

# 人工智能的“心識”限度 ——基於《成唯識論》的思考\*

朱敏  
哈爾濱理工大學

**摘要：**《成唯識論》中的唯識心法與人工智能之間沒有直接的聯繫，但唯識學中的“心”與“心所”較全面系統地分析了人類心理活動能力的各種可能，可以幫助我們思考人工智能發展中所遇到的一些問題，特別是人工智能的“心識”研究，使我們跳出人工智能研究的思維怪圈。本文主要從三個方面來談唯識心法和人工智能“心識”問題之間的間接聯繫：一是，《成唯識論》中對人類心理活動能力的各種分析；二是，從唯識心法的角度看人工智能方法的局限性；三是，思考人工智能的“心識”發展，主要涉及從唯識學角度來看人工智能的終極目標是什麼等問題。

**關鍵詞：**心、心所、人工智能、唯識

---

\* 本文為黑龍江省哲學社會科學研究青年項目“面向中文信息處理的自然語境下漢語隱喻理解的因果推理模式研究”(項目編號:16YYC05)和教育部人文社會科學研究青年基金項目“面向人工智能的漢語隱喻理解的邏輯推理模式研究”(項目編號:19YJC740123)階段性成果。

依託於計算機信息技術的不斷進步，人工智能領域近些年來實現了快速發展，與此相關的問題逐漸受到學界的普遍關注。面對人工智能，我們首先要解決“智能是什麼”、“人工智能是什麼”、“人工智能的基本實現手段是什麼”、“人工智能的終極目標是什麼”等基本問題，但時至今日依然沒有一個可調和不同 AI 學派之間分歧的說法。總體來看，“人工智能希望用計算機類比智慧過程和行為，並據此加深對大腦的功能和智慧本質的認識和理解，但不限於生物學或神經科學可解釋的意義。”<sup>1</sup>

約翰·塞爾將人工智能分為兩類：強人工智能與弱人工智能。強人工智能同弱人工智能的根本區別在於，如何理解程式化的計算機在解釋心靈過程中的作用。在弱人工智能看來，“計算機在心靈研究中的主要價值就是為我們提供了一個強有力的工具。例如，它能使我們以更嚴格的、更精確的方式對一些假設進行系統闡述和檢驗。”<sup>2</sup> 所以，從弱人工智能的角度來說，機器程序祇是用來解釋的工具。強人工智能則認為，“計算機不是祇研究心靈的工具，更確切地說，帶有正確程序的計算機確實被認為具有理解和其他認知狀態，在這個意義上，恰當編程的電腦其實就是一個心靈”<sup>3</sup>；“擁有心靈就是擁有程序”<sup>4</sup>。由此可以看出，在強人工智能這裡，計算機程序本身就是解釋對象——心靈。

那麼，究竟人工智能機器是否可以實現“強人工智能”，還是祇

---

<sup>1</sup> 董軍《人工智能哲學》，北京：科學出版社，2011年，第1頁。

<sup>2</sup> J. R. Searle, “Minds, Brains, and Programs,” *Behavioral and Brain Sciences* Vol. 3 (1980): 417-457.

<sup>3</sup> 同上。

<sup>4</sup> 約翰·塞爾《心靈的再發現》，北京：中國人民大學出版社，2005年，第168頁。

能在“弱人工智能”的範圍內徐徐圖之，其實不可避免的都要涉及到人工智能機器的“心識”究竟能發展到何種程度的問題，特別是機器意識的發展。我們知道大腦是一團皺巴巴、黏糊糊的物質，由互相交換電流訊號的神經細胞所組成的。但是，我們不得不思考這團灰色的物質是如何形成我們的“心識”的？換句話說，心智、意識等精神性的活動如何由又軟又黏的腦細胞這樣的物質性東西所產生？無論是人類腦科學研究還是人工智能機器研究都面臨這一問題，一直以來也有一些科學家嘗試對其進行解答。《成唯識論》中對“識”的認識以及對“心”與“心所”的劃分幾乎涵蓋了所有的心理現象和心理活動，當然也包括有意識的心理活動，這些理論可以為我們研究人工智能的“心識”問題提供理論研究方面的借鑒。

## 一、人類“心識”能力的唯識學解讀

人類的“心識”活動以神經活動為基礎，但神經活動作為一種物質性的活動自身無法意識到“心識”的產生，而且有些神經活動也無法直接產生“心識”。“根據現有的相關科學與哲學研究成果，人類意識的運行機制大致是這樣的：物質運動變化創生萬物，生物的生理活動支持著神經活動，神經活動湧現意識（有意識的心理活動），意識感受生理表現並指導意向性心智活動的實現，從而反觀認知萬物之理。”<sup>5</sup>一般地說，心理活動能力包括感覺、感知、認知、覺知、情感、行為、返觀等，其中的行為主要涉及與心理有關的意志、願望、

---

<sup>5</sup> 周昌樂《機器意識能走多遠：未來的人工智能哲學》，《人民論壇·學術前沿》，2016年第13期，第81-95頁。

情欲等，這些心理活動能力共同構成了人類的“心識”。唯識學所研究的物件就是“心識”問題。

《成唯識論》中玄奘法師為了傳唯識意，破除我執和法執，延續慣例將心理活動的主體分為“心”與“心所”兩大類。“能了別境，名為識故”，據此可知“識”的作用在於“了別”。“了”即是瞭解、認識，“別”即分別、辨別，“了別”就是認識功能。唯識學認為，“識”的產生，雖需要具備一定的條件，但根本上說，“識”起源於自己的種子，依賴“作意”轉變為現實活動。即，對“識”的形成來說，識的種子是因，根與境是緣。

“心識”所變現的“心”與“心所”，最早的劃分形成於部派佛教時期。一切有部的經書《阿毗達摩發智論》《大毗婆沙論》《雜阿毗曇心論》《阿毗達摩俱舍論》等中都有對“心”和“心所”的劃分。《成唯識論》中“心”、“心所”的劃分與世親《大乘百法名門論》中相同，將世間事物總結歸納為一百樣，分為五大類，包括心法、心所法、色法、不相應法和無為法，即所謂“五位百法”。通過將一切諸法分為五百位法，《成唯識論》建立了一個次第嚴謹的修證體系。《大乘百法明門論》卷一中對“五百位法”的具體說明如下：<sup>6</sup>

第一心法，略有八種。一眼識，二耳識，三鼻識，四舌識，五身識，六意識，七末那識，八阿賴耶識。

第二心所有法，略有五十一種，分為六位。一遍行有五，二別境有五，三善有十一，四煩惱有六，五隨煩惱有二十，六不定有四。

(一) 遍行五者：一作意，二觸，三受，四想，五思。

(二) 別境五者：一欲，二勝解，三念，四三摩地(可譯作“定”)，

<sup>6</sup> 《大正藏》編號 1614，第 31 冊第 855 頁，中欄第 20-29 行，下欄第 1-19 行。

五慧。

(三) 善十一者：一信，二精進，三慚，四愧，五無貪，六無嗔，七無癡，八輕安，九不放逸，十行舍，十一不害。

(四) 煩惱六者：一貪，二嗔，三慢，四無明，五疑，六不正見。

(五) 隨煩惱二十者：一忿，二恨，三惱，四覆，五誑，六諂，七驕，八害，九嫉，十慳，十一無慚，十二無悔，十三不信，十四懈怠，十五放逸，十六昏沉，十七掉舉，十八失念，十九不正知，二十散亂。

(六) 不定四者：一睡眠，二惡作，三尋，四伺。

第三色法，略有十一種。一眼，二耳，三鼻，四舌，五身，六色，七聲，八香，九味，十觸，十一法處所攝色。

第四心不相應行法，略有二十四種。一得，二命根，三眾同分，四異生性，五無想定，六滅盡定，七無想報，八名身，九句身，十文身，十一生，十二住，十三老，十四無常，十五流轉，十六定異，十七相應，十八勢速，十九次第，二十時，二十一方，二十二數，二十三和合性，二十四不和合性。

第五無為法者，略有六種。一虛空無為，二擇滅無為，三非擇滅無為，四不動滅無為，五受想滅無為，六真如無為。

從這“百法”可以看出，這五位都離不開識，識的自相是心法，識的相應是心所法，識所變是色法，識的分位是不相應行法，識的實性是無為法。由此可知，所謂“唯識”，就是說一切法不離識，除卻識之外的其他東西，祇不過是識的幻影而已。以上“百法”限於篇幅，這裡不能一一進行詳細說明，本文中我們祇探討“心法”和“心所有法”的內容。

心法，涵蓋三能變，初能變是作為第八識的阿賴耶識；第二能變是作為第七識的末那識；其餘的六識則是第三能變。第八識是根

本識，其他七識成為“轉識”。相較於第八識的眾多名稱，阿賴耶識、異熟識、一切種識是最重要的三個，分別對應自相、果相、因相這三種性狀。阿賴耶識是“非常非斷”的，它和一切種識一起表明此識能保藏眾生身心和物質世界等一切事物的種子，但此識是無形無象的。種子也祇是指此識“能遍任持世出世間諸種子”，即此識具有能生起現行事物的功能。異熟識表明此識是前世善惡業形成的果報。《成唯識論》中詳細論述了第八識的性質，並對第八識的存在進行了論證。

次第二能變，是識名末那。依彼轉緣彼，思量為性相。四煩惱常俱，謂我癡我見，並我慢我愛，及餘觸等俱。有覆無記攝，隨所生所系，阿羅漢滅定，出世道無有。<sup>7</sup>

第七識是末那識，其含義是“污染意”，它與第八識一樣也是一個超出人們經驗範圍的概念。末那識的生起以第八識為因緣，它的主要性質是“恒審思量”。“能審思量，名末那故。未轉依位，恒審思量所執我相；已轉依位，亦恒審思量無我相故。”<sup>8</sup>“第七識的‘自性’與‘行相’都是‘審思量’，即凡夫狀態的第七識始終明確清晰地思量第八識為自我，而成佛後的第七識始終明確清晰地思量無我的實相。”<sup>9</sup>第七識的“意”與第六識的“意識”是相區別的。末那識表示“意動是識”，其識之本身即名為“意”，而且末那識是第六識的根，即特有的依託對象。

次第三能變，差別有六種，了境為性相，善不善俱非。<sup>10</sup>

<sup>7</sup> 《大正藏》編號 1585，第 31 冊第 19 頁，中欄第 1-6 行。

<sup>8</sup> 《大正藏》編號 1585，第 31 冊第 22 頁，上欄第 22-23 行。

<sup>9</sup> 林國良《成唯識論直解》，上海：復旦大學出版社，2000 年，第 293 頁。

<sup>10</sup> 《大正藏》編號 1585，第 31 冊第 26 頁，上欄第 13-13 行。

第三能變識共包括六種不同的識，分別為眼識、耳識、鼻識、舌識、身識、意識，這六識的名稱是依據“根”建立的。如果依據“境”建立，六識就指色識、聲識、香識、味識、觸識、法識。但依據“境”建立的這一類名稱在成佛後就不再適用，因此，六識的名稱一般都是根據六根而建立。六識的本性及現行活動作用的落腳點在於“了境”，即認識辨別。

從現代心理學的角度看，“心法”中的前六識與心理學中的視覺、聽覺、嗅覺、味覺、觸覺以及意識大體相類似。但第七識末那識和第八識阿賴耶識是整體統一的心法，尚沒有現代認知科學中的概念可以與之相對應。這兩識主要強調自我意識，即對自我心識的審悟能力。

“心”的認知功能的真正實現離不開“心所”的幫助。“心所”，即“心所有法”，亦即“心數”，指依託心而生起的、屬於心的各種心理活動或心理機能。“心所”不僅可以幫助“心”完成認知活動，還可以幫助“心”實現情感、意志等其他功能。八識在時總會升起一定數量的“心所”伴隨其進行活動，沒有無“心所”伴隨的識，也無脫離識而獨立活動的“心所”。“心”祇能認識物件所具有的總體狀況，“心所”卻能從局部和整體兩個方面認識物件，而且還能幫助“心”完成認識。

按照《大乘百法明門論》的說法，心所有六類五十一種，即遍行心所五種、別境心所五種、善心所十一種、根本煩惱心所六種、隨煩惱心所二十種、不定心所四種。其中與心理活動關係密切的主要是“遍行心所”、“別境心所”和“不定心所”。

<sup>11</sup> 林國良《成唯識論直解·前言》，上海：復旦大學出版社2000年，第21頁。

“遍行心所”涉及一些基本的、普遍的心理活動或心理功能，它是在任何條件下都會伴隨著任何識生起的心所，是六類“心所”中最重要的一類。“遍行心所”包括五種，即觸、作意、受、想、思。“唯識學認為，識與根、境三者的和合能生起觸心所，而觸心所同時也加強了三者和合的力量，使識最終生起，即使識從種子狀態（僅僅是一種功能）過渡到現行活動的狀態。這也就是說，如果沒有觸心所，則根、境、識三者就不能會合，識也就不能最終生起。”<sup>11</sup>前五識的生起可以看出，根是眼、耳、鼻、舌、身等五種感覺機制，境是色、聲、香、味、觸等感覺物件。“作意心所是使識從種子狀態警覺而起，將識引向認識物件，從而使識、根、境三者和合。所以，識的生起必定依賴於作意，作意使識從種子轉變為現行。”<sup>12</sup>“遍行心所”中“觸心所”和“作意心所”是唯識學所獨有的描述心理活動的概念，而受、想、思則與六識結合，形成了心理活動中的情感活動、認知活動與意志活動。

“別境心所”是一些較為基本的心理功能，不具有“遍行心所”那樣的普遍性，祇是在一些特定的條件下纔產生。在“別境心所”中，“欲心所”相當於興趣，“勝心所”相當於理解、“念心所”相當於記憶，“定心所”相當於注意力高度集中的狀態，“慧心所”相當於判斷和推理等思維活動。“不定心所”指悔、眠與尋、伺，兩類各兩種。其中，“悔心所”指懊悔的心理，“眠心所”指睡眠狀態，“尋心所”與“伺心所”也相當於判斷與推理等思維活動。但與“別境心所”中的“慧心所”不同，“尋心所”和“伺心所”是對現前事物的判斷、推理等思維活動，而“慧心所”則是對非現前事物的判斷、推理。

---

<sup>12</sup> 同上，第22頁。

“善心所”共有十一種，分別為信、慚、愧、三善根中的無貪、無瞋、無癡、勤、安、不放逸、行舍以及不害等。“煩惱心所”則是指貪、瞋、癡、慢、疑、惡見這幾種。“隨煩惱心所”包括大、中、小三種，它因是隨從根本煩惱而產生的煩惱得名。小隨煩惱主要包括忿、恨、覆、惱、害、嫉、慳、誑、諂、諂、驕十種。中隨煩惱包括無慚、無愧兩種。大隨煩惱包括掉舉、昏沉、不信、懈怠、放逸、失念、散亂、不正知等八種。概括一下，“隨煩惱心所”總共二十種。正是因為“善心所”、“煩惱心所”及“隨煩惱心所”的活動，心識纔具有了善惡等倫理性質。

通過對《成唯識論》中“心”與“心所”的分析，唯識學對人類心識結構的解析已經形成了一個較為完整的體系，充分展現了心識活動的生起條件、生起過程及生成狀態。

## 二、人工智能“心識”的困境

目前人工智能“心識”能力方面的研究，困難之處就在於機器心智能力特別是機器意識的研究，也就是如何對有意識的人類心理活動能力進行計算表徵，並完成某種可以看作與此相關的計算任務。心理活動能力的計算表徵可以看作是人工智能的理論限度。這其中無意向心理活動的表徵和去意向性心理活動的表徵問題是目前計算理論與方法無法解決的問題。因此，“機器實現的人類意識能力不可能超越意向性心識的範圍”。<sup>13</sup>

從唯識學的心識結構角度來說，人工智能機器“心識”難以變現

<sup>13</sup> 周昌樂《機器意識能走多遠》(前揭)。

實在的“心”和“心所”。認識活動的升起需要感知物件，我們最主要的認識物件是物質世界，《成唯識論》中與之相應的說法即是“器世間”。唯識學認為，“所言‘處’者，謂異熟識由共相種成熟力故，變似色等器世間相，即外大種即所造色。雖諸有情所變各別，而相相似，處所無異，如眾燈明，各遍似一。”<sup>14</sup> 物質世界是由眾生的第八識變現的。對此，玄奘法師在《成唯識論》中有更進一步的討論，“然所變土，本為色身依持受用，故若於身，可有持用，便變為彼。由是設生他方自地，彼識亦得變為此土。故器世間將壞初成，雖無有情，而亦現有。”<sup>15</sup> 這樣，色界和欲界各地的眾生各自變現各自的物質世界。除了物質世界，第八識作為一切種識，還有第二類認識物件：種子與具有感覺機制的身體。“諸種子者，謂異熟識所持一切有漏法種，此識性攝，故是所緣。……有根身者，謂異熟識不共相種成熟力故，變似色根及根依處，即內大種及所造色。”<sup>16</sup> 有漏種子是第八識的認取物件，它們由第八識變現，與第八識具有同一主體。第八識的另一認取物件是具有感覺機制的身體。“粗略地說，眾生的身體是由眾生各自的不共相種子生起的，非感覺機制的那部分身體仍是共相種子生起的，所以後者可為眾生共同受用。……因此，眾生的身體既然也能為眾人共同受用，所以應該也是由本人與其他眾生共同變現。”<sup>17</sup>

“心識”得以生起的基礎在於第八識，而對於人工智能來說，如果現代認知科學中沒有與第八識相對應的概念，我們也就很難找到

---

<sup>14</sup> 林國良《成唯識論直解》，上海：復旦大學出版社，2000年，第157頁。

<sup>15</sup> 同上，第157-158頁。

<sup>16</sup> 同上，第160-161頁。

<sup>17</sup> 同上，第161頁。

一種可以在人工智能機器中實現第八識的技術手段。很明顯，人工智能也很難像人一樣擁有具有感覺機制的身體，更不用說這具身體得能由本人與其他眾生共同變現。“皮之不存，毛將焉附”，此種意義上來說，人工智能機器的“心識”難以變現出“心”和“心所”。同理，分析“作意心所”、“根”、“境”也會發現人工智能機器難以對此進行技術性表徵。因此，“心識”很難在人工智能中自然而然變現。

從形式系統角度看，哥德爾定理表明，基於形式化邏輯方法的機器，是不可能擁有像人類那樣的整體性智能表現的。這也是形式化方法難以克服的局限性。也就是說，“心”和“心所”所表現的人類心智現象是不可能完整地由形式化方法加以描述的，因而基於形式化邏輯的人工智能也無法真正刻畫實現這種類人的整體性智能。

哥德爾在其《論數學原理中的形式不可判定命題及其關係》的論文中提出了哥德爾不完全性定理，指出了邏輯形式系統不可克服的局限性。<sup>18</sup>

哥德爾不完全性定理 1：存在一個 PA 句子  $p$ ，使得：如果 PA 是一致的，則  $p$  在 PA 中不可證；如果 PA 是  $\omega$ -一致的，則  $\neg p$  在 PA 中不可證。因此 PA 是不完全的。

哥德爾不完全性定理 2：如果 PA 是一致的，那麼 PA 的一致性不能在 PA 內部證明。

哥德爾不完全性定理的意義之一就在於表明確實存在不可計算（證明）的問題，而且大量存在的不可計算問題對於人類自身也是不可計算的，如圖靈停機問題。邏輯形式系統的局限性直接的影響在於，我們無法給出“心”和“心所”這樣的心理活動能力一致性的

<sup>18</sup> 周昌樂《心腦計算舉要》，北京：清華大學出版社，2003年，第205頁。

形式描述。“在人類的意識現象中，存在著意識的自反映心理現象：我們的意識活動是自明性的。從邏輯上講，如果一個系統允許自涉，那麼該系統一定是不一致的，也就是說無法對該系統給出一致性的形式化描述。其實，人類的心理活動本來就是建立在神經集群活動的自組織湧現機制之上的。因此，出現意識的自明性現象是必然的。”<sup>19</sup> 我們所有的心理活動能力顯然不能為機器所操縱。機器可能在某些方面實現或者超越人類，如記憶、大規模計算等方面，但很難把心理體驗通過計算方法描述出來，因此靠邏輯機器的人工智能是不可能擁有人類全部的心理活動能力的，正如《成唯識論》中“心”與“心所”所分析的。

“心識”的生起有所“緣”和所“依”。所依據的“緣”和“依”不同，產生的“心識”也各有不同。因此，我們人類的“心識”是有創造性和靈活可變性的。人工智能的運行需要依賴於預先編定的程序，這樣來看，人工智能機器如何來體現創造性和靈活性呢？預先設定的程序使機器難以實現靈活性，更不用說具有突破性的創造能力。程序越是符合規則，人工智能就越容易掌控，複雜的程序祇不過是簡單規則的反覆運算產生。但是，人類社會新事物產生的基礎在於人面對問題、面對事物的靈活性和創造性。我們常常會出錯，有些創造性的錯誤造就了新物種、新事物的產生或者是文化的多樣性。而人工智能機器祇需要按照預先設定的程序運行，永遠不可能出錯，也就缺少由出錯可能帶來的創造性成果。

當然，世事無絕對，按照預先設定程序運行的人工智能機器不會出錯，但給機器編定程序的人是具有靈活性和創造性的，是會出錯

---

<sup>19</sup> 周昌樂《機器意識能走多遠》(前揭)。

的。也許某一天，一個錯誤的程序經過運行反倒可以解決人工智能在“心識”發展中所遇到的瓶頸。這時候，怎麼判斷建構的人工智能機器是具有人類“心識”能力的，成功的標準是什麼？英國數學家圖靈在其《計算機器與心智》的文章中就“機器能不能思維”這一命題，給出了一種測驗機器心智是否達到人類水準的測驗，即圖靈測驗。這樣，構建具有人類“心識”能力的機器是否成功，很重要的標準就在於看其能否通過圖靈測驗，內部的心智水準可以通過外部行為表現得以判斷。

所謂圖靈測驗，是指在兩個隔離的房間分別關有一個人和一臺機器，然後通過向人和機器同時提問來判斷誰是人，誰是機器。圖靈認為，如果通過“巧問”的設計，最終能夠正確地將人與機器區分開來，那麼說明機器不同於人，否則就說明機器與人在心智上沒有區別。當我們對同一問題重複提問的時候，人工智能機器很容易會暴露自己的缺陷，它缺乏機動的不可預見性反應能力，祇會按照預先設定的程式進行機械式回答。可是人在面對同一問題的多次反復提問時，往往會將自己的情緒、好惡等加入到問題回答的行為中從而進行表達，即使是同一個人針對同一個問題，答案也是不可預見的。此後，針對機器能否真正理解符號所代表的意義，塞爾又設計了“中文屋”測試。

根據上文的分析不難看出，《成唯識論》中各種“心”和“心所法”生起的基礎第八識阿賴耶識是人工智能機器很難具有的。人工智能機器是邏輯的，哥德爾不完全性定理指出基於形式邏輯的計算方法存在著嚴重局限性，也無法表徵“心”和“心所”所表現的人類心智現象。預先設定的程式也不具備人類的靈活性和創造性，無法保障人工智能機器向具備人類“心識”進行突破。因此，人工

智能可能在某些功能方面超越人類獲得突破，但它難以真正全面類比人類的全部心智能力，不可能真正實現人類“心識”。

### 三、人工智能“心識”發展的思考

《成唯識論》中“心”和“心所”的最終目標是為了破除“我執”和“法執”，從而獲得解脫。二執均有俱生（與生俱來）和分別（由思維而產生）兩種。分別的我執和法執是由第六識引起的。俱生的二執既有由第六識產生的又有由第七識產生的，而由第七識產生的二執是世世代代與生俱來、連續不斷的。那麼涉及到有“心識”的人工智能其最終目標是什麼呢？開展有修行悟道意識的機器人研究是否有可能？

探討具有修行悟道意識的人工智能機器人研究的可能性問題，主要涉及自我意識問題，一方面是要具備關於自我的意識——“我執”，另一方面要有內省反思意識，相當於“悟”的體驗。但是，從《成唯識論》中“心”和“心所”的角度可以看出，主體對自身本體感知活動的內省反思這種“自我”意識是需要破除的，我們的心理內部並不存在所謂的“自我”實體。所以，機器自我意識的人工智能研究方向是與唯識心法相背離的。

所羅門曾對人工智能進行量級劃分，分別記為  $T_i$  ( $i=1, 1a, 2, 3, 4, 5$ )，具體如下：<sup>20</sup>

$T_1$ ：認為每一種 UAI（尚未發現的智能算法）都具有心智能力，有待發現的這些 UAI 由純資料和算法構成，而無需考慮時間、豐富

---

<sup>20</sup> 周昌樂《心腦計算舉要》，北京：清華大學出版社，2003年，第212-213頁。

的執行機制和意義因素，也就是說，僅靠抽象固定的結構就可以產生心智。

T1a：對於 T1 作了一點擴展，加上時間因素（積累學習），認為結合了時間因素（積累學習），認為結合了時間因素的每個 UAI 具有心智能力。

T2：在 T1a 的基礎上再引入對程序控制的種種執行機制，認為這樣就可以產生心智了。

T3：認為單個算法不足以實現智能問題，必須考慮多算法的虛擬並行機制（又分為連續環境刺激，時間共用連續處理機和適當的計算機網絡三種情況），祇有採用這種虛擬並行主義機制，纔能產生心智能力。

T4：將上述的虛擬改為物理，也就是採用物理並行主義機制並行主義機制，就能夠產生心智能力。

T5：要擁有心智能力，至少部分子系統需要具有超計算的能力，比如採用物理、化學和生物等自然機制。

這六條的觀點很明顯是逐漸減弱的，概括了人工智能心智能力發展的不同量級。從中不難發現，人工智能的發展實際上是朝著人類智慧逼近的方向發展的。但是我們又不能真正讓人工智能“心識”的發展達到或超越人工智能，具體原因恐怕要從倫理學角度來進行探討。

那麼，人工智能的發展終將歸於何處？很顯然，對於這個問題的回答，一方面需要依賴未來邏輯計算方法的突破；另一方面，我們不得不思考不斷完善、快速發展人工智能是否真正有必要。最近，隨著量子計算、腦機融合和類腦構造等先進技術的發展，為人工智能機器的“智能化”實現提供了一種全新的途徑，於是人工智能的前景又出現了新的光明。人工製造類人機器以及腦機合成體，除了技

術之外，我們還不斷地在嘗試讓機器像人一樣成為智慧生命體。但是要想實現智慧生命體又絕不僅僅是技術問題，它還涉及一系列其他的問題。從唯識學意義上來說，能達到人類“心識”的人工智慧生命是不可能的。

另外，我們還不得不面對一個問題，如果機器真的擁有了類人智慧，那我們進行人工智能探索的意義何在？是為了有一天能夠製造出取代自己的“人造物”嗎？這種不斷完善和發展的人工智能是否真正有必要呢？會不會最終促使人類的奴化，甚至走向消亡？這似乎又陷入了人類在不斷完善自身、追求完美過程中的一個“怪圈”。

人工智能“心識”問題的研究無論是理論方面，還是應用方面都是任重而道遠的。一方面，人工智能的發展努力逼近人類整體性智能，在某些方面甚至可以超越人類；但另一方面，我們又擔心人工智能的高度快速發展會對人類社會發展造成不可逆的不利影響。《成唯識論》中的“心識”分析有助於更全面地認識人類有意識的心理活動，推動人工智能“心識”的理論研究，也有助於我們對人工智能進行思考。

## 參考文獻

### 東亞語研究

周昌樂《心腦計算舉要》，北京：清華大學出版社，2003年。

周昌樂《機器意識能走多遠：未來的人工智能哲學》，《人民論壇·學術前沿》2016年第13期，第8-95頁。

林國良《成唯識論直解》，上海：復旦大學出版社，2000年。

約翰·塞爾《心靈的再發現》，北京：中國人民大學出版社，2005年。

董軍《人工智能哲学》，北京：科學出版社，2011年版。

### 西文研究

Searle, J. R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences* Vol. 3 (1980): 417–457.

