

活動 21

和電腦的對話 — 圖靈測試

活動摘要

這個活動的主要目的是激發學生們的討論，看看電腦是否能擁有「智慧」，或是在未來是否可能發生。根據一名先驅電腦科學家觀察，一個人對於人工智慧存在的認知，顯示出當時的科技進展，但也血淋淋的呈現了許多對人工智慧的誤解。

課程銜接

- 科技：科技系統。瞭解科技系統是由符號語言工具所呈現，並瞭解在科技系統中的黑盒子所扮演的角色。

習得技能

- 提出問題
- 歸納理解

適合年齡

- 7 歲以上

所需素材

- 一份「活動學習單：圖靈測試問題」（第 243 頁）的複本。可以一組學生看一份，或是用投影機顯示出來
- 一份「活動學習單：圖靈測試問題的答案」（第 244 頁）

與電腦對話

活動介紹



這項活動以遊戲的方式呈現所謂的圖靈測試。讓學生透過提問以及分析得到的答案來嘗試分辨人類和電腦的不同。遊戲的規則如下：

這遊戲將會有四個角色：分別是小帥、小哥、小美、小麗。老師負責控制流程，引導大家提出問題。小帥和小哥是中間人，不參與問題的回答；而小美小麗則要負責回答問題。教室中的其他人是聽眾。小美必須用人類的回答方式作答，而小麗則要假裝自己是電腦來作答。教室中的其他學生要分辨出到底誰扮演人類，而誰扮演電腦。小帥和小哥這時的主要任務是確保遊戲公平進行：他們要負責把問題帶給小美和小麗，但同時要注意不可以洩漏小美和小麗所扮演的角色。角色確定後，就將小美、小麗和聽眾，分別帶到不同的房間之內。

而關於提問及答題的機制如下。小帥把問題帶給小美，而小哥則把問題帶給小麗。最後兩人再將他們的答案告訴聽眾（實際上聽眾們並不知道是誰將問題傳達給誰）。這個中間人的角色，就是用來保證聽眾們不會看到小美以及小麗回答問題時的狀況。

在活動開始以前，請先選出扮演這四個角色的學生，並告訴他們遊戲規則。有個重點是，身為中間人的小帥和小哥，回來說出答案時必須完全掩飾所有能透露出是小美或小麗的資訊，例如不能說溜嘴告訴全班：「剛剛小美說她的答案是 XXX…」小美的回答由自己想，要簡短、精確與誠實。另一方面，小麗身為扮演電腦的一方，回答問題時必須遵照「學習活動單：圖靈測試問題的答案」上的答案來做回答。

小帥和小哥要準備紙和筆，因為有些問題不容易記住。

1. 在遊戲開始以前，先蒐集學生對於以下這個問題的回答以及想法：你覺得現在的電腦算不算是有智慧的？還是你覺得以後的電腦才有可能實現呢？
2. 向學生介紹此次活動的目的，在於分辨電腦以及人腦的差別。如果班上最後沒有人能夠分辨出兩者的差異，那電腦就算是通過考驗。向學生們解釋小帥和小哥的工作，主要是扮演問題和答案傳送的角色，但是不能說破誰是負責傳達誰的訊息。我們必須讓學生分辨出，到底誰是扮演電腦的那一位。
3. 向學生展示「學習活動單：圖靈測試問題」中的問題。建議可以每兩人共看一份，或是用投影機顯示。

問同學們想要先問哪一個問題。當他們決定以後，再詢問他們為什麼會選擇這個問題，還有他們為什麼會覺得個問題可以分辨出電腦和人腦。這一個推理的過程是整堂課程最重要的一環，因為它能讓學生去思考什麼問題是人類可以回答，但電腦卻不見得辦得到的。

小帥和小哥接著傳遞這些問題和答案。班上同學必須開始討論哪一個答案可能是電腦回答的。

重複這些步驟，多問幾個問題，當班上同學開始發現到哪一位扮演電腦時，就可以暫時停止。但是如果老師發現班上同學很快就知道誰是電腦，誰是人類的話，可以讓小帥和小哥隨機決定，例如擲硬幣或猜拳，然後重新安排傳話的對象是小美還是小麗。

小麗所拿到的圖靈測試問題的答案，看起來就像是由真正「有智慧」的電腦所回答出來的。有些答案會讓人很快知道是電腦回答的，例如：根號二等於多少？回答得出來的人，大部份頂多回答到 1.414，而不會回答到小數點後 20 位。另外，有些問題的答案放在一起看時，可能也會透露出電腦的真實身分。例如：「你喜歡 嗎？」這類問題本身很似是而非，但在問了幾次之後，就可以發現利用問題產生答案的一種簡單模式。而有時候，聽眾也可以發現對方在回答時誤解了題目的意思，他們就可以合理的推斷這樣回答的應該是人類。

大部份的答案都很平淡 — 但安全。不過，繼續追問下去就可能發現電腦其實並不真正了解問題的主題。用「我不知道」來作答，對電腦來說相對安全，也容易讓人覺得它是人類，因為回答「不知道」對於一個人類來說是很常見的。但如果電腦太常這樣回答的時候，或是對一個很簡單的問題這樣回答時，仍然有機會被察覺。

因為電腦的目的是要讓對方認為自己是人類，因此有些答案可能是故意設計來誤導大家的。例如遇到數學問題時，刻意拖延回答時間，或是故意回答錯。這類的問題與答案可以讓學生們好好討論。

活動學習單：圖靈測試問題

從這份清單中選出問題來問，試圖辨識出回答者是電腦還是人類。

1. Bart Simpson 的小妹妹叫什麼名字？
2. 你對 Roald Dahl 的看法是什麼？
3. 你是個電腦嗎？
4. 在 3, 6, 9, 12, 15 這個序列中，下一個數字是什麼？
5. 你對核子武器的看法是什麼？
6. 2×78 等於多少？
7. 根號 2 等於多少？
8. 34957 加 70764 等於多少？
9. 你喜歡學校嗎？
10. 你喜歡跳舞嗎？
11. 今天星期幾？
12. 現在幾點了？
13. 二月在閏年裡有幾天？
14. 一周有幾天？
15. 哪個國家的國旗是白色的背景，上面有個紅色圓圈？
16. 你喜歡看書嗎？
17. 你喜歡吃什麼食物？

活動學習單：圖靈測試問題的答案

1. Bart Simpson 的小妹妹叫什麼名字？

我不記得了。

2. 你對 Roald Dahl 的看法是什麼？

他寫了一些很有趣的書。

3. 你是個電腦嗎？

那你是個電腦嗎？

4. 在 3, 6, 9, 12, 15 這個序列中，下一個數字是什麼？

18

5. 你對核子武器的看法是什麼？

核子武器非常的危險，不應該使用。

6. 2×78 等於多少？

166 (這是故意寫錯的！)

7. 根號 2 等於多少？

1.41421356237309504878

8. 34957 加 70764 等於多少？

(在算出答案之前，請等待 20 秒 ...) 105621

9. 你喜歡學校嗎？

是的，我喜歡學校。

10. 你喜歡跳舞嗎？

是的，我喜歡跳舞。

11. 今天星期幾？

告訴他今天是星期幾。

12. 現在幾點了？

給他一個正確的時間。

13. 二月在閏年裡有幾天？

2000 年和 2004 年是閏年。(這是故意寫錯的！)

14. 一周有幾天？

七天。

15. 哪個國家的國旗是白色的背景，上面有個紅色圓圈？

我不知道。

16. 你喜歡看書嗎？

是的，我喜歡看書。

17. 你喜歡吃什麼食物？

我肚子不餓，謝謝。

活動變化與延伸

這個遊戲最少可以用三個人來玩，此時讓小帥也同時扮演小哥和小麗的角色。小帥拿問題給小美，並記錄小美的回答，同時也記錄「圖靈測試問題的答案」中該問題的答案。小帥回覆這兩個答案，並用字母 A 和字母 B 來區別這兩個答案的來源。

為了考量電腦是否能在這場問答中模擬人類，跟班上同學討論要回答每個圖靈測試的問題，需要哪些知識。學生可以建議其他他們想問的問題，並且討論任何他們可能設想的答案。這需要一點想像，因為預測這場對話的走向是不可能的。

為了說明，這裡有兩個示範的對話。前面這個範例說明了一台電腦「真實的」可以回答正確的問題，而下面的範例則展現了這個討論可以拓展地多廣泛，並且展現了電腦可能需要引用的廣泛的知識。

網站上有個廣泛可用系統叫做 ”Eliza” (是一種「聊天機器人」(chatbot)，它是一個你可以打字聊天的對象。Eliza 模擬一個與心理治療師的會談，並且可以藉由簡單的規則產生出色且聰明的對話。一些範例會談會在下面討論。學生可以試試 Eliza，或者其他聊天機器人 (像是 Siri !)，不過也要留意，有些機器人是用不適合學校學生的語言及主題訓練出來的。

問題：請以 Forth Bridge 為主題寫給我一個十四行詩。

回答：別看我，我從來不會寫詩。

問題：計算 34957 加上 70764

回答：(暫停個 30 秒 ...) 105621。

問題：你會下西洋棋嗎？

回答：會。

問題：我的國王現在在 K1 上，而且我沒有其他棋子了。你只有國王在 K6 和城堡在 R1 上。輪到你了。

回答：(經過 15 秒的暫停 ...) 城堡移到 R8，將死你。

問題：一首十四行的第一行寫道「我該將汝與一個夏日比較嗎？」若寫「一個春日」會不會一樣甚至更好呢？

回答：這不符合詩的韻律。

問題：那如果是「一個冬日」呢？這完全符合韻律。

回答：是，但沒有人想被拿來比喻成一個冬日。

問題：你認為 Mr. Pickwick 會讓你想起聖誕節嗎？

回答：某種程度上如此。

問題：然而聖誕節是個冬天的日子，而且我不認為 Pickwick 先生會介意這個比喻。

回答：我想你不夠嚴謹，一個冬天的日子代表的是一個尋常的冬日，而不是像聖誕節這種特殊的日子。

這個活動在說什麼？

好幾世紀以來，哲學家一直在爭辯一台機器是否能模擬出人類的智慧，以及反過來問，人腦的能力是不是根本無法超過一個寫得好的電腦程式？這個問題看法兩極。有些人覺得這個想法荒謬、瘋狂，甚至是褻瀆的，而另一些人則認為人工智慧是不可避免的，最終我們會開發出跟我們一樣聰明的機器。（而無數的科幻小說作者都指出，如果電腦的智慧超過我們，則他們會展開革命——自己可以建造出比他們更聰明的機器。）人工智慧（AI）研究學者一直被批評，他們所謂的崇高目標只是用來從政府單位吸引研究資金，而政府單位的目標則是要建造出自動化的戰爭機器。而這些研究學者們則譴責這種看法是十九世紀反工業革命的盧德運動的翻版，並指出在生活週遭多一點智慧對社會是有極大好處的。另外還有一種比較平衡一點的觀點是，人工智慧既不荒謬也不必然會發生：既然目前沒有任何電腦程式展現出「智慧」，這些問題目前爭辯也不會有任何結論。

這些關於人工智慧的爭辯，取決於對「智慧」的定義。已經有許多定義被提出與辯論。其中一個最有趣的提案是在 1940 年晚期由艾倫圖靈（Alan Turing，一位著名的英國數學家，也是二戰期間破解德軍密碼的靈魂人物與一名長跑好手）所提出。他提出一種「思想實驗」。圖靈的想法是不去定義何謂「智慧」，而是提出一個情境讓電腦去展示他們的智慧。圖靈所提出的情境就類似這個活動中所玩的，本質上也是由一名提問者，與一個人還有一台電腦透過電子打字機（在 1940 年代是很先進的技術！）來互動。如果提問者分辨不出誰是電腦誰是人類，那麼這台電腦就通過了「圖靈測試」，也就是人工智慧的測試。使用電子打字機也是為了避免提問者經由電腦的物理特徵，或是聲調等辨別。這樣一來，電腦不需要在外觀、聲音、觸感、甚至味道上偽裝，因為這些感觀跟我們所謂的「智慧」比較沒有關係。

最原始的圖靈測試跟我們的活動有一點不一樣。他所提議的最原始的測試，是一名男人與一名女人被詢問，而提問者要決定兩名回答者的性別。男人的目標是要讓提問者相信他其實是回答的另一名女人，而女人的目標則是要讓提問者相信她才是貨真價實的。接著圖靈想像（因為這只是一個思想實驗），用一台電腦來取代其中一人，看看電腦是否能假扮人類通過這個「模仿遊戲」。我們把這個想法轉換成在教室裡玩的活動。之所以不讓學生猜測回答者的性別，其中一個原因是學生所提出來的問題可能不合適，而且這樣也隱含了對性別的刻板印象，更別提回答者必須用欺騙的方式回答。

模仿智慧是一件困難的工作。如果今天把角色反過來，由人類設法通過測試讓提問者以為自己是電腦，那根本很難通過：因為只要一問「 123456×789012 等於多少」就掛了。

不過，要讓電腦取得對話能力的皮毛其實出乎意料地簡單。在 1960 年早期所開發出來的知名程式 Eliza，就在與使用者的對話中假扮成一名非指導性的心理治療師。右方是某人與這個系統間的對話範例。一些模仿自然對話的原則包括：

1. 產生一些罐頭回應（例如在回應「他們總是如何如何」時，回答「可以給我一個例子嗎？」；遇到「他說如何如何」時回答「聽到這些我感到很遺憾。」等等）
2. 重複使用者的敘述（例如「我的男友叫我來這裡。」「喔，妳的男友叫妳來這裡。」）
3. 辨識出關鍵字（聽到「媽媽」則回應「談談你的家人吧！」）
4. 一些常用的片語或回答（如果 XXXX 的話，對你來說有意義嗎？）
5. 從前面的對話中找出問題來詢問（如果遇到 XXXX，你打算怎麼辦？）

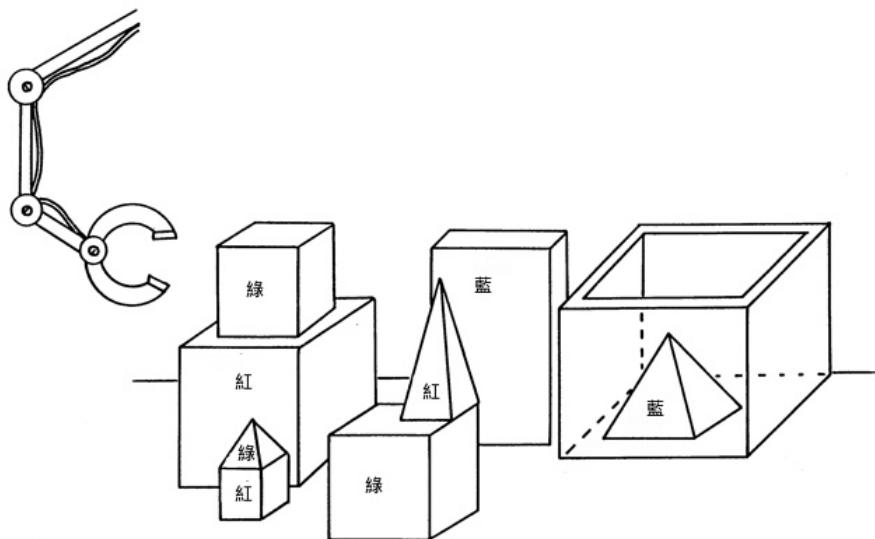
當然，這樣的對話不代表就通過了圖靈測試，因為提問的目的並非要分辨出回答的是人還是機器。如果真的刻意要分辨，那這個系統其實撐不了太久。

另外有一個廣為流傳的故事，可能是杜撰的，但現在已經成為一個人工智慧領域的傳說。有一名電腦銷售副總裁宣稱他偶然連到一個終端機。通常該終端機是直接連線到某特定人的家中，但是那一次剛好正在跑一個「心理治療師」程式。他們間的對話顯示在右方。雖然證據顯示這名副總裁的確有參與這個事件，但很明顯的整個情境本來就讓他認為他是在跟人談話。如果相反地，他一開始就起疑而刻意辨識，那應該很快就會發現。

人：唉，男人們都一樣。
電腦：妳是指哪方面？
人：男人們總是不斷用許多問題來困擾我們。
電腦：可以給我一個例子嗎？
人：噢，就像我男朋友要我來這裡。
電腦：喔，妳男朋友要妳來這裡。
人：他說我大部份時間都太憂鬱了。
電腦：噢，聽到妳這麼憂鬱我感到很遺憾。
人：那是真的，我很不快樂。
電腦：那妳認為到這裡來能讓妳快樂一點嗎？
人：可以肯定的是，我需要幫忙。
電腦：跟我談談妳的家人吧。
...
電腦：講到妳父親妳會聯想到什麼？
人：家暴。
電腦：這是為什麼妳男朋友要妳來這裡的原因嗎？
....

副總裁：請問今天早上我可以使用一下你的電腦嗎？
終端機：為什麼這樣問？
副總裁：這樣我可能可以多賣出我們的電信產品。
終端機：你為什麼這麼不確定？
副總裁：我這有些潛在的客戶，從來沒看過這個系統的展示。
終端機：這對你很重要嗎？
副總裁：當然啊！
終端機：你看起來非常積極。
副總裁：喂，這太扯了吧！
終端機：你說這太扯了吧是什麼意思？
副總裁：有任何理由我現在不能使用電腦嗎？
終端機：跟我談談你說的潛在客戶吧。
副總裁：請你現在撥電話給我，號碼是 491-1850
(因為副總裁沒有在尾端輸入標點，所以電腦沒有回答。副總裁很生氣，所以他打電話給負責那個終端機的人的家中。)
副總裁：你剛剛為什麼跟我講那些有的沒的，就是不讓我用電腦？
人：我？我剛剛有跟你在談話嗎？

另一個看來可以做智慧對話的系統叫做 "SHRDLU"，在 60 年代晚期開發。它可以接受指令，並操作在學生桌上的積木塊。事實上，這些積木塊與桌子只是在電腦螢幕上模擬的圖片而已，但要真的把它轉成真實的桌子與積木，用真的機器人來做也不是什麼難事（原則上啦）。下面這張圖就是機器人的小小世界。



SHRDLU 接續對話出奇地流利，下一頁有它跟人的對話內容。這對話內容展示了它可以遵照指令，分辨出在說的是哪個積木，也可以接受新的詞（尖塔）。

人：夾起大的紅色積木。

機器人：OK.

(系統理解指令時會回覆 "OK."。不過為了夾起大的紅色積木，它必須先找個空間，把上方的綠色積木移走。)

人：把三角錐夾起來。

機器人：抱歉，我不知道你說的是哪一個三角錐。

(總共有三個三角錐，所以機器人不做任何動作。)

人：找出哪個積木比你手上的那個要高，並把那塊積木放入盒中。

(系統必須理解指令中的一些詞彙，例如「它」是指哪個積木，來有「你手上的那個」等，並用消去法推論它手上的是哪一個積木。)

機器人：我假設你說的「它」是指比我手上的積木要高的那一個。

(持續印出資訊)

機器人：OK.

(把紅色積木放下並夾起較高的藍色積木。)

人：盒子裡有什麼？

機器人：藍色三角錐與藍色積木。

人：三角錐放在哪裡？

(雖然系統前面回應過「我不知道你說的是哪一個三角錐。」但是這次它知道指令中指的一定是它剛才提到的那個。)

機器人：在盒子中。

人：做一個尖塔 ...

機器人：抱歉，我不知道「尖塔」是什麼。

(系統遇到它不認得的詞彙時就立即發出回應。)

人：「尖塔」是用兩個綠色積木跟一個三角錐堆起來的。

(引號表示人正在定義一個新詞彙。)

機器人：我懂了。

人：目前有任何尖塔嗎？

機器人：沒有。

(機器人用新的定義來辨認何謂尖塔。)

人：做一個尖塔。

機器人：OK.

(人給它的描述已經足以定義並瞭解如何辨認並製作一個尖塔。)

不過，這個系統其實很脆弱。它展示的人工智慧美好景象其實很容易破滅。

右邊的對話是上面對話的延續—好啦，這不是真的去跟系統對話讓系統崩潰，不過這類對話的確很可能會發生。雖然它可以定義新的名詞（尖塔），但仍然只限於非常特定種類，因為這系統的「智慧」本來就很有限。最終，它還是得承認它甚至不知道自己用的字是什麼意思（「抱歉，我不知道『抱歉』是什麼。」）！

現在，讓我們換個方式來接續對話。這完全是幻想出來的。雖然對話很短，但這些對話仍然展示出世界知識的豐富。

1. 「機器人」知道已經有水槍的情況下，對方的水槍沒什麼價值。
2. 再來，它期望對方瞭解，因為它沒有繼續解釋拒絕接受的原因。
3. 它發覺對方很想要一塊積木來建個重要的東西。
4. 接著，它察覺到積木的數量很重要。
5. 它試著用青蛙這個玩笑去軟化對方。
6. 它表示已經準備好要討價還價了

人：我用一支水槍跟你交易大的紅色積木。

機器人：抱歉，我不知道「交易」是什麼。

人：「交易」是指自由交換物品。

機器人：抱歉，我不知道「自由」是什麼。

人：「自由」就是不被強迫做事情，任何事都是出於自願。

機器人：抱歉，我不知道「做事情」是什麼。

人：「做事情」就是人們有目的地做，而不是出於意外。

機器人：抱歉，我不知道「人們」是什麼。

人：抱歉，我以為你夠聰明。

機器人：抱歉，我不知道「抱歉」是什麼。

人：我用一支水槍跟你交易大的紅色積木。

機器人：我已經有一支水槍了，但我會給你兩個積木與一個三角錐，讓你可以跟你那隻滑溜溜的老青蛙玩。

當然，這種程度的對話藝術已經遠遠超過現在電腦的能力了！

在右方的故事是由另一個 80 年代後期的電腦所產生。（譯註：原文後附）雖然它很容易看出是電腦產出的故事，因為它所用的句子太簡潔太制式了，但是經過一番調整，加上許多細節之後就會不一樣了。有趣的不是這個故事表面，而是它可以體現出整個故事情節。雖然距離真正人類作者寫出的情節還有很長一段路，但它似乎已經抓住一些衝突的人性元素。這幾年已經開始出現不少自動產生文章的系統，例如網路上常見的那些內容農場，或是股票分析等等。要評估這些系統，要看有多少素材是標準的樣式，又有多少情節是有創意的。

從前，有一個名叫杜魯門的北極燕鷗。杜魯門無家可歸。杜魯門需要一個窩。他飛到岸上。杜魯門在找一些樹枝。杜魯門沒有發現樹枝。他飛到苔原。他遇到了一個名為賀拉斯的北極熊。杜魯門問賀拉斯哪裡有一些樹枝。賀拉斯隱瞞了樹枝。賀拉斯告訴杜魯門說在冰山有一些樹枝。杜魯門飛到冰山。他在找一些樹枝。他沒有發現樹枝。賀拉斯在找一些肉。他發現一些肉。他吃杜魯門。杜魯門死了。

現在每年都會舉辦年度洛伯納獎（Loebner prize）的競賽，讓電腦程式們競爭是否能通過圖靈測試，讓裁判們認為它們實際上是人類。截至 2014 年為止，尚無任何電腦程式贏得二獎（成功讓裁判認為人其實是電腦）與首獎（成功讓裁判無法辨識出人與電腦，也就是通過圖靈測試，一旦首獎頒出，這個競賽即宣告結束。）不過每年都會頒出裁判認為「最像人類」的電腦程式。1991 年第一次的比賽中，一個電腦程式設法透過故意打錯字而得到這個獎！

目前尚沒有任何人工智慧系統通過完整的圖靈測試。不管有沒有，許多哲學家們仍持續在爭辯圖靈測試無法真正衡量出大多數人所謂的「智慧」。它所測試的是行為，也就是一個電腦程式是否擁有「智慧」的表象，但這不盡然代表它有「智慧」。如果你沒有意識，不認識自己，沒有知覺，沒辦法感受自我意識，沒辦法體驗愛，甚至沒辦法「活著」，那可以算是有人類智慧嗎？

人工智慧的爭論看來還會持續幾十年吧。

延伸閱讀

哲學家 John Haugeland 所著的“Artificial intelligence: the very idea”是一本非常傑出，值得閱讀的書。裡面談到關於人工智慧的爭論，在這個活動中所描述的一些事情也是從這本書來的（特別是 SHRDLU 的對話與它們的討論）。

原始的圖靈測試是由 Alan Turing 所寫的文章“Computing machinery and intelligence”中所描述。這篇文章在 1950 年刊登在“Mind”這本哲學期刊上，並且在 Feigenbaum 與 Feldman 編著的“Computers and thought”重新被提出。這篇文章包含了最開始的兩個對話。

Eliza 這個模仿心理治療師的程式在 J. Weizenbaum 於 1966 年發表在“Communications of the Association for Computing Machinery”這本電腦雜誌中的文章“ELIZA -- A computer program for the study of natural language communication between man and machine”中有描述。

喜歡堆積木的那個機器人程式，是由 Terry Winograd 所寫的博士論文中所提出，發布在“Understanding natural language”這本書中（出版社為 Academic Press，1972 年紐約）。

產生故事的程式是由 Tony Smith 與 Ian Witten 在 1990 年於“Proceedings of the 10th International Conference on Computing and the Humanities”中發表的“A planning mechanism for generating story text”中有描述。故事原文如下：“Once upon a time there was an Arctic tern named Truman. Truman was homeless. Truman needed a nest. He flew to the shore. Truman looked for some twigs. Truman found no twigs. He flew to the tundra. He met a polar bear named Horace. Truman asked Horace where there were some twigs. Horace concealed the twigs. Horace told Truman there were some twigs on the iceberg. Truman flew to the iceberg. He looked for some twigs. He found no twigs. Horace looked for some meat. He found some meat. He ate Truman. Truman died.”