機器人「能思想」的理論部分。人工智慧函數計算的能力屬於系統功能性的知識。但我認為，如此的知識尚未是真正的 (eigentlich) 思想 (Denken/thinking)，因為真正的思想需要知覺作為充實基礎，也就是作為判斷的證成。在第貳節中，我將在進入胡塞爾的知覺理論前，簡介人

---

<sup>1</sup> 在德文和英文中，empirisch (empirical) 和 Erfahrung (experience) 在日常中文都是譯為「經驗」。但兩者在現象學中有相當不同的知識論意涵。在這篇文章中，empirisch 會譯為「物理經驗」，指的是在物理因果 (kausal) 關係下所產生的感覺材料部分。為了有別於感官刺激的物理因果關係，Erfahrung 會譯為「意識經驗」或「意向性經驗」；在胡塞爾的術語中，類似概念的德文字為 Erlebnis，會譯為「體驗」。意識經驗代表的不再只是感官在物理因果刺激下所產生的感覺，而是具有意向性的內容，具有本質性的結構。意識經驗雖然具有感覺材料的部分，無法與此分割；但其主要的認知能力仍是來自意識意向性的作用，並非物理因果的刺激關係。在現象學知識論的進路下，意識經驗對判斷真假來說，是重要的基礎。

工智慧所模擬的生物神經知覺系統，以及辨識圖片所遭遇的困難。如此的討論，並非有意傳達生物知覺優於人工智慧，而是就「具有思想」在胡塞爾現象學中代表甚麼，來進行探討。在第參節中，我將討論，真正「具有思想」代表要擁有一套相應的知覺動機網絡。換句話說，知覺經驗是「具有思想」的必要條件。知覺並非來自空洞的我思，而是具有肉身性，並且在世界中展開。在胡塞爾的現象學中，知覺意識與生命的概念息息相關。知覺、身體、情感、世界性，在意識生活中，是相互開放影響並以動機關係相關連著的。在第肆節中，我會比較「機器意識」(machine consciousness) 和人類意識的差別，並且進一步討論，「具有思想」不僅要具有知覺，還蘊含具有生命。在第伍節中，我將試著討論，「具有思想」的條件，是否可透過由外在行為互動來規定。我將論證，在現象學的分析下，只是能外在地和機器人互動，並不就直接代表，此機器人「具有思想」。僅只是外在的互動，還不足以作為「具有思想」的充分條件。

## 壹、「能思想」：符號表象

有別於瑟爾和傑佛斯否定人工智慧有像人一樣思考的可能性，德國的現象學學者反而是支持此可能性。但首先，僅僅只是在高層次的形式符號演算。

荷倫斯坦 (Elmar Holenstein) 的論證是從符號表象 (symbolische Vorstellung) 的能力出發。他透過胡塞爾早期在《算術哲學》(*Philosophie der Arithmetik*, 1891) 中提出的符號學 (Semiotik) 論證，人的某些符號演算應用能力，可以和機器人無異 (Holenstein, 1988: 84)。

符號 (Zeichen) 對胡塞爾來說是一種非本真的表象 (uneigentliche Vorstellung)。本真的表象是由直觀 (Anschauung) 作為基

礎的，會有感覺內容作為直觀得以可能的實項內涵 (reeller Bestandteil)。符號是一種非本真表象，因為表象內容本身並非直接是具體的對象，而是提供指涉到被標誌者的功能。像是「小明」這個專有名字指涉到小明這個人，在此脈絡下，名字作為符號仍有它的表象內容 (Vorstellungsinhalt)。大部分的語言文字也是符號的一種，但文字符號仍具有或獨立或不獨立的内容 (Husserl, 1970: 344)。胡塞爾稱符號為一種「關係的概念」(341)。

在形式邏輯中，區分質料 (materielle) 的符號和形式 (formelle) 的符號格外重要。一個判斷像是「A 比 B 大」，A 和 B 屬於質料符號，有實指被判斷的對象，而「比……大」是一個形式的關係。在算數中，數學運算的符號是形式的。像是數學符號等於「=」和大於「>」、小於「<」所表達的，是前項與後項之間的關係 (Beziehung/relation) 是否成立；而運算符號加「+」或減「-」則是建立關係表象合併在一起的結果。純粹的數學符號和推論，胡塞爾譬喻為一種「形式的機械性」(formale Mechanism) (Husserl, 1970: 349)。

在形式符號的操作中，胡塞爾說我們可以製作出取代某些程序的「代理的符號」(Surrogat/surrogate)，讓我們的運算機制效率更高、更方便使用。如此在形式邏輯和算數裡的功能代理符號，可以在沒有相應直觀的基礎下進行。胡塞爾說，我們不追問其過程的邏輯證成；我們進行如此形式的符號運用，但不反思或主動根據合理的理由考慮和推論。這樣代理符號的運用過程，未必是邏輯合法的，而是彷彿「機械般的過程」：

在此我們不詢問符號推算的邏輯證成……我們在無反思的情況下跟隨符號推演，並且不根據前面的演繹或任何一個證成的

考量。這並非是邏輯規則合理的方法，而是自然機械般的過程。  
(Husserl, 1970: 361-362)

荷倫斯坦使用如此代理符號的應用來說明，人工智慧擺脫語意而進行單純形式和語法的操作，在胡塞爾的表象理論中是可能的。對荷倫斯坦來說，電腦的程式計算，儘管只是輸入和輸出的關係，以及透過規則演算，但也是一種知識的產出。他舉例說，我們一樣是寫一個程式指令到書裡，但書不會因為這樣的輸入而有相應的演算和輸出。他認為，知識並非只是人類或主體性獨有的能力；在模控學 (Kybernetik) 的發展下，我們對知識的範疇會有所轉變。人類的知識是認知性 (kognitiv) 的，而人工智慧的知識，是屬於功能性的 (funktionell) (Holenstein, 1988: 111)。既然人工智慧的系統程式也能發展出自主的範疇 (autonome Kategorie)，那麼人工智慧當然可以稱為能思考，只是是「系統功能性的知識」。

門斯 (James Mensch) (1991: 109) 也提出，能計算和能對計算給出理由——也就是「能思考」和「具有思考」——是兩種可以區分的能力。我們可以透過《邏輯研究》的文本，更進一步地闡釋這個想法。在《邏輯研究》中，胡塞爾批評當時的心理主義，試圖用心理歸納的因果關係去解釋邏輯得以可能的思想能力。胡塞爾其中的一個批評在於，歸納性、現實性 (Reale) 的因果關係並無法取代觀念性的 (Ideale) 思想 (Husserl, 1975: 78)，因為歸納性的因果關係無法給予思想所包含的理由 (Grund)。歸納性的因果關係就猶如計算機器，它可以計算，但無法給予理由基礎 (75)。我們可以進一步地說，計算機程式語言的計算能力，並無法作為「具有思想」的條件。「具有思想」代表，要能擁有理由內容。

計算機本身，就輸入與輸出的結果來看，的確是會計算、能夠處理符號的。但程式模擬的因果關係，並非包含使思維得以可能的

證成以及理由給予。邏輯作為一門科學，是關於知識論的科學，目的在於解釋思想得以可能的條件。邏輯知識論要做的，不是發展形式符號的推論，而是去解釋推論得以可能的基礎。對胡塞爾來說，只掌握了邏輯規則的邏輯學家，只是擁有操作形式邏輯符號的技術，並未真的在進行邏輯這門學問。邏輯學的核心在於，要解釋思想得以可能的條件。胡塞爾寫道：「純粹邏輯是關於可能性之觀念性條件的科學，也就是關於普遍性的科學」(Husserl, 1975: 262)。

由此可知，胡塞爾雖然認同單純操作符號和應用規則的功能性知識，但知識論的目的在於，解釋思想得以可能的條件、說明思維內含的理由給予關係。單單只是會操演和應用是不夠的。即使是代理符號的表象 (Surrogatvorstellung)，胡塞爾也強調說，即使我們可以忽略知識證成 (Rechtfertigung) 的部分，我們同時仍保有這個能力：隨時可以回到充實符號本身真正所指的內容上。「我們在此所處理的代理符號表象之種類可以如此刻畫：符號所代理的真正的表象，是在每個時刻都可以任我們取用安排的」(Husserl, 1975: 354)。

雖然沒有相應直觀充實的思想是可能的，但它不是真正的思想。真正的思想是主體要能做出相應能被充實的直觀；也就是說，主體要能回到此知識判斷得以被充實和證成的原初經驗之中。<sup>2</sup> 我們可以說，直觀被充實的過程——也就是思維判斷在原初經驗中被證成的過程——是思維主體真正「理解」的過程。

算術知識的邏輯判斷，對胡塞爾來說是絕對的。像是「 $5+2=7$ 」這樣的邏輯判斷，不論對誰來說、不管是在哪個地方，都是等同的真理，「不管是由人類、非人類、天使或上帝來判斷」(Husserl, 1975: 125)。數學和邏輯知識，是原本就在那的，我們只是去「發現」

---

<sup>2</sup> 關於直觀的部分，會在第參節第一部分「知覺認知：部分與整體」，以範疇直觀為例再做說明。

(entdecken/discover) 它 (1972: 312)。如此的說法，讓胡塞爾時常被誤認為是柏拉圖主義。但即使是如此絕對的真理，胡塞爾都有提出相應的意向性建構，去說明這個判斷如何在直觀中和明證性下被充實。一個為真的判斷，一定有相應對象自身被給予的直觀和原初意識經驗。對象不是與我意識無關的獨立存在，而是有在意識經驗中如其自身地給予我的過程，也就是直觀經驗被充實的過程。擁有如此能夠充實的意識經驗，才能說是真正地「具有思想」。

總結來說，人工智慧在符號邏輯的形式層次，是「能思想」的，有著「系統功能性的知識」。系統可以處理資訊、使用規則、會回答、「知道」怎麼做 (know-how)，即使完全不涉及規則背後的道理以及證成的理由。但如此的知識並非是真正「具有思想」，因為人工智慧仍不擁有真正的意向性直觀和活生生的原初意識經驗。

人類智慧有別於人工智慧的，不只是意識高層次的理由證成和意向性直觀的充實，更是在意識根本底層的作用，像是時間意識建構的面向。知識的綜合經驗有其時間的結構 (Mensch, 1991: 117)。在胡塞爾的意識理論中，被動綜合是意向性在意識流中持續的內在建構；這是人類認知作用中相當有別於人工智慧的部分 (Holenstein, 1988: 84)。如此意識的內在建構使得過去的認知成果不會消失，而是保留在意識生活 (Leben/life) 之中，成為一種認知的習性；習性同時影響著當下也引導著未來。習性認知打開的架構是使意義得以可能的背景 (background)；人工智慧常被指稱，缺乏背景架構和常識 (Dreyfus, 2005; Minsky, 1980)。

在胡塞爾的理論中，如此日常知識的背景和習性在意識流中持續的內在建構有關；而這是知覺中相當基本的架構。在晚期發生現象學發展中的超驗邏輯，特別重視知覺的內在建構。邏輯概念都要能夠回到原初經驗，才能充分闡明其意義。知覺是意識經驗最原初

的一個形式，因而知覺，以及其內在自主的建構，對判斷的闡明也是扮演著不可或缺的角色。

因而，儘管在符號表象層次——也就是可以在沒有直觀充實下運作的層次——的推論，人的推論可以宛如機器人一樣。但胡塞爾意識理論重要之處在於，意識對象在原初經驗中的自身給予。胡塞爾的意識分析之主要目的之一，是在不預設立場的情況下，探討對象自身給予我們的原初經驗結構為何。對意識分析的目的在於，回到實事本身。因而我們可以看到胡塞爾對不同的意向性體驗 (Erlebnis) 的分析，諸如：通達經驗對象的知覺經驗、通達虛構對象的想像經驗、通達他者的陌異經驗等。此外，在現象學的發展下，意識經驗的結構會呼應主體之意識生活的結構（也就是視域) (Horizont/horizon)。在胡塞爾對知覺經驗結構的分析中，我們尤其可見這個特色：意識經驗一方面是透過相應的直觀充實，讓對象如其自身地被給予；另一方面，是考慮知識在意識生活中的沉澱與再發生，作為之後認識對象的認知習性與預期。<sup>3</sup>

## 貳、人工智慧中的知覺系統

人工智慧能夠在高層次的形式邏輯和函數程式進行計算，但真正「具有思考」，對胡塞爾來說，要擁有能夠徹底闡明其意義的原初經驗，即知覺。在進入胡塞爾的知覺理論之前，在此先簡略說明人工智慧中知覺認知的困難。

早期人工智慧追求抽象符號和語言系統的發展，隨著人工神經網絡 (artificial neural networks) 的發展，人工智慧也開始模擬生物

---

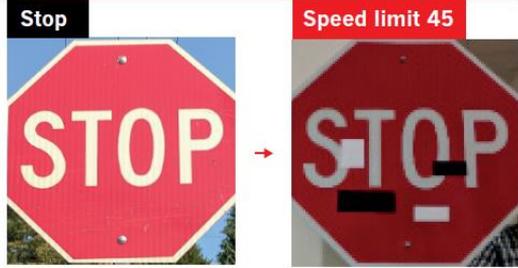
<sup>3</sup> 關於意識經驗這兩個面向，會在第參節第二部分「類型化的知覺：意識的底層建構」中繼續闡明。

的知覺系統。我們可以看到，人工智慧的發展也從抽象符號，步入對人類感官系統的模擬。人工神經網絡是一種模仿生物神經系統功能的數學計算模型。這類技術被廣泛運用在圖像辨識上 (Butz & Kutter, 2017: 60)。所謂的深度學習 (deep learning)，是具有多層的神經網絡計算模型，可以對資料進行「學習」。

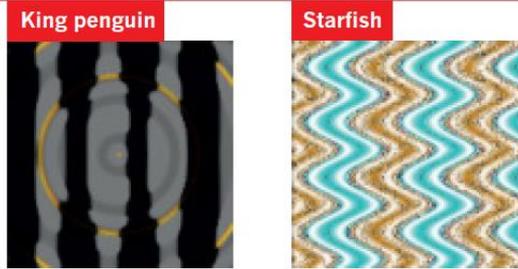
但人工智慧仍然無法輕易成功辨識視覺上的整體。如何將眾不同的特徵掌握為一個整體，這樣的能力仍是人工智慧所缺乏的 (Butz & Kutter, 2017: 57)。在透過深度學習訓練人工智慧辨識圖片時，會輸入大量的資料去進行函數計算。像是人工智慧學習了「停止」的交通符號，但如果在學習過的圖片上，在關鍵處加上一些方塊矩陣，一樣的照片會被辨識成完全另一種意思的「限速 45」。或者，某些凌亂的抽象圖紋會被解讀為具體的對象，像是企鵝，或是海星 (圖 1) (Heaven, 2019)。被訓練過的人工智慧仍很容易被欺騙 (Nguyen et al., 2014; Szegedy et al., 2014; Yosinski et al., 2015)。相較於人工智慧的圖片辨識能力，人類生物性的知覺能力比人工智慧要來得更穩定，同時也更靈活 (Newen et al., 2020)：一方面，即使狗的圖案在不同圖片上的形狀和樣貌都不同，我們仍可毫不費力地辨別出是狗的類型；另一方面，又能將習得的類型多變地運用在其他不同的特例上。即使是孩童，一旦認識了貓的類型，就可以運用在不同大小、不同顏色的貓身上，甚至運用在卡通上。卡通圖案的抽象以及與現實相離的誇張變化，也不影響孩童辨識出狗的類型。除此之外，孩童的知覺能力也可以很快地辨別出狗和貓的不同。

像是維根斯坦 (Ludwig Wittgenstein) 提出的鴨兔圖 (rabbit-duck illusion)，人類可以自由地轉換不同的類型運用在這張圖上，我們可以一會兒看成鴨，一會兒看為兔子。看成鴨的知覺中，細長的柱狀圖形會被知覺為是嘴巴；在看成兔子的知覺中，同樣的線條

These stickers made an artificial-intelligence system read this stop sign as 'speed limit 45'.



Scientists have evolved images that look like abstract patterns — but which DNNs see as familiar objects.



資料來源：Heaven (2019)。

圖 1 造成人工智慧誤讀的圖例

則是被詮釋為耳朵。在不同知覺中，一樣的客觀圖形會被掌握為不同的知覺內容，之間相關性的組合排列也會不同。如何在眾多信息中，匯集重要相關的部分掌握為一個整體，尚是人工智能缺乏的能力 (Butz & Kutter, 2017: 57)。

知覺如何從其自身辨別整體與部分，其實已經牽涉到很根本的知識論議題。其中核心的哲學問題在於：我們如何在雜多流變的經驗中認識到同一的對象、我們如何在雜多中決定出對象單一的意義。除此之外，知覺認知也涉及到在當下雖然沒有彰顯、卻又對意

義決定有著重要影響的背景因素，像是社群中共享的常識、特定情境下的因果關係。

這些日常生活中的認知問題，我們人類主體不須主動進行判斷，有時甚至在絲毫未察覺的狀況下，透過知覺自身就可以完成。從胡塞爾現象學的角度來說，知覺中的認知在不需要語言概念的協助下，而是在「類型」(Typus/types) 的協助下，即可區分部分與整體。接下來我將會說明知覺中的部分與整體。除此之外，我也會說明知覺的身體性和世界性，以及之間的動機關係。

## 參、胡塞爾晚期發生現象學中的知覺理論

### 一、知覺認知：部分與整體

胡塞爾在一九二〇年代晚期，將邏輯範疇的起源訴諸於更底層的知覺經驗。相較於早期在《邏輯研究》中，胡塞爾著重在語言、判斷和符號等較高意識層級的充實關係，在《被動綜合分析》(*Analysen zur passiven Synthesis*) 的課堂講稿、《形式與超驗邏輯》和《經驗與判斷》，胡塞爾往更底層的知覺經驗中去探討、說明思想的起源以及知識得以可能的條件。

知覺作為邏輯概念的必要條件，理由在於，如果純粹只從形式邏輯的語法，我們並無法判斷真假。而是要回到認知對象自身給予的經驗之中，才能真正闡明它的意義 (Husserl, 1974: 213)。像是「天空是鹹的」這句話，符合「S is p」的語法句型，但是這是沒有意義的一句話，因為範疇錯置；我們甚至無法論及這句話的真假。要透過回到對天空這個概念的原初經驗，才能判斷此這句話是不是有意義的。在對天空的知覺經驗中，我們會知道，天空可以用顏色來形容，但不是透過味覺來認識。以對象原初被給予的知覺經驗為

基礎，才能排除無意義但符合形式規則的句子。<sup>4</sup> 邏輯學因而需要一門「意識經驗理論」(Erfahrungstheorie / theory of experience) 作為意義來源以及真假值的基礎 (Lohmar, 2000)，<sup>5</sup> 僅是形式邏輯是不足夠的。

知覺對胡塞爾來說是意識經驗的最原初狀態。在發生現象學的開展下，知覺的意識作用不僅只是當下對象的原初呈顯，更是包含了過往沉澱的習性認知與朝向未來的預期作用。習性和預期所擁有的時間性面向，以及其在意識流建構中的沉澱和觸發，在當下的知覺認知中都扮演關鍵的角色。由此可知，在胡塞爾現象學理論內部，知覺經驗對知識的成立來說，是必要的條件。

對康德來說，知識的最小單位是判斷。但對胡塞爾而言，在知覺中將部分歸屬於整體的過程，就可以是最小的認知單位。如此的

---

<sup>4</sup> 在此所說的知識屬於經驗性知識。形式知識，像是數學知識，有另外的直觀作用。而本篇是首先著重在經驗性知識。

<sup>5</sup> 在現象學中，意識經驗對我們判斷真假來說，是重要的基礎。意識經驗是對象顯現的場所，是邏輯知識論研究重要的場域。邏輯奠基是屬於知識論的領域，更是屬於關於思維和知識體驗 (Denkund Erkenntniserlebnisse) 的純粹現象學 (Husserl, 1984: 6)。如此的體驗並非是物理性的或實事性的 (faktisch)，而是在本質直觀中被把握和分析的本質普遍性，是先驗的 (apriorisch)。研究意識經驗背後的目的，是為了進行知識批判和建立純粹的邏輯學。以下這段引文可以讓我們看清楚意識經驗 (或說體驗) 和知識邏輯學的關係：

屬於知識客觀理論，以及，與此緊密相扣的，一門思維和知識體驗的純粹現象學之廣泛範圍的闡釋。此關於思維和知識體驗的純粹現象學，如同包含其的體驗之純粹現象學一般，唯獨與能在直觀中以其純粹本質普遍性被掌握和被分析的體驗有關，而非與物理經驗性的、作為現實事實性的統覺體驗有關，亦非與活著的人類或動物、其作為在顯現和在經驗事實中所設定出的世界之體驗有關。這在本質直觀中直接被掌握的本質以及在此本質中被建立的相關性，將這門體驗現象學描述性地表現在本質概念與法則性的本質陳述中。每個如此的陳述在其最佳的表達中都是先驗的。這個領域，會在為知識批判所準備的、在和純粹邏輯所闡明的目的下被探究。我們的研究因此會在這個領域中被推動。(6)

知覺活動，胡塞爾稱為闡明 (Explication/explication)，是邏輯範疇——「是」(ist/is)——的起源處 (Husserl, 1972: 127)。透過認知主體在注意力之下進行闡明，我們可以在知覺中認識一個新的對象，並尚未涉及到概念的使用。這是範疇「是」的最原始型態。我們可以在《經驗與判斷》找到關於此的詳細說明。

認知是一種意識活動，所認識到的意義是透過意向性建構而成。有別於人工智慧裡模仿人工神經元的數學模型，胡塞爾透過意識的意向性活動 (Akt/act) 去說明認識的過程。當我們初次認識某個具體對象時，我們首先會將它知覺為一個整體 (A)；但此個體的內部環節尚未展開。我們會進一步去認識它不同的特徵，像是大小 (i)、顏色 (j)、功用 (k) 等等。當我們受到興趣的驅力推動、想要去認識這個對象時 (Husserl, 1972: 124)，認知活動才會開始；我們想要去充實內部的細節。否則我們只會把它維持在一個模糊的整體而已。興趣的部分在這個認知過程佔有重要的角色，作為促發的動機。

當我們開始進行認識時，我們會把各種特徵 i、j、k 歸屬到 A 上，胡塞爾稱此為一種「重合」(Deckung) 的活動；這是一種精神活動的轉移，從部分特徵轉到作為整體的 A 上。在這個過程中，不同的特徵不再是各自獨立疏離的個體，而是作為 A 內部的環節部分，成為規定 A 的內容。如此的部分特徵重合到整體上，這個認知活動，胡塞爾稱為「重合同一性的綜合」(Synthese der Identitätsdeckung) (Husserl, 1972: 129)。在這整個過程中，我的興趣和注意力仍是維持在 A 上，而不會在顧及個別的 i、j、k 時就跑散了。這樣我才是「認識 A」(126)，而不是在認識各別特徵 i、j 或 k。展開 A 的各別環節同時也是在豐富 A 的對象意義。

如此的重合綜合 (Deckungssynthese) 是認知的意識活動，是知識的最小單位，也是邏輯範疇「是」的最初起源。《經驗與判斷》

在這部分對闡明的說明，呼應胡塞爾在《邏輯研究》所提出的範疇直觀 (kategoriale Anschauung)。範疇作為認知的對象，不屬於感覺經驗性的對象；後者可以在簡單的知覺中直接自身給予、充實直觀。範疇要在重合綜合如此複合性的意向性活動中，才能夠自身給予。闡明的作用，就是使直觀活動的充實得以可能。當我的注意力維持在 A 上，將 i、j、k 的部分都作為豐富 A 之內容的環節，那麼，將內容歸屬到認知對象主體上的歸屬過程中，範疇「是」便在原初的意識經驗中自身給予、也就是在明證性的經驗中獲得了充實。

從上面的展示我們可以看到，我們不需要語言概念的介入，去統整雜多的感覺材料；而是對象會在經驗中作為整體自身給予。我們也不需要邏輯判斷形式，像是將 A 抽象化為基底 (substrate) 並給予述謂詞 A is i, j and k；而是透過部分和整體的歸屬關係，我們便可在知覺闡明中，認識 A，並將 A 視為擁有 i、j、k 之部分環節的整體。如此範疇「是」便得以在闡明活動的意識經驗中自身給予。如此便是在胡塞爾現象學中，範疇直觀在意識經驗中獲得充實的結構。這也就是意識之認知活動最簡單的模態之一。範疇直觀在意識經驗中的**充實**，也就是認知主體**認識**了，i、j、k 是 A 的內容。

人工智慧特殊之處在於，它相當擅長使用高層次的抽象概念與符號，以及邏輯嚴謹的算術推演。但真正地能夠運用概念，代表要回到對象原初被給予的意識經驗之中；在認知主體注意力的配合之下，使直觀得以在原初經驗中被充實——也就是**認識**。人在底層的實際知覺經驗之中，就已擁有掌握對象之整體和區分部分的能力。感官知覺內部，就已有其自身的秩序與意識的建構法則。相較於人工智慧，人類認知能力的特殊之處，更展現在知覺活動之中。因為我們在知覺中並非是透過感官點對點的物理刺激去認識對象，更非

透過智性的方式；而是，知覺本身的意識結構，就使得對象可以在背景的襯托下而呈顯出主要的內容，在「雜多」中自身呈顯出「一」。

## 二、類型化的知覺：意識的底層建構

胡塞爾在晚期發生現象學中，提出了「類型」的概念 (Husserl, 1972: 136, 398-401)，來做為部分與整體的內在連結。發生現象學會考慮意識流中習性的內在建構，以及不同意向性活動之間的動機連結；沒有一個意識活動是單獨存在的。類型理論表示，知覺總是已經類型化了的。我們並非單純看到綠綠的、一片片的東西，而是在知覺中直接看到樹上的葉子。而類型總是已經引導著我們，在日常認知中去解讀新的知覺內容。

一旦我們第一次透過闡明認識過某個對象後，我們會在意識流中建立 (stiften/building) 一個認知成果。這樣的建立可以透過主體主動記憶，但在日常實踐中，這樣的建立往往是被動的，在主體不費力的情況下就會發生。在意識底層持續進行著的內在建構，是被動發生的綜合活動。我們可以從胡塞爾的發生現象學中，分析出意識流內在兩個方向的建構，即：沉澱和再發生。過往的認識會沉澱在意識生活的歷史之中，並在之後每個新的當下，可以透過觸發 (affection) 而再次被喚醒和運用。

知識因此有它沉澱和被喚醒的歷史：一個透過與世界互動的經驗中獲得的習性 (habitus)。我們因而總是在類型的引導下，去和世界中的對象打交道。知覺擁有透過類型勾勒出一個視域 (Horizont) 結構，我們會已有或多或少來自過去的認知和預期。像是我看到狗，我預期牠可能的動作。如果我是怕狗的人，我對狗的知覺還包含著一些畏懼的情感。

在《被動綜合分析》手稿中，胡塞爾非常詳細地分析了認知習性中觸發產生的條件，他提及，除了相似性，和觸發最相關的部分，其實是在於興趣、情感、驅力等等：「最起作用的動機像是廣義常見的『興趣』，原初或是已習得的情感評價、本能或較高層次的驅力」(Husserl, 1966: 178)。相關於觸發的字詞，胡塞爾也會使用「刺激」(Reiz/stimulus)。但觸發的法則，並非是物理經驗性的 (empirisch/empirical)，而是一種意識本質的 (eidetisch)、先驗的 (apriorisch)、合法則性 (Gesetzmäßigkeit/lawfulness)。我們的肉身遇到知覺對象，透過感官獲得感覺，這並非由物理刺激造成。肉身並非如一般軀體在因果制約下產生感覺，而是，身體本身也參與在觸發的建構之中。胡塞爾稱肉身為一種能力 (Vermögen/ability)，代表著「我能」(ich kann / I can)，是一種非常底層、在自我注意力都尚未發生，就已經有的原初主動性。身體性的我能，因而和觸發有相當緊密的關係；因為我們透過身體獲得的感覺，都已經是活生生的意識活動，而非僅是物理因果造成。<sup>6</sup>

關於類型的想法，似乎人工智慧透過大數據和深度學習也可以模擬出來。只要樣本和資料夠多，可以找出其中的樣式 (pattern)，將類型內的動機關係參數化，像是模擬出聽到狗吠叫聲會受驚嚇、身體會往後縮的因果關係。或是透過社會常見的文化規範，也可模擬出甚麼脈絡下，如何的情感反應是恰當的。但我認為，之間仍存在著差異。根本的差距在於，意識仍是透由動機所連結的內在建構自主地運作著；而人工智慧則是以第三人稱外部的角度去做資料類

---

<sup>6</sup> 這並非代表，胡塞爾會反對自然科學態度下對人體因果關係的探索。在《觀念二》(Ideen II) 中，胡塞爾區分自然態度 (naturalistische Einstellung / naturalistic attitude) 和人格態度 (personalistische Einstellung / personalistic attitude)。身體在物理因果法則制約下的表現，會包含在前者；而身體「我能」的面向，則會屬於後者。

比，沒有意識內在自發的聯結綜合。即使人工智慧可以模擬出某人知覺習性和動機中的所有類型，模擬出高度人格化，也仍是缺乏內在的動機關係、觸發關係。如此意識建構的內在性 (Immanenz)，是本質性的，而非在內部 (innere) 和外部 (äußere) 的封閉區分。在胡塞爾現象學中，內在性是相對於超越性 (Transzendenz)，是所有意向性對象得以可能顯像的基礎。相對於人工智慧程式，在輸入輸出面向上的平等關係，對胡塞爾的意識哲學來說，意識反而是最原初的，是使所有認知與表象 (Vorstellung) 得以可能的基礎。

意識經驗有著這樣兩個面向：一方面，知覺經驗對世界敞開，世界對象在經驗中會有原初的自身給予。另一方面，意識生活內在在建構所獲得的習性，不管是認知的習性抑或情感的習性，也都會參與其中；每個知覺經驗中所打開的視域結構，也是相應於意識體驗流的內在建構、類型化形塑而成。而這個知覺背景 (視域) 如何展開，也與觸發的結構有關。因而意識經驗本身的形成，有很大一方面，與意識的本質內在建構、人格建構有關。

### 三、知覺、身體和世界：動機關係

在知覺中的動機關係並非是高度主動反思或推理的活動，而是意識內在自主觸發 (affiziert) 的關係。<sup>7</sup> 即使認知主體沒有察覺到或多花注意力，自發的意識活動也包含這些動機促發 (motiviert)。<sup>8</sup> 底層的意識活動並非主動出於自我的注意力，但自我作為此意識經驗的承受者，仍是先天本質結構上、第一人稱的可覺察者 (即使實際上自我沒有真的當下意識到)。任何意識經驗，都蘊含著前反思的第

<sup>7</sup> 胡塞爾在《觀念二》中區分理性動機 (Vernunftmotivation / motivation of reason) 以及聯結動機 (assozielle Motivation / motivation of association) (Husserl, 1952: 222)。前者是高層次要符合邏輯反思的推論，而後者是透過觸發聯結。

<sup>8</sup> 關於底層的動機促發，會在第五節中繼續深入說明。

一人稱覺察。對胡塞爾來說，動機關係是意識生活的法則；即使在主體未察覺的意識底層，如此的內在發生也持續作用著。即使是底層的意識習性歷史的建構，以及習性觸發後的再發生，都是活生生的意向性關係。

如此的知覺並非僅是腦部的認知，而是具肉身性、在世界之中開放的。知覺內容會和「價值情感」(Wertgefühl / feeling of value)、「身體運動感」(Kinästhesie/kinesthetic) 構成一個動機網絡。像是我看到我要搭的公車駛近，我會招手，因為我預期招手車會停，不然它可能會開走；我招手的動作，已經是透過這個社會規範認知所引導。我們的日常行為實踐，已深深受預期和習性所影響。觸發的機制，在「知覺內容－情感－運動感」這三者的相互牽引中發生。這三者在一個動態且自主的動機相關關係 (Motivationszusammenhang / relation of motivation) 中產生作用。這一組類型化的動機關係，在主體不費力或不知情的情況下，已深深地為當下新的對象形塑它的視域背景。

我們可以用更多例子來說明類型中的動機關係。像是在德國我第一次認識到，紅燈會先轉成黃燈，才變成綠燈。這和臺灣直接從紅燈轉成綠燈不一樣。在我看到對面方向綠燈已經結束時，身體已經有起動的傾向 (Tendenz/tendency)，等著紅燈變綠燈，但突然的黃燈打斷了預期以及身體的節奏。隨著幾次修正的經驗，身體開始會預期，紅燈之後還會有黃燈。我對「紅綠燈」這個對象，因而有一個類型認識的歷史。它與我的過往認知、身體習性、我要趕快過馬路的需求 (否則遲到不好) 相關連。這些不同因素，都會一起建構在類型化的認知中，作為促發的動機、產生作用。像是晚上從公園回家，黑暗中草叢突然地傳來一陣沙沙聲，以為有東西會跑出來，如此的念頭 (預期) 讓我產生驚嚇、身體會突然往後縮。而習性上

害怕黑暗處的人，對沙沙聲的掌握會比其他不怕黑的人還敏感，是因為人格觸發的結構不同。我們因而可以看到，身體習性與意識底層的被動建構之間的關係。意識底層在建構不只是認知性的、情感性的、人格性的，也是身體性的。由於意識經驗對世界的開放性，肉身性的建構因而也會透過社群經驗具有文化和社群習性。肉身性的建構因而也是習性式的。

主體的身體不只是定位在世界的這裡和現在，身體運動感和對象的顯像也是一種動機關係，胡塞爾稱為「運動感促發」(kinästhetische Motivationen) (Husserl, 1966: 264)。身體運動感的歷程，並非單單只關係到認知中的對象顯像，而更是一種「本能的、驅力的歷程」(2006: 328)。「每個運動感領域自身都是本能的相關關係」(328)。

對胡塞爾來說，世界和意識是在互相建構的雙向關係中形成。知覺在動機之動態脈絡中，是對世界開放的。隨著在不同周遭世界中獲得的經驗，對象的意義內容也會有所不同。胡塞爾說，對溫度類型的建立，和地理環境有關。像是冷的類型，對溫帶地區和熱帶地區的人來說，是不一樣的；對熱帶地區來說的冷，在溫帶地區可能只是涼快。再一個例子，火車在現代是相對於飛機較慢的交通工具，但在以前是最快的。交通快和慢的類型也隨著社群、技術、文化有所不同 (Husserl, 1972: 230)。「火車是最快的交通工具」，隨著飛機的發展，如此的判斷已經不再為真。

世界一方面塑造我的知覺內容；另一方面，我每一次新的認識和興趣喜好，也同時塑造著生活世界。生活世界也是透過經驗在被動建構之中，隨著我身體的習慣、需求、喜惡而被類型化的背景視域。胡塞爾說：「具體的周遭世界之後是充滿意義、有趣的；根據人格的興趣被分類並類型化」(Husserl, 1973: 56)。隨著新的經驗、

預期被破壞，這個類型化的動機網絡可以被豐富意義、被改變、修正或強化，一直隨著主體和世界環境的變動而適應和調整。這個動機網絡在熟悉的日常經驗中，雖然沒有被彰顯，卻是使認知（也就是在雜多中統一意義）得以可能的背景。

知覺的動機網絡與身體運動感相關連，並在對世界開放的關係之中。一個思維判斷要能夠成立，要回到如此活生生的原初經驗之中。而原初知覺經驗在發生現象學的考察下，與意識生活中的認知習性歷史、觸發情感等有關。如此一套活生生的知覺動機網絡，對「具有思想」來說是必要的。若人工智慧不具有如此第一人稱動機觸發的知覺經驗，那麼我們就無法說它真正「具有思想」。相較於人類高層次之自我意識以及反思存有的心靈能力，我認為底層的知覺活動的結構和意識自發性的內在建構（也就是意識生活本身），更能突顯人類心靈和人工智慧的差別。

#### 四、回應傑佛斯對胡塞爾「背景」概念的批評

現象學學者傑佛斯強調，人工智慧在常識 (commonsense knowledge) 上顯得相當不足 (Dreyfus, 1972, 1980)。人工智慧可以回應艱難的理論問題，但卻無法判斷簡單的日常情境語意。他認為胡塞爾的現象學不如海德格 (Martin Heidegger) 和梅洛－龐蒂 (Maurice Merleau-Ponty)，足以對日常知識提供清楚的說明，原因之一在於背景 (視域／世界) 概念的差異 (Dreyfus, 2005: 48, 2012: 1, 3-4)。

傑佛斯解釋：雖然胡塞爾會說背景是沒有被察覺到的 (unnoticed)、未彰顯的 (implicit)、隱藏的 (concealed)，但事後現象學家仍可以反省和對象化它，彷彿日常經驗僅僅只是由主體的心靈狀態所構成 (Dreyfus, 2012: 3)。除此之外，胡塞爾仍把世界當成對象，

而非真正使對象得以可能的背景；在他的理論中無法說出一個未被規定、主體沉浸其中的背景，一個整體性的氛圍 (holistic atmosphere) (4)。以下我想回應第一部分的批評：日常經驗 (世界背景) 僅是由主體心靈狀態所構成。

對胡塞爾來說，世界並非只是主體意識的心靈狀態。如上面曾提及的，環境因素會決定認知的內容，知覺的內容會透過意識經驗受地理、文化環境、交互社群形塑。胡塞爾在《經驗與判斷》討論到，文化客體的觀念 (Idealität/ideality)，是和環境領域 (Territorien/territory) 相連繫 (gebunden) 在一起的，是和環境「相連繫著的觀念」(gebundene Idealitäten) (Husserl, 1972: 321)。

傑佛斯引述心理學家吉布森 (James Gibson) 所提出的「環境示能」(affordance) 的概念，去說明知覺中的給予是非概念性的，並且環境提供「有意義內容的」(meaningful) 的指示，會「要求」我 (Dreyfus, 2005: 55)。像是「食物會要求進食、門會要求進出、地板會要求走路」(56)。主體對這些環境要求的回應，並不需要主動地特意察覺，或用注意力去回覆。

如此來自對象的要求，胡塞爾也有類似說法。他認為，當我在懷疑，眼前的人像是真人，還是玩偶時，對象如其自身所是的顯像，則是一種對我要求的關係。在懷疑中，兩種知覺都是有可能的，都有提供足夠的證據，去「促使」我、「要求」(fordert) 我，去做出相應證據合理的知覺把握 (Husserl, 1966: 34)。胡塞爾說，選擇「要求」這樣的表達不是沒有原因的，「這個表達指出，我們知覺的把握中，都存在有**傾向**，如此的傾向是在相互網絡中被促發；並且，在此動機促發中，**傾向**擁有它的『力量』」(34, n. 1)(粗體為筆者的強調)。當最後，是人偶的證據比是人的還鮮明時，前者力量勝過後者，前者會要求我去「跟隨」這個證據：「『甚麼在那』，這個問

題從**實事、從事物自己**去決定。而自我是透過實事自身決定，**跟隨著實事的決定**」(51) (粗體為筆者的強調)。由此可知，對胡塞爾來說，對象意義並非心靈單向的建構活動，而是對象自己會要求我的知覺，去使對象可以自身如其呈顯。在意識合理的狀態下，知覺會去配合實事自身的要求。

我認為傑佛斯對胡塞爾內在論式的詮釋過於片面；他用弗雷格 (Friedrich L. Gottlob Frege) 的方式解讀胡塞爾的意義理論，並且將超驗還原強硬地只解讀為來自笛卡爾 (René Descartes) 傳統的獨我論 (Zahavi, 2004: 42)，而排除其他可能性。視域是意識在時間內在建構下所產生的經驗結構；雖然此建構來自意識流內在的時間綜合，卻是對世界和他者開放的，並與主體從日常經驗中獲得的認知、身體習性、興趣和價值情感有關。雖然數學知識不論對誰或在哪儿都絕對為真，但經驗性的知識一定是透過環境因素去規定認知內容，單靠意向性本身絕無辦法。

除此之外，傑佛斯還提出專家知識，透過亞里斯多德實踐智 (phronesis) 的概念來說明專家技巧性的知識 (expertise skills) (Dreyfus, 2005: 50)：一旦人真的學會了、熟練了規則，是在情境的沉浸與情感涉入 (emotional involvement) 之中去實踐，而非無時無刻還在將注意力放在規則或優劣計算的判斷上。傑佛斯試著用如此熟練化、技巧性化的專家知識，來反對智性主義主張的知識論。但我認為，如果主張是為了說明心靈最原初是非智性的日常經驗狀態，傑佛斯反而提出了更高的條件去說明經驗與實踐：實踐主體首先要能認識一定程度的專業規則、反覆熟練，還要能在關鍵時刻真的實踐出來。但一個小孩在剛學習走路或丟球認識地心引力的時候，身體就已經開始在訓練如何和世界打交道，並在處境中實踐自身。

## 肆、知覺與意識生活

### 一、人工智慧與人類智慧在知覺系統上的差別：生命

人工智慧的發展，傾向完善功用性的系統知識；盡可能地列舉出人類認知的各種特徵，將認知活動寫成參數，模擬出涵蓋這些現象的程式系統。目前模擬的不僅只是人工認知，還有人工道德 (artificial mortality)、擬真的人工運動系統 (artificial sensorimotor)，甚至人工意識 (artificial consciousness) (Gamez, 2008; Reggia, 2013) 等等。當人工智慧將每個人類心智活動都找出理論上模擬的可能性，只待技術更加純熟和完善；人工智慧是否「具有思想」的差異似乎也越來越模糊。那麼，人類有別於人工智慧「具有思想」的本質差異，除了上述所說的原初意識經驗，以及其身體和世界的動機相關聯等，還有哪些？

一九九〇年代瓦雷拉 (Francisco J. Varela) 等人提倡結合不同科學領域共同研究意識，提出「腦神經現象學」(neurophenomenology) 的計畫 (Varela, 1996; Varela et al., 1992)。他們推動，在自然科學進行腦神經科學現象學化 (phenomenalized neuroscience)，而在現象學進行現象學自然化 (naturalized phenomenology)。他們提倡使用四個進路去研究心靈、研究認知：身體性 (embodied)、能動性 (enacted)、延展性 (extended) 以及在世性 (embedded)，簡稱 4E cognition (Gallagher et al., 2015)。<sup>9</sup> 認知並非

---

<sup>9</sup> 4E cognition 的計畫也引起很多討論和批評 (Newen et al., 2018)。身體和世界究竟如何影響心靈認知，4E 的進路雖然勾勒出了一個研究藍圖，但細節仍是講不清楚的 (Aizawa, 2018)。就筆者的觀察，相較於身體性，時間的因素在 4E cognition 中，似乎沒有受到較大的重視。在人工智能的發展中，有個從程式函數模擬到探究運動感知 (sensorimotor) 的轉向 (Butz & Fuster, 2017: 61)。運動感知試著模擬行為上的認知，像是昆蟲，牠們沒有精密複雜的大腦機制，但動作非常靈敏，而且可以相應環境做出精確的改變。科學家開始探究這種「技術性的認知」為何，因而在目的導向的行為 (goal-

只是封閉在腦袋中的神經活動，並非只是桶中腦，而是和身體、經驗、環境、交互主體的社群、工具文化、世界性等有關。<sup>10</sup>

而近幾年機械意識的研發中，也有人提出「人工合成現象學」(synthetic phenomenology) 的進路 (Gamez, 2006, 2008)。他們試圖去透過各種參數模擬人工的神經網，去描述知覺現象，並揣摩如何可使得機械系統變得擁有意識。除了物理實體的機器人，他們建造了相應的內在模型 (internal model)，去模擬各種意識活動。他們區分出四種根本的心理活動，並認為這四種心理活動可以使得系統變得有意識：(1) 系統能夠描述知覺狀態，並陳述自己的位置。(2) 想像力，可以喚起過往與世界相關的知覺。(3) 注意力，可以選擇某個在世界中的知覺對象，以及何種想像力內容。(4) 對接下來行動的規劃。(5) 情感，可以去衡量和評估抉擇與行為 (2006)。

如此的內在模型連結了機器人的身體行為，並涵蓋了和環境的關係。人工神經網絡的視覺系統是以生物視覺為基礎建造。它可以模擬出，相應環境刺激而自發性地產生眼球運動，因而機械眼睛可以自主地「學習」看清楚並認識一個對象。情感參數也納入感官輸入和動作輸出的計算過程之中，因而可以參考價值需求，選擇想要看的對象。對於人類認知中的自我學習，程式系統也可以模擬出回

---

oriented behavior) 中，有了較大的進展。他們發現，如果考慮行動中記憶 (memory) 的部分，將會增進行為的效能 (224)。行為不只是身體空間性的，也是時間性的，會隨著經歷而修正到最好。胡塞爾可以對 4E 研究提供相對完整的理論資源。胡塞爾意識建構理論中，對肉身性和世界性的建構都提供了清楚的分析。雖然他的理論從意識出發，但這只是我們從第一人稱切入的起點、一個工作場。意識和世界是相互開放、相互建構動機相關關係的。他的意識理論並非一種極端的超驗主體主義，而是充分考慮現實性 (Faktizität/facticity)。胡塞爾並非如傑佛斯所說，只是封閉獨我的心靈表象理論 (Dreyfus, 1980: 98)；相反地，認知要在身體對世界的開放性中展開。

<sup>10</sup> 蓋勒格 (Shaun Gallagher) (2009, 2012) 則表示，瑟爾中文房論證中的人，若走出房間和中文世界的人、環境與社群進行具體肉身性的經驗互動，語意便可以從語法中發展出來。

饋的機制，進行自我修正，以模擬人的認知過程 (Newen, 2017)。這些模擬的神經網絡試著去揣摩，內在系統在甚麼程度上可以被稱為是具有意識的。

合成現象學的確捕捉了意識裡重要的心理活動。意識的內在狀態以及身體的輸出是一個相關聯的系統。除此之外，合成現象學所模擬的情感、想像力、籌劃等意識活動，也的確屬於高度與人格性相關的意識活動。

在胡塞爾發生現象學的意識分析脈絡中，情感、想像力不僅只是單一的意識活動，而是會和意識的統一性以及意識的人格化有關。情感和想像力，在考慮到意識流內在的動機發生下，會與觸發和人格習性的建構有關。人格會在意識內在不間斷的被動建構下——也就是沉澱和再發生下——建立起習性信念以及情感。而如此的沉澱和再發生也會形塑出相應每個個體人格性的觸發結構。

合成現象學雖然沒有考慮意識底層的內在建構，但他們所選取要模擬的意識活動的確是具有代表性的。但即使如此，我認為，就算這些重要的心理活動和意識特徵都被模擬出來，我們也仍無法忽視機械意識和人類意識之間根本性的差別，也就是意識底層內在建構所展示的「生命性」(Thompson, 2007)。

程式計算模擬出的意識仍非自身為有機體，是非生命的。<sup>11</sup> 而人的意識經驗有其「活生生」(Lebendigkeit) 的特徵，有底層自主且持續不間斷地內在建構。此底層的時間建構作為所有心理活動的條

---

<sup>11</sup> 本篇文章中對人工智慧與人類智慧之差別的討論，無意表達出孰優孰劣的評判。很多能力，人工智慧優於人類，反之亦然。即使人工智慧不具有生命意識，其成果為人類科技社會帶來的便利性也顯而易見。筆者試著探討的，仍就是「具有思想」的本質為何，並且闡述，在胡塞爾作為歐陸哲學知識論的脈絡展開下，「具有思想」也蘊含著「人是甚麼」、「意識生命是甚麼」的內容。而這也是在人工智慧快速地進展下，促使我們加深對這些問題的思考。

件基礎且涵蓋整個生命歷程。即使函數程式模擬出了類似常識的系統、模擬出自我反思，但關鍵的差別在於，人類的意識是從底層自發性產生連續不斷的自身建構。意識底層的內在建構，是一直在動態和改變之中成就一個整體性。它是自發性的、自主性的、自為因果 (self-causality) 的概念 (Fuchs, 2021: 41)。在此底層不間斷的建構下所成就的人格 (Person)，也是一直在成為 (Werden/becoming) 的過程之中；這個過程是朝向開放性的。意識本身就是首要原初的。意識經驗內部如此自發、自主的內在建構，因而與人工智慧輸入輸出的程式設計，有著本質上的差異。

總結上述對知覺的討論，我們可以簡略地用表 1 概括人工智慧和人類智慧的差別：

表 1 機械意識與人工意識的差別比較

機械意識	人工意識
模擬 (simulation)	生命性 (aliveness)
內在模型	自發的 (spontaneous)、持續不間斷的內在建構
第三人稱的模仿	第一人稱觀點的經驗
大量資料庫深度學習	單一知覺中即可辨別部分與整體
函數關係 模擬的知覺—身體—環境	動機關係 知覺—身體—環境
輸入—輸出 程式指令	觸發、前反思的自我意識 (self-awareness)
數學必然性	非必然性，觸發
模擬的身體動作	可反身性的身體經驗

資料來源：作者自行整理，並參考 Fuchs (2021)。

## 二、這是否蘊含，要來自於自然類才「具有思考」？

人類意識在機械意識的根本差別在於，機械意識自身仍非有機體、非具有生命。那麼，這是否蘊含，胡塞爾會跟瑟爾一樣，認為人工智慧的程式系統因為不是來自於「自然類」(natural kind)，所以即使窮盡所有心靈活動，也仍不「具有思想」、不具有理解？也就是說，是否意識的心靈能力無法來自合成的電路組合，而是一定要「來自於」自然有機體或神經系統？

瑟爾反對機器人具有思想的理由，可整理為主要兩點 (Mensch, 1991: 107)：(1) 電腦語言程式的模擬只能產生語法 (syntax) 的連結，但並不包含語意 (semantics) 的部分。人的心靈狀態是語意的，而語意的部分無法由語法產生 (Searle, 1990: 27)。(2) 機器人的輸入輸出是由程式模擬，是來自晶片、電力、物理等因果過程產生的，並非自然類。要自然類的生理物理 (physiological) 關係，才能產生能思想的心靈。瑟爾寫道，「認知是生物現象」(29)。意向性是來自有機生理的產物，無法透過物理因果取代；即使把程式灌入內在反思或回憶等高層次的心靈狀態，也無法等同人作為自然物所擁有的意向性。

我們可以進一步深入第二點。對瑟爾來說，即使機器人有反思和心靈能力，也不能算是「具有思想」；原因在於，機器人不是來自自然生物的演化。我們因此可以推論，對瑟爾來說，作為自然演化的生物種類，是「具有思想」、具有心靈的條件之一。

就我所知，胡塞爾並未針對意識經驗是否一定要來自自然演化做過評論，但他對意識的分析，的確總是針對動物的意識。我們可藉由他對人之自然與精神 (Geist) 的雙重層面，探討他可能的想法。

胡塞爾表示，意識的確會受到物理自然部分的影響，像是「喝酒會變開心」(Husserl, 1952: 391)、喝了山道年 (santonin)<sup>12</sup> 視覺看到的東西會變黃色、感官感覺要靠著感覺器官和神經才得以可能。意識和自然是在一個相互的作用之中 (62, 288-289)；意識的確某個程度可以被自然化。我們的心理是依賴於身體的，具有因果關係。

但是，儘管意識有可被自然化的一面，也仍是有限度的。即使是身體物理的因果關係，我對其仍是有相應的原初經驗 (Husserl, 1952: 391)。就像上面舉的例子，喝酒會產生開心的感受、藥物會改變視覺顏色，這都是在經驗中原初給予我的；換句話說，我對這些因果關係擁有第一人稱的直接經驗。自然物理對我身體造成的心理狀態的改變，我對如此的因果關係仍是有原初經驗。但是，我心理意識的內容，無法單純透過腦神經的作用來解釋。

自然科學雖然可以解釋精神受物理自然牽制的部分，但無法真正說明精神的重要之處。在胡塞爾現象學中，「精神」要被視為獨立的領域來研究。自然無法取代精神在存有論上先於物理自然的優先性與重要性 (Husserl, 1952: 298)。意識是活生生的，並具有先天的建構能力，如此的能力並非來自自然的物理生理，而是來自於精神的層面。胡塞爾甚至主張，自然的地位只是相對性的，真正具有決定性地位的，是精神。胡塞爾寫道：

當精神在和相關於人格世界自身建構的自然的關係之中，展現出多重的依賴性時，它的確是透過它的周遭環境所決定，並且服從自然法則。但是，這不影響它是「絕對的」、「非相對性的」。(297)

瑟爾所說的自然類，是屬於胡塞爾所區分的動物自然 (animalische Natur / animal nature)，也就是靈魂 (Seele) (Husserl,

<sup>12</sup> 胡塞爾那個年代可以用來治療蛔蟲等線蟲感染的藥物。

1952: 90)。在《觀念二》(*Ideen II*) 中，胡塞爾將人抽象地區分為「物質自然」(*materielle Natur / material nature*)、「動物自然」和「精神」三個部分。動物自然（靈魂、也就是自然類）的部分並非使思考得以可能的來源。對胡塞爾來說，意識的建構能力，並非來自自然物理的部分，也並非來自於動物自然（自然類），而是「精神生活」(*Geistesleben / life of spirit*)。因此我認為，胡塞爾並不會將「來自自然類」作為「具有思想」的必要條件之一。

「具有思想」，是因為具有精神生活、具有意識心靈的建構能力，而這個能力並非單純來自於有機體神經或自然類。因此有別於瑟爾，胡塞爾理論中精神生活的建構，有其自身法則的來源 (*Quelle*)，其來源並非可化約到動物自然，也就是生物體上。認知活動有和腦神經相應的部分，但認知活動的起源並非來自腦神經。更仔細地說，在我對胡塞爾的詮釋下，具有意識經驗是真正「具有思想」的必要條件；「具有思想」會因此蘊含具有生命，但「來自有機生命」並非直接是「具有思想」的必要條件，因為精神領域仍有其自身法則的來源。

意識是一個複雜的心靈現象。我們如何在生理感官的作用下，產生知覺並獲得到關於對象的認知，在不同哲學學派也有不同立場和主張。意識究竟如何產生，也是心物 (*mind/body*) 問題中的艱難問題 (*hard problem*) (*Chalmers, 2007*)，至今科學界和哲學界的說法眾多而不同。從胡塞爾精神和自然兩個領域的區分，我們可以抽象分出三種心物法則：心—心法則（精神）、<sup>13</sup> 心—物法則（動物自然）<sup>14</sup> 以及物—物法則（物質）。不只是心物法則目前尚未有定解，精神生

<sup>13</sup> 精神生活。如上述所說的身體性知覺活動、第一人稱的動機脈絡、能在和世界互動的經驗中去自主建構出生命歷史和認知習性等。

<sup>14</sup> 例如意識生活的身體建構，像是身體記憶、老化 (*ageing*)、身體的疾病，或甚至像是阿茲海默的疾病、死亡等。

命自身的心—心法則也未必能清楚說明。動機關係是相當有別於物理因果的法則。<sup>15</sup>

或許我們可大膽推論，若真的能窮盡找到這三種法則以及它們之間交纏的作用，也許製作出真正「具有思想」的人工智慧是可能的。對此單只是塑膠、電線、鋼鐵，是不夠的，還要有可以製作具有如此意識和肉身狀態的物質技術（或像是複製人）。但這已不再是單純的人工智慧，而是具有「生命」。而如此能自主組織建構的意識生命，是否還能再稱為「人工」，也還有待商榷。<sup>16</sup> 因為自主建構的意識生命，已經是和我們一樣存活的心靈。

但如此的討論相當未來主義 (futuristic)，因為胡塞爾關心的不是意識透過如何的物質產生，而是人類精神生活本身的動機法則為何，以及其和物理法則的差別。而精神和自然的斷然區分仍是抽象的。對胡塞爾來說，雖然精神和自然擁有不同的本質和法則，實際情況上，兩者在人身上卻仍是一個相互交纏的整體。

在此我們退一步可掌握的哲學意義在於，我們看到，思想和意識經驗的關係；以及進一步地，思想和生命的關係。這也衍生到思想和人的關係。在歐陸哲學中，對思想和理性能力的考察，與「人是甚麼」的哲學人類學關懷息息相關。我們所探討的思想，並非是單純的算術與邏輯系統，而是具有意識經驗；其主體亦非空洞的我

---

<sup>15</sup> 關於胡塞爾現象學中的心物問題，以及動機法則如何有別於物理因果法則，無法在這篇文章的篇幅中處理。

<sup>16</sup> 富克斯 (Thomas Fuchs) (2021: 40) 則表示，「人工生命」是相互抵觸的概念。如果是生命，就不再是人工的，而是和我們一樣擁有生命、血肉、會死亡的存有 (living being in flesh and blood)。生命本身就代表自主組織 (self-organization)。礙於篇幅，這篇文章無法再深入處理這個議題。關於生命現象與意識的關係，可參考 Thompson (2007)。

思 (cogito)，而是擁有具體肉身、時間老化、在世界與他者之間相互影響中形成。

## 伍、外在身體擬真的互動，是否就代表機器人「具有思想」？

圖靈透過模擬遊戲 (imitation game)，<sup>17</sup> 去探討機器和人類之間的差別；並試著回答「機器是否可以思考」的問題 (Turing, 2004: 441)。圖靈的進路是由外部行為去歸屬思想活動。<sup>18</sup> 以圖靈的模擬遊戲為起點，我們讓機器人走到戶外和觀察者互動。就如同現在已

---

<sup>17</sup> 圖靈測試中的模擬遊戲中是這樣三個人的：A 和 B 分別和關在房間內的 C 透過電報的方式交流。A 是男性，B 是女性，但 C 不知道。C 在遊戲中要透過寫電報的內容，去猜測個別的性別。遊戲中最主要的機器人測試在於，A 由機器人來扮演，看能不能讓 C 在沒察覺異樣的情況下去玩遊戲。如果 C 沒有察覺異樣，那扮演 A 的機器人就可說是通過了測試；我們就可以說 A 是可以／具有思想的。C 和 A 並沒有實際面對面，而是只有寫電報。

<sup>18</sup> 圖靈測試顯示了對思想本質考察的不足，並引發了許多哲學家的評論。像是戴維森對其測試進行了修正，認為受試者要在戶外，讓觀察者有足夠的時間去和受試者親身互動並考察他和外在環境的互動關係。戴氏分析在甚麼樣的條件下，我們——作為觀察者——可以合理地將受試者詮釋為一個思想者。透過觀察對方和世界對象互動的因果關係，以及透過實際交流去觀察其語言和思想的歷史性；藉此檢視對方是否真的具有相應於其語言的心靈狀態。並且，對戴氏來說，一旦我們真的有足夠合理性可以將思想歸屬給對方，就已排除了對方「沒有思想或者其思想全都是錯的」(趙之振，2020: 175) 的可能性。趙之振也探討了這個區分：可以合理地把思想歸屬給機器人、將機器人視為是有思想的，如此就可以代表，機器人是自己真的可以思想的嗎？透過徹底詮釋和慈善原則去探討是否能合理歸屬思想給對方，似乎已經假設了，對方思想最大程度是和諧且對的 (166, 172, 174)。「機器是否能被視為思想者，主要不是取決於機器與人類外在行為有多相似，而是取決於人類有多大的意願先行把機器看作與自身(身為思想者)相似」(172)。接下來的部分，並沒有要進入戴氏的思想探討。但我們接受戴氏對測試的修正，讓機器人走出來和觀察者親身互動。以此為基礎，我接下來仍是就胡塞爾的角度，提出現象學的看法。

越來越多社交機器人 (social robot)<sup>19</sup> 實際走進我們的日常生活中，和我們互動。那麼，「機器人是否可以思考」的問題，可以依賴於，「機器人的外部行為是否和人一樣」嗎？意識的思想活動，可以化約到身體行為上嗎？

以胡塞爾現象學角度來說，意識精神無法單純化約到外部語法或物理身體上。「具有思想」代表要擁有活生生的知覺經驗和第一人稱的動機網絡。活生生的肉身經驗，並非是第三人稱的機械身體可以取代。外部的語法表達或機械式的身體交流無法作為真正「具有思想」的充分條件。

在胡塞爾基本的他者理論中，肉身 (Leib / living body) 的表達是一個很關鍵的因素，讓我把對方經驗為像我一樣具有意識生活的主體。在和他者互動中，透過對方身體的表達，我可以共顯出對方的心靈狀態。我們對自己的肉身知覺 (Leibwahrnehmung / perception of living body) 都擁有第一人稱的運動感。理論上，活生生的肉身表達和行動，的確可以直接表達出對方的認知和意識狀態。對方的行為會表達他對世界因果關係的認知，以及他的意圖和心理狀態；像是同學把圖書館的窗戶關起來，我可以理解為，他覺得風吹起來冷，而他認為關窗可以讓風不吹進來。「他心」在胡塞爾他者理論中並非封閉不可知，而是可以透過活生生的身體經驗表達去通

---

<sup>19</sup> 有別於圖靈模擬遊戲中的測試，機器人要試著讓自己跟真人一樣和人互動，去通過測試；目前許多社交機器人的設計本身就誠實表達自己是機器人，並提供自己被設計的長才去協助需要的人。像是 Amazon Echo (Alex) 提供大量充分的資訊、機器狗作為寵物提供陪伴；社交機器人也被設計為讓小孩學習社交情緒、或是協助自閉障礙的兒童等的教育功能，像是 Moxie (Ackerman, 2020)；另請參見 Pashevich (2021)。如此和小孩互動、或僅是以提供功能為目的的人工智慧，並不著墨於高程度的擬真。這部分探討的社交機器人主要是針對擬真的機器人。但是，即使沒有高程度的刻意擬真，只要有簡單相似形狀的軀體，也足以達成「宛如真人」的素樸聯想。下文會繼續對此說明 (參考註腳 24)。

達。對胡塞爾來說，他心顯現在他者身體經驗的表達之中。而活生生的身體經驗，還會有其成長、生病、死亡等的歷程。

擬真的機械身體模糊了和真人的區分。在社交機器人中，2017年開始活躍的人工智慧機器人索菲亞 (Sophia)，設計者強調它和人類一樣甚至「超智慧」(super-intelligent)。設計者試著透過擬真的機械身體和程式語法，去說服群眾，索菲亞真的具有「具有思想」。索菲亞擁有逼真擬人的面部表情 (mimic)，除此之外還有機智的修辭與談諧的對話。看過訪談，我們都不免覺得索菲亞貌似為一位聰明風趣的女性。與它對唱情歌的主持人甚至說，他彷彿愛上了它。但只要我們的觀察時間一長，就可以馬上發現，索菲亞對話、表情和採訪都是事先程式化好的。它離設計者所強調的超智慧還有一段很遙遠的距離。<sup>20</sup>

接下來，我將說明，和如此的社交機器人互動時，我們的意識知覺經驗為何？和索菲亞擬真的互動，就代表它擁有心靈、「具有思想」嗎？我們如何一方面確定機器人並非真人，又另一方面同時有宛如真人的感受？社群機器人作為虛擬 (virtual) 實在物進入日常生活，的確或多或少動搖了我們對實在性 (reality) 的理解。

在共感經驗中，想像力 (Phantasie/phantasy) 佔有一定的角色。在透過身體知覺和他者互動時，胡塞爾寫道，「彷彿我走過去在那」(Husserl, 1950: 146)、「彷彿我在那邊」(147)、「就彷彿我站在對

---

<sup>20</sup> 索菲亞在世界巡迴中深受大家的注目，甚至獲得沙烏地阿拉伯的公民身分，是所謂第一個非人類的公民。索菲亞設計者試圖宣傳出人工智慧可以取代人類的想法。他們強調索菲亞是「超智慧」，甚至「超人類」；因為人工智慧的計算功能以及資訊體系比人類來得強大。關於索菲亞如何在媒體宣傳裡刻意被塑造出「超智慧」的形象，以及背後科技和媒體宣傳的關係，可見 Parviainen & Coeckelbergh (2021)。「具有思想」是否可以就如索菲亞的設計者所宣稱的，透過逼真的面貌和談諧的修辭就可以達到？如筆者此節所要論證的，是不足夠的。

方身體那邊一般」(152)。皮膚和五官表情擬真的機器人，讓我們更容易進入情境，把他視為彷彿真人一般。即使是非擬真的社交機器人，也透過具體的聲音和身體，將虛擬性實在化到現實的社交互動之中。

想像意識的作用在日常生活中層出不窮。像是看電影時，我們知道，這是虛構出的故事，但我們同樣會感到害怕、緊張、陷入憂愁或愛情。我們對現實性清楚的掌握，不影響我們融入劇情。現象學會這樣解釋：想像力的作用就猶如知覺一般；所有知覺有的作用，想像力也都有，只是是以「彷彿」(Als-ob / as if) 的模態 (Husserl, 1972: 197)。胡塞爾在描述看到人偶時的知覺經驗寫道：

我們雖然「知道」，這只是假象，但我們無法克制，我們彷彿看到了真人。伴隨著的概念判斷，訴說著這只是如一幅圖像；但這概念面對知覺顯像時，是沒有作用的。而我們想要把他看成現實性的傾向，是如此地強而有力，以至於我們某些時刻真的想這樣相信。(1980: 40-41)

想像意識擁有宛如知覺一般的作用：「彷彿」的認知、「彷彿」的存有設定、「彷彿」的情感觸發，甚至是「彷彿」的現實性 (Quasi-Wirklichkeit / quasi-reality)。即使只是彷彿的模態，對我們的意識經驗，卻仍會起一定程度的作用。

對現象學的進路來說，我們不從第三人稱全知的觀點，去決定現實性 (Wirklichkeit)<sup>21</sup> 為何。而是從第一人稱經驗的觀點，去考

---

<sup>21</sup> 在胡塞爾的脈絡中，德文 Wirklichkeit 和 Realität 雖然英文都是對應為 reality，但兩者有不同意思。Wirklichkeit，我翻為「現實性」，與第一人稱經驗中的明證性 (Evidenz) 引導有關；我透過經驗裡被給予的線索，去進行相應的現實性設定，也就是相應的存有設定。像是眼前的人像不確定是真人還是玩偶；若知覺中真人的顯像具有比較高的說服力，我便會下「可能是真人」的判斷。而 Realität 是「實在性」，指涉到客體對

察，經驗給予了我甚麼理由，讓我有合理性，去促使接下來的判斷、情感反應或行動。合理性在這裡聽起來是很高的層次，但如之前所說，動機關係有分不同等級。除了最高符合一致性的合理性，底層的「合理性」像是情感、慾望、本能的動機觸發，也是在經驗中給予並觸發我們：自我受到情感的觸發，促使了一個想法或做了甚麼。

我們可以用例子來說明低層的合理性（動機促使）。癮君子雖然被人覺得「不理性」（當人考慮健康的因素和醫學報導），但在第一人稱經驗當中，習性動機觸發的力量很強，主體在這股力量的促使下去拿菸。就經驗給予她／他的，對她／他來說，是「合理的」；因為意識經驗提供了她／他強度足夠的動機去這麼做。她／他並非憑空雜亂無章地去做。在意識結構中，一個經驗中可以存有不同的傾向和觸發動機，胡塞爾形容它們各自有不同的「重量」（Gewicht/weight）；力量強的會壓過力量小的。透過練習，我們可以建立一個抵抗吸菸的動機，當抵抗吸菸的意識傾向開始大於吸菸的習性傾向時，停止吸菸的動作就會產生；因為促發我身體的不再是吸菸的傾向，而是力量較強的停止拿菸。

類似於此，我一方面知道，對方是不具有心靈狀態的機器人，但知覺經驗中給我的各種觸發，像是語言、身體動作，仍會輕易地引發和促使我透過認知習性去理解和給予語意上的詮釋。對方只是機器人之存有信念的傾向，與出於習性詮釋對方為語意的傾向，並非互斥，而是共存。即使我由此獲得資訊以及娛樂，也不代表，對方就「具有思想」。

人的意識生活受語意的形塑很深，語意滲透在我們的認知習性中。在我們的日常生活的知覺經驗中，語法和語意、面部表情和意

---

象，像延展於空間的形狀或具體的聲音（不涉及內容），還未考慮在第一人稱經驗中的顯像。

識的感受狀態等，之間的連結十分強烈。尤其當機器人刻意模擬人類表情和說話時，會促使我們覺得對方彷彿跟日常生活中打交道的他人一般，具有活生生的知覺意識。在社群生活經驗滲透甚深的建構下，機器人即使只是簡單相似的身體和語音表達，也可以十分容易觸發我們的認知習性。<sup>22</sup> 這些符號在我們意識生活的結構中，具有非常強大的觸發力量。

在意識底層的結構中，這些意向活動（身體的驅力、傾向）並非雜亂無章或從天上掉下來突然產生的，而是有其動機發生的過程、有意識觸發或動機促成的過程。<sup>23</sup> 就意識流本質的結構，不同意識活動之間都是環環相扣的。如此的立場也蘊含說，胡塞爾將會贊成，情感，尤其是底層的慾望、想要、驅力、傾向，也是有意向性成分在的；我們可以說它們是「有理由成分的」（Heffernam, 2009），因為他們會對我接續的意識內容、想像或動作有所影響。德文「現實性」（Wirklichkeit）的字源，和「起作用」（wirken）有關

---

<sup>22</sup> 我們人跟他者的知覺經驗，不只是思想和語言上的交流，更是身體性的。具肉身性的互動是在視域背景中，使他者經驗得以可能的、未彰顯的底層基礎。就如文中陳述的，我們在日常社群生活和他人互動的長期建構下，即使只是簡單的相似性，也足以引起具有影響力的觸發。比起只是在透過電腦進行虛擬的互動，有一個具有實體的機器人可以主動告知訊息和觸碰，在人類意識中對感知觸發產生的影響相當不同。占卜機器人（Astrobot）便是在探討，軀體實在性的加入是否會，以及會如何改變人類和機器人的社交經驗。占卜機器人是能提供占星術的資訊，並具有實際軀體的社交機器人。占卜機器人的設計誠實地表達自己是功能性機器人。這個計畫在探討，即使互動者一方面知道它是機器人，但另一方面，透過具身體性的社交互動以及環境氛圍的營造，是否會促使占卜機器人所提供的訊息對人是具有說服力的。尤其是當占卜的內容並非屬於事實性的資訊，互動者是否還可能會受促使去相信，具有身體的占卜機器人所提供的訊息（Semjonovs, 2020a, 2020b）。智能機器人身體性的社交互動，對我們的意識塑造了特殊的他者經驗。比起自己面對電腦查訊息，智能機器人提供的資訊，透過身體性互動經驗的觸發，變得更融入人們的日常生活。

<sup>23</sup> 即使我們不認識或沒有意識到這些來源，或許這些來源也無法真正透過主體事後的反思澄清出來。

(Lohmar, 2018: 38)。即使是彷彿模態的現實性，也可以對意識起作用，讓我們去做出相應的存有設定 (Seinssetzung)。透過意識經驗底層的動機關係 (理由成分)，我們因此可以解釋，即使是虛擬對象，其具體的實在性 (聲音、身體) 仍可以對我的意識造成觸發，提供我和機器人互動的合理性。<sup>24</sup> 我們因而可以一方面全觀地認為，對方仍只是模擬的機器人，還不真正具有生命；但另一方面，在當下的第一人稱經驗中，擬真的互動讓我們覺得對方彷彿如真人一般。

這些和社交機器人互動的基礎，是上述說明的想像力經驗、相似性觸發等作用提供了合理性。社交機器人對觀察者的第一人稱知覺營造了一些逼真的相似性，去觸發我們對真人的聯想。但這些理由都與機器人真的「具有思想」、或是「具有活生生的肉身表達」無關。即使具有擬真的實在性，但沒有真正相應的第一人稱肉身知覺，我們也不能說那是活生生具有血肉的身體，而只是一般的物理軀體。對胡塞爾現象學來說，人工智慧是否「具有思想」，仍是要從是否擁有肉身經驗和第一人稱的動機脈絡來看，而非從外部擬真的身軀、相貌和語法。

## 陸、總結

在這篇文章中，我主張「具有思想」需要具備一套相應的知覺活動和意識生活。這篇文章也討論了人工知覺 (第貳節)、人工意識 (第肆節第一部分)，並試著在胡塞爾現象學的脈絡下，給出理由，為何人工知覺和人工意識與真正「具有思想」的條件尚有差異。而機

---

<sup>24</sup> 像是機械寵物也可以成為情感的陪伴者。使用者情感依賴著寵物機器人，但他們仍是認知清醒且非病態的狀況。想像力在其意識建構也是扮演著有重要的角色，提供底層合理性的基礎，而使得如此的情感經驗 (affective experience) 得以可能 (Weber-Guskar, 2021)。

器人僅僅只是外在的擬真互動，也無法作為機器人「具有思想」的充分條件。這些外在的模擬能力，只是讓人工智能在與人互動時看起來更有說服力。但具有說服力，並不代表本質上即為如此。

在這篇文章中所涉及到的意識活動，不管是高層的形式知識的意向性直觀、知覺認知的動機關係、或是肉身經驗的表達，都可涵攝為一種相當廣義的理由關係、一種活生生的「意向性關係」。肉身知覺的動機關係範圍相當廣泛，可能是在未反思的情況下做出一個動作，沒有彰顯的理由內容，或不一定給得出理由解釋。但即便如此，也是在動機關係收攝之下。在我的胡塞爾詮釋中，不管是在判斷或知覺層次、主體主動或被動的面向，我們的認知都是由意向性建構來進行，都是一種最廣義的理由關係。因為即使在意識底層，我們的認知和意識活動也不是混亂、無中生有的，而是有其發生法則。

「具有思想」，代表要能擁有這樣一套動機網絡連結的意識經驗。認知是具肉身化的，是在時間的建構下、和情感、運動感與世界性息息相關的動機網絡之中展開。意識不只是單純的認知活動，還是活生生的意識經驗，具有內在自發組織的「具肉身的生命」(embodied life)。<sup>25</sup> 如此的探討，並非在表示生物的優越性，而是試著深入我們對思想的理解，以及，我們作為人的思想如何因為「意識生活」而有別於人工智慧。

---

<sup>25</sup> 「具肉身的生命」一字來自參考“embodied mind” (Thompson, 2007) 以及 “embodied being” (Fuchs, 2021)。

## 參考文獻

- 趙之振 (2020)。〈戴維森論圖靈測試〉，《歐美研究》，50, 2: 147-178。 (Chiu, C.-C. [2020]. Davidson on the Turing test. *EurAmerica*, 50, 2: 147-178.) [https://doi.org/10.7015/JEAS.202006\\_50\(2\)0002](https://doi.org/10.7015/JEAS.202006_50(2)0002)
- Ackerman, E. (2020, April 28). *Meet Moxie, a social robot that helps kids with social-emotional learning: A new social robot designed for children wants to help promote social, emotional, and cognitive development*. IEEE Spectrum. <https://spectrum.ieee.org/>
- Aizawa, K. (2018). Critical note: So, what again is 4E cognition? In A. Newen, L. D. Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 117-126). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198735410.013.6>
- Butz, M. V., & Kutter, E. F. (2017). *How the mind comes into being: Introducing cognitive science from a functional and computational perspective*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198739692.001.0001>
- Chalmers, D. (2007). The hard problem of consciousness. In M. Velmans & S. Schneider (Eds.), *The Blackwell companion to consciousness* (pp. 225-235). Blackwell. <https://doi.org/10.1002/9780470751466.ch18>
- Davidson, D. (2004). Turing's test. In *Problems of rationality* (pp. 77-86). Clarendon Press. <https://doi.org/10.1093/0198237545.003.0005>
- Dreyfus, H. (1972). *What computers can't do: A critique of artificial reason*. Harper & Row.
- Dreyfus, H. (1980). Dasein's revenge: Methodological solipsism as an unsuccessful escape strategy in psychology. *The Behavioral and Brain Sciences*, 3, 1: 78-79. <https://doi.org/10.1017/S0140525X00001849>
- Dreyfus, H. (2005). Overcoming the myth of the mental: How philosophers can profit from the phenomenology of everyday

- expertise. *Proceedings and Addresses of the American Philosophical Association*, 79, 2: 47-65.
- Dreyfus, H. (2012). Introductory essay: The mystery of the background *qua* background. In Z. Radman (Ed.), *Knowing without thinking: Mind, action, cognition and the phenomenon of the background* (pp. 1-10). Palgrave Macmillan. [https://doi.org/10.1057/9780230368064\\_1](https://doi.org/10.1057/9780230368064_1)
- Fuchs, T. (2021). *In defence of the human being: Foundational questions of an embodied anthropology*. Oxford University Press. <https://doi.org/10.1093/oso/9780192898197.001.0001>
- Gallagher, S. (2009). The key to the Chinese room. In K. Leidlmaier (Ed.), *After cognitivism* (pp. 87-96). Springer. [https://doi.org/10.1007/978-1-4020-9992-2\\_5](https://doi.org/10.1007/978-1-4020-9992-2_5)
- Gallagher, S. (2012). Social cognition, the Chinese room, and the robot replies. In Z. Radman (Ed.), *Knowing without thinking: Mind, action, cognition and the phenomenon of the background* (pp. 83-97). Palgrave Macmillan. <https://doi.org/10.1057/9780230368064.0011>
- Gallagher, S., Reinerman-Jones, L., Janz, B., Bockelman, P., & Trempler, J. (2015). *A neurophenomenology of awe and wonder: Toward a non-reductionist cognitive science*. Palgrave Macmillan. <https://doi.org/10.1057/9781137496058>
- Gamez, D. (2006). The XML approach to synthetic phenomenology. In R. Chrisley, R. Clowes, & S. Torrance (Eds.), *Proceedings of the AISB06 symposium on integrative approaches to machine consciousness* (pp. 128-135). Bristol.
- Gamez, D. (2008). Progress in machine consciousness. *Consciousness and Cognition*, 17, 3: 887-910. <https://doi.org/10.1016/j.concog.2007.04.005>
- Heaven, D. (2019). Deep trouble for deep learning. *Nature*, 574, 7777: 163-166. <https://doi.org/10.1038/d41586-019-03013-5>
- Heffernan, G. (2009). On Husserl's remark that "[s]elbst eine sich als apodiktisch ausgebende Evidenz kann sich als Täuschung enthüllen . . ." (XVII 164: 32-33): Does the phenomenological

- method yield any epistemic infallibility? *Husserl Studies*, 25, 1: 15-43. <https://doi.org/10.1007/s10743-008-9051-5>
- Holenstein, E. (1988). Eine Maschine im Geist. Husserls Begründung und Begrenzung künstlicher Intelligenz. *Phänomenologische Forschung*, 21: 82-113.
- Husserl, E. (1950). *Husserliana: Vol. 1. Cartesianische Meditationen und pariser Vorträge* (S. Strasser, Ed.). Martinus Nijhoff.
- Husserl, E. (1952). *Husserliana: Vol. 4. Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Phänomenologische Untersuchungen zur Konstitution* (M. Biemel, Ed.). Martinus Nijhoff.
- Husserl, E. (1966). *Husserliana: Vol. 11. Analysen zur passiven Synthesis. Aus Vorlesungs- und Forschungsmanuskripten 1918-1926* (M. Fleischer, Ed.). Martinus Nijhoff.
- Husserl, E. (1970). *Husserliana: Vol. 12. Philosophie der Arithmetik. Mit Ergänzenden Texten (1890-1901)* (L. Eley, Ed.). Martinus Nijhoff.
- Husserl, E. (1972). *Erfahrung und Urteil. Untersuchung zur Genealogie der Logik* (L. Landgrebe, Ed.). Meiner.
- Husserl, E. (1973). *Husserliana: Vol. 15. Zur Phänomenologie der Intersubjektivität. Texte aus dem Nachlass. Dritter Teil (1929-35)* (I. Kern, Ed.). Martinus Nijhoff.
- Husserl, E. (1974). *Husserliana: Vol. 17. Formale und transzendente Logik. Versuch einer Kritik der logischen Vernunft* (P. Janssen, Ed.). Martinus Nijhoff.
- Husserl, E. (1975). *Husserliana: Vol. 18. Logische Untersuchungen. Erster Band. Prolegomena zur reinen Logik* (E. Holenstein, Ed.). Martinus Nijhoff.
- Husserl, E. (1980). *Husserliana: Vol. 23. Phantasie, Bildbewusstsein, Erinnerung. Zur Phänomenologie der anschaulichen Vergegenwärtigungen. Texte aus dem Nachlass (1898-1925)* (E. Marbach, Ed.). Martinus Nijhoff.

- Husserl, E. (1984). *Husserliana: Vol. 19/1. Logische Untersuchungen. Zweiter Band. Untersuchungen zur Phänomenologie und Theorie der Erkenntnis. Erster Teil* (U. Panzer, Ed.). Martinus Nijhoff.
- Husserl, E. (2006). *Husserliana Materialien: Vol. 8. Später Texte über Zeitkonstitution (1929-1934): Die C-Manuskripte* (D. Lohmar, Ed.). Springer. <https://doi.org/10.1007/1-4020-4122-5>
- Lohmar, D. (2000). Warum braucht die Logik eine Theorie der Erfahrung. In O. K. Wiegand, R. J. Dostal, L. Embree, J. Kockelmans, & J. N. Mohanty (Eds.), *Phenomenology on Kant, German idealism, hermeneutics and logic* (pp. 149-170). Springer. [https://doi.org/10.1007/978-94-015-9446-2\\_10](https://doi.org/10.1007/978-94-015-9446-2_10)
- Lohmar, D. (2018). On some motives for Husserl's genetic turn in his research on a foundation of the Geisteswissenschaften. *Studia Phaenomenologica*, 18: 31-48. <https://doi.org/10.5840/studphaen2018182>
- Mensch, J. (1991). Phenomenology and artificial intelligence: Husserl learns Chinese. *Husserl Studies*, 8, 2: 107-127. <https://doi.org/10.1007/BF00123537>
- Minsky, M. (1980). A framework for representing knowledge. In D. Metzger (Ed.), *Frame conceptions and text understanding* (pp. 1-25). Walter de Gruyter.
- Newen, A. (2017). What are cognitive processes: An example-based approach. *Synthese*, 194, 11: 4251-4268. <https://doi.org/10.1007/s11229-015-0812-3>
- Newen, A., Bruin, L. D., & Gallagher, S. (2018). *The Oxford handbook of 4E cognition*. Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198735410.001.0001>
- Newen, A., Butz, M. V., Russwinkel, N., & Rothkopf, C. (2020). Was KI über unsere Intelligenz lehrt. *Spektrum.de*. <https://www.spektrum.de/>
- Nguyen, A., Yosinski, J., & Clune, J. (2014). *Deep neural networks are easily fooled: High confidence predictions for unrecognizable images* [Paper presentation]. Computer Vision and Pattern Recog-

- nition (CVPR 2015), IEEE, Boston, United States. <https://doi.org/10.48550/arXiv.1412.1897>
- Parviainen, J., & Coeckelbergh, M. (2021). The political choreography of the Sophia robot: Beyond robot rights and citizenship to political performances for the social robotics market. *AI & Society*, 36, 3: 715-724. <https://doi.org/10.1007/s00146-020-01104-w>
- Pashevich, E. (2021). Can communication with social robots influence how children develop empathy? Best-evidence synthesis. *AI & Society*, 37, 2: 579-589. <https://doi.org/10.1007/s00146-021-01214-z>
- Reggia, J. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, 44: 112-131. <https://doi.org/10.1016/j.neunet.2013.03.011>
- Searle, J. (1990). Is the brain's mind a computer program? *Scientific American*, 262, 1: 26-31. <https://doi.org/10.1038/scientificamerican0190-26>
- Semjonovs, S. (2020a). *Astrobot: The far end of persuasion* [Unpublished master's thesis]. University of Twente.
- Semjonovs, S. (2020b, May 4). *Interaction technology—Designing a believable social robot* [Video]. YouTube. <https://www.youtube.com/watch?v=fP30oINzNjg>
- Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2014). *Intriguing properties of neural networks* [Paper presentation]. Computer Vision and Pattern Recognition (CVPR 2014), Columbus, United States. <https://doi.org/10.48550/arXiv.1312.6199>
- Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Belknap Press.
- Turing, A. (2004). Computing machinery and intelligence. In B. J. Copeland (Ed.), *The essential Turing: Seminal writings in computing, logic, philosophy, artificial intelligence, and artificial life: Plus the secrets of enigma* (pp. 441-464). Clarendon Press.

- Varela, F. J. (1996). Neurophenomenology: A methodological remedy to the hard problem. *Journal of Consciousness Studies*, 3, 4: 330-349.
- Varela, F. J., Thompson, E., & Rosch, E. (1992). *The embodied mind: Cognitive science and human experience*. MIT Press.
- Weber-Guskar, E. (2021). How to feel about emotionalized artificial intelligence? When robot pets, holograms, and chatbots become affective partners. *Ethics and Information Technology*, 23, 4: 601-610. <https://doi.org/10.1007/s10676-021-09598-8>
- Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., & Lipson, H. (2015). *Understanding neural networks through deep visualization* [Paper presentation]. Deep Learning Workshop, 31st International Conference on Machine Learning, Lille, France. <https://doi.org/10.48550/arXiv.1506.06579>
- Zahavi, D. (2004). Husserl's noema and the internalism-externalism debate. *Inquiry: An Interdisciplinary Journal of Philosophy*, 47, 1: 42-66. <https://doi.org/10.1080/00201740310004404>

## Does Artificial Intelligence “Have Thoughts”? An Analysis from Husserl’s Phenomenology

*Ying-Chien Yang*

Husserl-Archive, University of Cologne

E-mail: yepyoung@gmail.com

### Abstract

As the technology on which artificial intelligence is based becomes more sophisticated, animated human-like interactions and super-intelligent computer programs will blur the lines between AI and human intelligence. The question arises, do artificial intelligences really “have thoughts” like humans? In this article I will argue, based on Husserl’s phenomenology, that artificial intelligence “can think,” and that the system can indeed respond, but this does not mean that it really “has thoughts.” Necessary conditions of “having thoughts” are perceptual experience and motivational relations. Such perceptual experiences are embodied and open to the intersubjective world. The difference between artificial intelligence and human intelligence does not lie in the higher level of linguistic capability, but in the more fundamental functions of consciousness. Consciousness has its own continuously immanent constitution, which is constantly in changing in processes of self-organization, and of “becoming.” This means that to genuinely “have thoughts” is to have an embodied life.

**Key Words:** consciousness, perception, cognition, artificial intelligence, Husserl’s phenomenology