

如何用 AI 教小孩作數學

許聞廉

中央研究院資訊所特聘研究員

實驗室簡介

- 1992年，發表自然輸入法，並將語意分析模組轉化為線上對談機器人
- 2000年，與業界合作，企業客服機器人
- 2005-2013 參與各種國際競賽：
 - 中文問答、斷詞、文本蘊含(entailment)
 - 英文生物專名辨識，relation extraction
- 2016年，智慧型知識管理與查詢（企業內部客服），智慧型對話系統
- 2020年，中文語音辨識及理解，改錯，摘要與文句生成

國際競賽名列前茅

斷詞、問答、推論、基因名稱辨識

- 1st in open track of ACL SIGHAN 2003, traditional Chinese word segmentation.
- 1st/8 in the NTCIR6 2005 CLQA Chinese Question Answering (44%)
- 1st/13 in the WS CityU closed track of the SIGHAN 2006 Word Segmentation (97.2%)
- 2nd/8 in the NER CityU closed track of the SIGHAN 2006 Named Entity Recognition (88%)
- 1st/8 in the NTCIR6 2007 CLQA Chinese Question Answering (57%)
- 1st in the NTCIR6 2007 CLQA English-to-Chinese Question Answering (34%)
- 1st/10 in BioCreAtIvE II.5 2009 Interactor Normalization Task (2008, Watson)
- 1st/12 in the NTCIR9 2011 Recognizing Inference in TExt (RITE) (66.1%)
- 1st/9 in the NTCIR9 2013 Recognizing Inference in TExt (RITE) (66.1%)
- 1st/10 in BioCreAtIvE 2017 Interactor Normalization Task

109年度大型活動-2020科技大擂台

- 科技部舉辦2020科技大擂台人工智慧競賽，內容包含文章閱讀簡答、申論及對話機器人。參賽者如果成績超越人類者，可獲獎金：第一名：兩千萬；第二名：500萬
- 基於本團隊過去在中文自然語言處理學界之研究成果，特邀本團隊合作製作訓練試題共1000餘題，文本200餘篇。其中每一題都註明不同的推論類型。另外，設計了對話機器人20個主題及其下不同之對話腳本。
- 共有99隊、逾400人報名挑戰。初賽錄取11隊，進行最後之決賽。決賽成績：參賽者的最高成績大約為人類測試者平均分數之60%



計畫主持人與科技部前瞻應用司司長楊琇雅、臺灣人工智慧實驗室創辦人杜奕瑾、科技部部長陳良基舉辦競賽記者會

小學數學解題範例

我們目前在語言理解上的研究計畫

Tools

General Batch ScriptStep

FAQ Type Infomap Primary Math

Map File Path C:\Users\iasl\Desktop\math\l8l204\小學數學ontology.db

Use Map GUID 43874373-ladb-4364-alf3-f65ad4fb42fe

小明有20元，小明有5顆蘋果，小華有10顆蘋果，一顆蘋果2元，小明買給小華4顆蘋果，小華吃了一顆，小華剩下幾顆蘋果？

student: 小明有20元，小明有5顆蘋果，小華有10顆蘋果，一顆蘋果2元，小明買給小華4顆蘋果，小華吃了一顆，小華剩下幾顆蘋果？

teacher: 由於「小華有10顆蘋果」在題目中是第一個描述「小華」所擁有的「蘋果」數量=10顆

teacher: 「10顆蘋果」是「小華」原有的量

teacher: 「小華」是「收受者」

teacher: 所以「小華」的「蘋果」

teacher: $10顆 + 4顆 = 14顆$

teacher: 「小華吃了一顆」是「重量減少」的概念 $-1顆$

teacher: $14顆 - 1顆 = 13顆$

teacher: 答案：小華剩下13顆蘋果

teacher: .

萬丈高樓平地起

小學生作數學的困難何在？

1. 數字計算：容易算錯
2. 將文字關係轉成算式：文字題題目看不太懂
(國文)

2.1 代名詞處理

爸爸給小明兩個蘋果，他又給小華兩個？

2.2 省略回填

2.3 中文文法鬆散

2.4 每個字都看得懂，但是不知道題目要我們做什麼（句意弄不清楚）。

我們的解題目標

- 儘可能理解句意，提高答案正確率
- 能自動產生解說（與工程師互動）
 - 翻譯成小學生容易理解的語言
 - 解說文字與公式並進
 - 系統易於進行錯誤分析
- 解題過程的分析（瞭解人類的學習過程）
 - 小朋友解題過程紀錄與歸納
 - 「困難點」的分析
- 疑難問答，提出暗示或建議

如何教機器看懂文字題？

- 自然語言理解是 AI 最為棘手的問題
- 人類語言具有顯著的歧義性。
 - 20 個字的句子會具有數百、數千種可能的句法結構（絕大多數都不具有意義）。機器內部的數學模型訓練了許多參數來預測哪一個結構最有可能。
- 一般自然語言研究會遇到的種種困難：
 - 指示代名詞、省略、倒裝句、文本蘊含(entailment)、文句剖析、本體論(ontology)、推論
- 小學數學解題需要的自然語言理解，知識範圍較為侷限，句型較為單純，但以上的困難全都存在。

電腦處理中文的難處

中文斷詞的困難

- 用音轉字系統來說明：

皮膚科

柯醫師

皮膚科醫師

治療這種過敏皮膚柯醫師最有經驗了

辨識語言的chunk很重要

中文處理的難處(I)

文法寬鬆

- 中文沒有單複數、時態的標註，通常是看上下文決定。
- 動詞、名詞混用（他游泳很快，他很會游泳）
- 惜字如金
 - 從台北到高雄的距離是360公里
 - The distance from Taipei to Kaohsiung is 360 kilometer
 - 台北到高雄是360公里
 - 台北高雄360公里
- 中文以「語意」為主。

中文處理的難處(II)

詞序自由

- 用正規的句型分析很難說清楚，但很適合用 pattern 來分析，譬如：

- 他作業寫得不錯

- S O V 補語

- 作業他寫得不錯

- O S V 補語

- 寫作業他很不錯

- V O S 補語

- 我有十個蘋果，六斤180元，四斤200元的梨子四個。

中文處理的錯誤類型

- 斷詞錯誤：一**台大**冰箱，還要幾**張才**夠？
- POS(詞性)錯誤：近幾年的**研究**發現 (將「研究」標為動詞)
- Parser錯誤：
 - **完成**清掃家裡的工作 Finish the job of house cleaning
 - **完成** {[清掃家裡] 的 **工作**} --- (完成, 工作)
 - **完成**清掃家裡的垃圾 Finish cleaning the household garbage
 - 完成 {**清掃** [家裡的 **垃圾**]} --- (完成, (清掃, 垃圾) event)

中文的描述比較是圖像式的

- 中文的parser遇到標點符號就認為句子結束。系統無法知道一句話的意思到哪裡才結束。
- 我有十個蘋果，六斤180元，四斤200元的梨子四個。
- Google 翻譯上句話的結果
 - I have ten apples, six catties (斤) of 180 yuan, four catties (斤) of 200 yuan pears.
 - 應該是I have ten apples that cost 180 yuan per 6 catties, and 4 pears that cost 200 yuan per 4 catties

語言理解和影像的關係

- 文字經

- 許多

- 人

- 事

- 成語

- 物

- 描述

- 事

- 海



- One picture is worth a thousand words

- 光靠文本中的文字不足以瞭解其文意

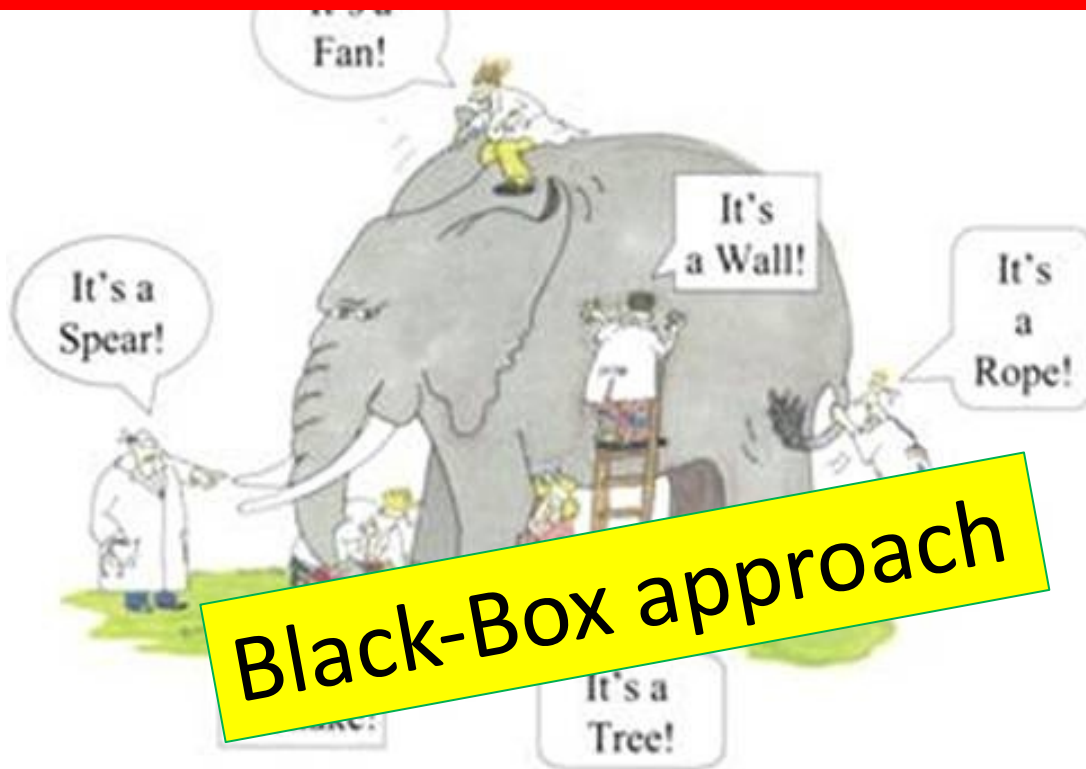
Property Description

- 修飾語裡面有許多的property description需要事先處理
- 這裡面的「換句話說」非常普遍，就像問答系統「意思相近的說法」一樣。
- 請問台北有沒有可以唱卡拉OK，能容納40人，並且有停車優惠的川菜餐廳？
- 我想明天搭高鐵下午四點以前到高雄，從南港出發。
- 靠近 P(換句話說)
- 離 P 不遠
- 距離 P 很近
- 與 P 距離不遠
- 到 P 很近
- 到 P 很快
- 從 P 很快就可走到
- 從 P 很快就到
- 可走路到 P
- 走路就到 P
- 走路一下就到 P

句子的理解

機器學習 瞎子摸象

When we can't see the whole picture (don't understand),
we create the features and learn their weights



統計與深度學習系統容易犯的錯誤

- 這些model大多在預測句子中的數字會形成怎樣的公式，不會給解釋。
- 小明有5個蘋果，吃了2個，又買了3個，請問小明現在還剩下幾個蘋果？
 - $5 - 2 + 3 = 6$
- 小明有5個蘋果，吃了2個，又買了3個梨子，請問小明現在還剩下幾個蘋果？（冗句）
 - $5 - 2 + 3 = 6$ **X**
- 一盒蘋果有10顆，小明吃了3顆，還剩下幾盒？（單位換算）
 - $10 - 3 = 7$ **X**

何謂理解一個句子？

- 知道這個句子的人、事、時、地、物，以及其所蘊含 (entailment) 的意義（相對於應用系統）
- 小明給小華3個蘋果
 - 小明少了3個蘋果
 - 小華多了3個蘋果
- 對應的算式
 - 小明原有蘋果數量 $- 3 =$ 小明現有蘋果數量
 - 小華原有蘋果數量 $+ 3 =$ 小華現有蘋果數量

句子理解範例(一)

- 小明買給小華3個蘋果
- 和「小明給小華3個蘋果」不太一樣
 - 小明買3個蘋果給小華（重組句子）
 - 小明少了3個蘋果的錢
 - 小華多了3個蘋果
- 三本漫畫書210元，如果買五本要花多少錢？
 - 一本漫畫書的價錢 = $210 \div 3 = 70$
 - 我們稱此為「單位量」的句子，可看成是reference

句子理解範例(二)

- 汽船以30公里/時的速率，在河中順流航行170公里，費時5小時，以相同速率返回原處需幾小時？
- 順流速率 = 汽船速率 + 水流速率
- 航行距離 = 170公里
- 返回原處：逆流航行170公里
- 逆流速率 = 汽船速率 - 水流速率

句子理解範例(三)

- 在一條公路的兩旁，每隔4公尺種一棵榕樹，路的兩端都種，共種了474棵榕樹。這條公路有多長？
- 線性、兩邊：公路的兩旁
- 間隔：每隔4公尺種一棵榕樹
- 端點：路的兩端都種
- 總數：474棵榕樹

「比」字句的理解

pattern recognition

- 爸爸的年齡比媽媽大三歲（歲是「年齡」的單位）
- 爸爸比媽媽的年齡大三歲
- 爸爸比媽媽大三歲（紅球比白球多三個）
- 完整的說法是「爸爸的年齡比媽媽的年齡大三歲」
- 比

| - 年齡

| - 爸爸

| - 媽媽

| - 大

| - 3

類比式機器學習

模擬人類「舉一反三」的學習模式

單字學習和概念學習

- 語言學習中，要能將一個單字學好，必需將這個單字的相關概念或一組單字合起來學才有效。
 - 和這個單字合起來使用的combination，或者說chunk
 - words (e.g., book, pen)
 - polywords (e.g., by the way, upside down)
 - collocations, or word partnerships (e.g., community service, absolutely convinced)
 - institutionalized utterances (e.g., I'll get it; We'll see; That'll do; If I were you . . .; Would you like a cup of coffee?)
 - sentence frames and heads (e.g., That is not as . . . as you think; The fact/suggestion/problem/danger was . . .)
 - text frames (e.g., In this paper we explore . . .; Firstly . . .; Secondly . . .; Finally . . .)

水果的概念

- 小明給小華每公斤100元的富士蘋果兩個
- 水果的相關概念中有
 - 個數
 - 價錢
 - 品牌
- 小明給小華100元的富士蘋果兩個
- 我有十個蘋果，六斤180元，四斤200元的梨子四個。

相似句

相似句 (I)

有相似句型及解法的句子（有時需要rewrite）

還需要在計算意義上相似

- 值量題

- 一本字典375元
- 一本漫畫72元
- 一打鉛筆賣84元
- 一打鉛筆值144元
- 一件白上衣賣320元
- 一個西瓜平均重2.45公斤

- 容量題

- 一本練習簿有30張紙
- 一本課外書有366頁
- 一列火車有7個車廂
- 一星期有7天
- 一串珠子108粒
- 一束玫瑰有9朵

Clustering Similar Sentences based on relabeling (generalization)

Similar structure or solution in primary math.
(often needs relabeling for generalization)

• 值量題

- 一本字典375元
- 一本漫畫72元
- 一打鉛筆賣84元
- 一打鉛筆值144元
- 一件白上衣賣320元
- 一個西瓜平均重2.45公斤
- [物件] V [物件值]

• 容量題

- 一本練習簿有30張紙
- 一本課外書有366頁
- 一列火車有7節車廂
- 一星期有7天
- 一串珠子108粒
- 一束玫瑰有9朵
- [物件] 有 [部件數量]

句子的關聯性

從pattern進行推論

- 一個句子後面可能接什麼樣的相關句？
- 類似一個詞可以構成其相關的概念一樣，一個句子也可考慮其相關句。這種相關性通常和應用密切相關。
- 甲、乙同時起跑，繞300米的環行跑道跑，甲每秒跑6米，乙每秒跑4米，第二次追上乙時，甲跑了幾圈？
 - 在上面這類問題中，常常會有哪些個句子？
 - 還狀（幾圈），線性（多久），速率，追上

關聯句

- 一盒巧克力有十顆，五盒有幾顆？
 - 「單位值」的概念
 - $A \times B = C$ 問 C ?
- 一盒巧克力有十顆，吃了兩顆，還剩幾盒？
 - 「屬性值」的概念
 - $A - B = C$ 問 C ?
- 小明有一盒巧克力，吃了兩顆，還剩五顆，請問一盒巧克力有幾顆？
 - 條件句有未知數
 - $A - B = C$ 問 A ?

關聯句(2)

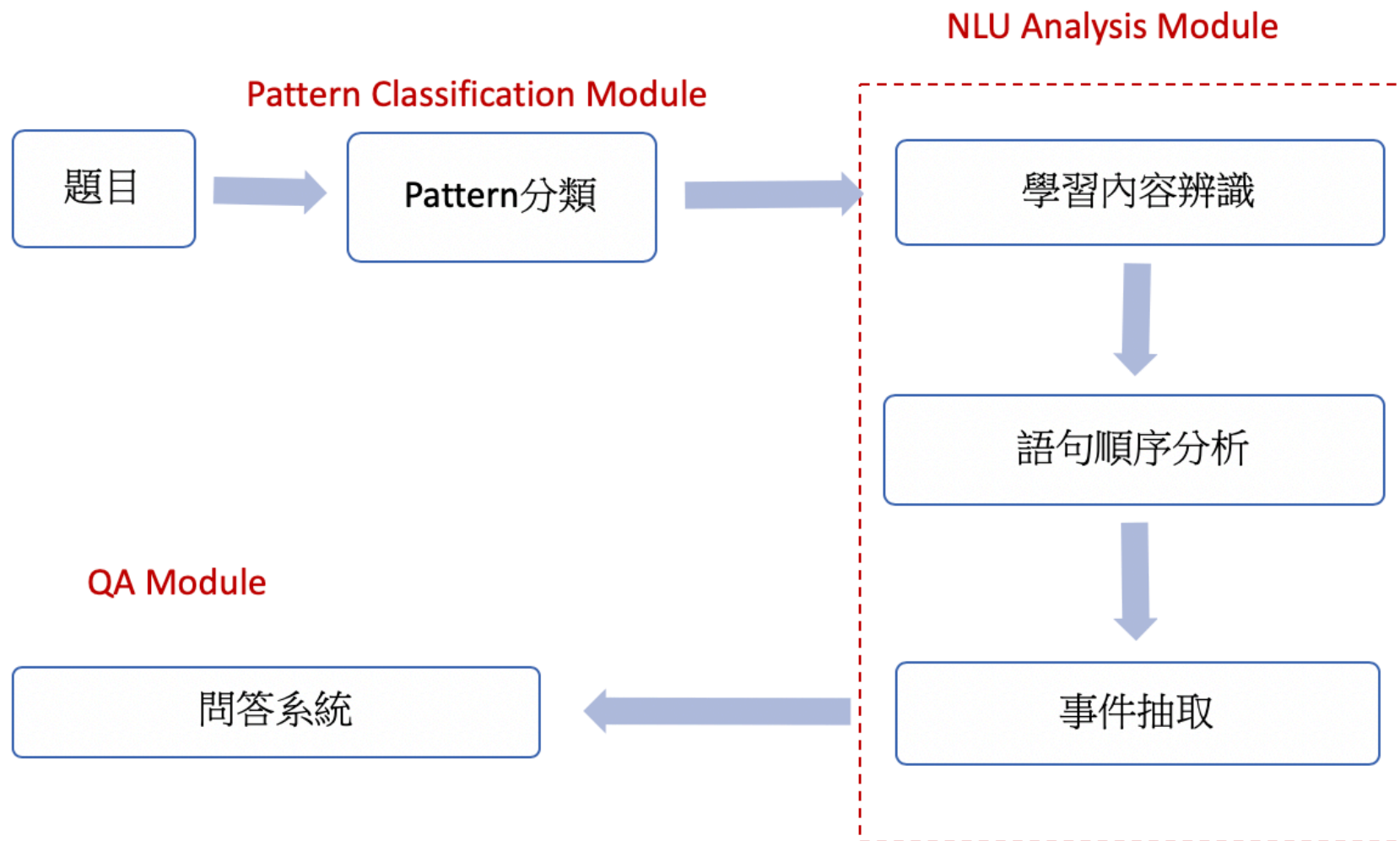
- 爸爸有十顆巧克力，吃了兩顆，還剩幾顆？
 - 吃了兩顆（沒頭沒尾），去前面找。
 - 爸爸吃了兩顆巧克力
- 小明喝了兩杯果汁，小華三杯。
 - 小華三杯（沒有動詞，受詞），去前面找。
 - 小華喝了三杯果汁

如何和學生互動？

小學數學教學輔助系統

- 在實際教學現場，老師多從事團體教學，對於學習能力落後的學生，我們需要特別一對一的關照。
- 我們研究出一套專門應用在數學文字題的教學輔助對話系統，透過與學生對談的「智能家教」，輔助進度落後的學生找出學習問題，提升學生的學習能力。
- 這套系統同時也能分析出學生的程度，以及在哪個解題階段出現問題。

小學數學教學輔助系統-架構



小學數學教學輔助系統

- 實際找學生做研究



小學數學教學輔助系統-範例

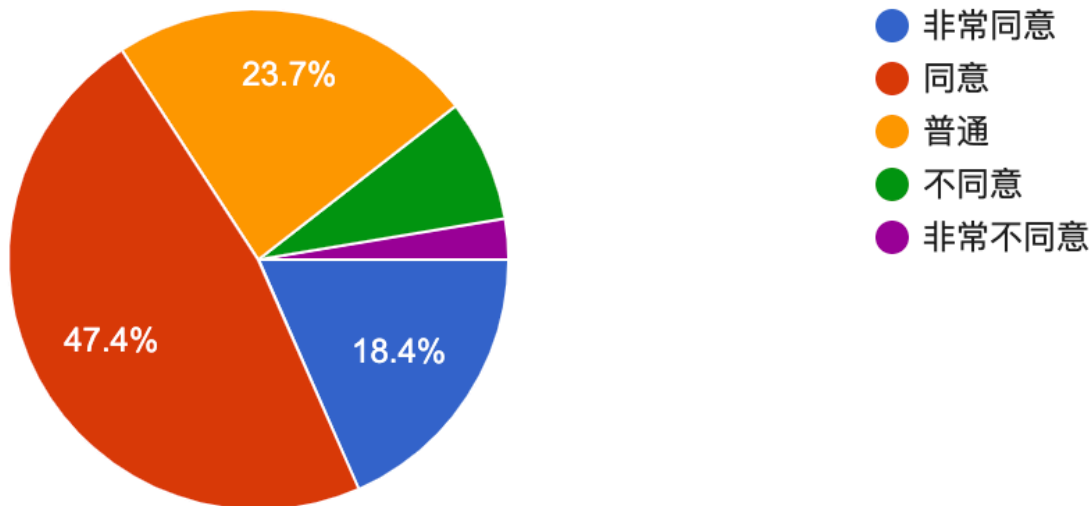
- **ROBOT(佈題)**：陳阿姨買了4公斤黃豆，做豆漿用了2.54公斤，還有多少公斤的黃豆？
- 生：6公斤（錯）
- **ROBOT**：不對喔！我把題目改一下，再試著做做看吧！
 - 爸爸買了10公升牛奶，喝了2.6公升，還有幾公升牛奶？
- 生：7.4公升（對）
- **ROBOT**：恭喜你答對了！其實這兩題解法是一樣的喔！
- **ROBOT**：我們再仔細看看這個題目：其實只是把
 - 「爸爸買了10公升牛奶」換成「陳阿姨買了4公斤黃豆」
 - 「喝了2.6公升」換成「做豆漿用了2.54公斤」
 - 「還有幾公升牛奶？」換成「還有多少公斤的黃豆？」

小學數學教學輔助系統-範例

- **ROBOT**：所以請你再想一想原本的題目：陳阿姨買了4公斤黃豆，做豆漿用了2.54公斤，還有多少公斤的黃豆？
- 生：6公斤（錯）
- **ROBOT**：不對喔！請問你的答案是如何算出來的？
- 生： $4 + 2.54 = 6$ （無法掌握減法pattern）
- **ROBOT**：在這個題目中，「用了」有減少的意思喔，算式應該是 $4 - 2.54 = ()$ ，括號內是多少呢？
- 生：1.46公斤
- **ROBOT**：很好！

小學數學教學輔助系統

- 老師們看過對於整個對話流程給予正面的回饋。



- 實驗後給學生做問卷，學生也認為這樣的系統對學習是有幫助的。

Q & A

