



New Constellations

Question and background

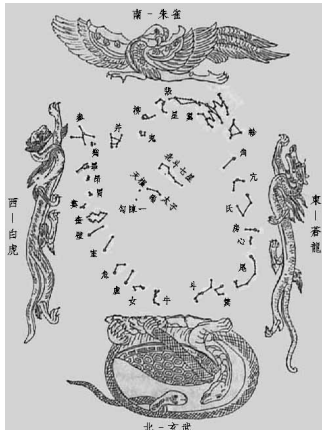


Fig.1 Chinese Twenty-Eight Constellations Chart
(Courtesy of the Astronomy Laboratory, Department of Physics, Cheng Kung University, Taiwan, China)

Vega and Altair flank the Milky Way;

Orion and Antares part night and day.

- Liweng Rhythm

Since time immemorial, people have observed the starry sky and seen groups of stars (and other objects) forming structures that are easy to find and identify – constellations. Different cultures have developed their own systems for dividing the starry sky into constellations: Greco-Roman - from which the modern system adopted by the International Astronomical Union (IAU) derives – as well as Chinese, Indian, and others.

In the modern astronomy, the 88 commonly accepted constellations are defined as areas of the celestial sphere, rather than simply sets of stars – this excludes “empty” areas that are not included in any constellation. However, the purpose of distinguishing constellations is still to facilitate the visual identification of areas of the celestial sphere when viewed with the naked eye. At the same time, for historical reasons, constellations have rather irregular shapes corresponding to various images – mythical heroes, animals, and objects.

Is it possible to redefine the constellations in a more convenient way, while retaining their identifiability with the naked eye and the existence of an “emotional” or “symbolic” description that facilitates in memorisation and naming?

Tasks

1. Study the attached dataset from Yale Bright Star Catalogue (Hoffleit *et al.*, 1991, 5th Revised Edition; <https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=V/50>) – it contains 9,110 stars visible with naked human eye. Construct a visualisation of the celestial sphere, taking into account that stars’ right ascension (RA) and declination (DE) are basically spherical coordinates of the star on the celestial sphere. Ensure that the visualisation allows you to perform a standard observational astronomy action – finding the North Star (Polaris) using the ‘edge’ of the Big Dipper (Ursa Major).



Fig.2 Drawing of Orion constellation from the catalogue *Prodrum Astronomiae* by Johannes Hevelius, 1690 (Image from Internet)

Important additional information:

- a. “asu.tsv” is the main Catalogue, it is a text file with fields delimited with “;” and comment lines starting with “#”; important fields are “_RAJ2000” – decimal right ascension in the J2000 epoch, “_DEJ2000” – decimal declination, “HR” – star number, “Vmag” – visual magnitude; a few objects have empty RA and DE fields and these objects should be excluded.
 - b. Visual magnitude is an inverse logarithmic measure of visible brightness relative to the visible brightness of Vega: $m \approx -2.5 \log_{10} \left(\frac{F}{F_{Vega}} \right)$, where F is the visible brightness. Hence, $m_{Vega} = +0.03$, $m_{Sirius} = -1.46$, $m_{Polaris} = +2.02$. Bigger m means a less visible star: difference of 1 in m corresponds to an approximately 2.512 times difference in visible brightness. $m = +6$ or $m = +6.5$ is usually considered to be the limit of visibility.
 - c. To identify famous stars, you can use the attached file “asu_names.tsv”, which provides association between HR numbers and names. Useful keys are HR_{Polaris}=424; HR_{Dubhe}=4301; HR_{Merak}=4295; HR_{Vega}=7001; HR_{Sirius}= 2491.
 - d. To identify standard IAU’s constellations, you can use the attached database “asu_constellations.tsv” (Davenhall *et al.*, Constellation Boundary Data, 1989, downloaded from <https://cdsarc.cds.unistra.fr/viz-bin/cat/VI/49>); in “_RAJ2000” and “_DEJ2000” fields, it provides “corner points” of constellation boundaries, with the name of constellation given in the “cst” field in the standard 3-letter form; full names of constellations can be obtained, for example, from here: <https://cdsarc.u-strasbg.fr/ftp/cati/vizier/constellations.htx>; note that Serpens is a two-part constellation encoded separately as “SER1” and “SER2”.
2. Propose a model for dividing stars into groups – i.e. “new constellations” – that look identifiable, separated from each other, convenient for practical observations with the naked eye, and possibly include other convenience considerations. The division criterion should be clear and quantitative, and the proposed division should be as optimal as possible in terms of this criterion.

You are free to define the “new constellations” either as sets of visible stars, or as segments of the celestial sphere. Dimmer, less visible stars can be omitted from the analysis; although full classification of stars is better.

Note that some of convenient options for quantifying the pairwise proximity of points on a sphere are the “great-circle distance” and the length of a chord connecting the points. Both have relatively simple expressions that use the longitude and latitude (here, RA and DE) of the points.

3. For the resulting “new constellations”, create a system of names associated with the images “visible” in the pattern of the stars in each constellation. Since this task is essentially an image

recognition problem with an extremely wide range of possible images, artificial intelligence (AI) systems that work with both graphic images and text may be a suitable intermediate tool. The ability to recognise an “interesting” object in the star pattern can be back-propagated into the criterion of constellation formation in the previous Task as one of the factors.

Submission

Your team's solution paper should include a 1-page Summary Sheet. The body cannot exceed 20 pages for a maximum of 21 pages with the Summary Sheet inclusive. The appendices and references should appear at the end of the paper and do not count towards the 21 pages limit.

Special Note

The principles of *Honor System* of IMMC also applies to the use of LLMs (large language models) or generative AI tools. If a team uses any large language model or generative AI tool in completing the modeling tasks, from problem research, model development, programing to paper/report writing, the team must make honest, open and transparent disclosure, including making in-text citations and detailing relevant content in the "References" section. It should be recognized that although large language models or generative AI have the advantages of productivity tools, they also have obvious shortcomings and pose risks to users (such as AI-generated content containing AI hallucinations or possible plagiarism in the produced content). Whether or not using large language models or generative AI tools itself in the team's work has no impact on judges' evaluation; the judges seriously remind every team that if any AI tool would be used, use it correctly, honestly, open and transparently.



新星座

问题与背景

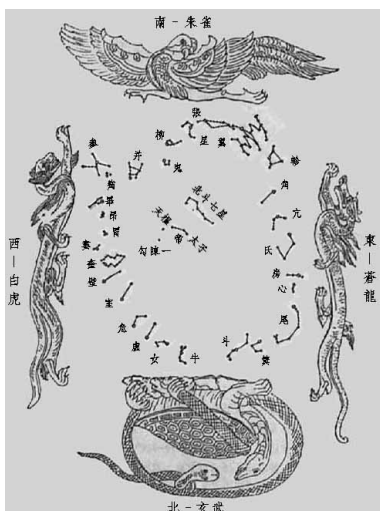


图1 中国二十八星宿图

（感谢中国台湾成功大学物理学系天文学实验室的图片）

牛女二星河左右；
参商两曜斗西东。

——《笠翁对韵》

自古以来，人们观测星空时发现群星（及其他天体）会构成易于寻找和辨识的结构——星座。不同文化发展出各自划分星空的星座体系：希腊-罗马体系（国际天文学联合会采用的现代星座体系即源于此）、中华体系、印度体系等等。

在现代天文学中，被普遍认可的 88 个星座被定义为天球区域，而非简单的星群组合——这种划分方式排除了不属于任何星座的“空白”区域。然而，区分星座的目的仍是为了方便通过肉眼直观辨识天球区域。同时由于历史原因，星座的形状往往不规则，对应着各种意象——神话英雄、动物以及器物等。

我们是否能用更便捷的方式重新定义星座，同时保留其肉眼可辨识的特性以及便于记忆命名的“情感化”或“象征性”描述？



图2 约翰内斯·赫维留的 *Prodromus Astronomiae* 星表中的猎户座插图，1690 年（图片来自互联网）

任务

1. 请研究题目附件数据集——摘自《耶鲁亮星目录》（Hoffleit 等，1991 年，第五次修订版；可下载自：

<https://vizier.cds.unistra.fr/viz-bin/VizieR?source=V/50>）。该数据集包含 9,110 颗肉眼可见的恒星。构建一个天球可视化图，要求正确利用恒星的赤经（RA）和赤纬（DE）作为恒星在天球上的基本球面坐标，并确保该可视化图可用来完成标准天文观测操作——通过北斗七星（大熊座）的“斗边”找到北极星。

重要补充信息：

- a. 数据集当中的“asu.tsv”是主目录文件，为分号“;”分隔的文本文件，以“#”开头的行是注释行；重要字段包括：

“_RAJ2000”：J2000.0 历元的十进制赤经

“_DEJ2000”：J2000.0 历元的十进制赤纬

“HR”：亮星编号

“Vmag”：目视星等

有少量天体的赤经和赤纬字段为空；此类天体应被排除。

- b. 目视星等是相对于织女星（Vega）亮度的逆对数量度：

$$m \approx -2.5 \log_{10} \left(\frac{F}{F_{Vega}} \right), \text{ 其中 } F \text{ 为可见亮度}$$

因此， $m_{Vega} = +0.03$ ， $m_{Sirius} = -1.46$ ， $m_{Polaris} = +2.02$ 。 m 数值越大表示恒星越暗： m 星等相差 1 级对应约 2.512 倍的亮度差异。通常认为 $m = +6$ 或 $m = +6.5$ 是肉眼可见的星等极限。

- c. 识别著名恒星可使用附件中的“asu_names.tsv”，该文件提供 HR 编号与星名的对应关系。关键参数包括：北极星 $HR_{Polaris}=424$ ；天枢星 $HR_{Dubhe}=4301$ ；开阳星 $HR_{Merak}=4295$ ；织女星 $HR_{Vega}=7001$ ；天狼星 $HR_{Sirius}=2491$ 。

- d. 识别国际天文学联合会（IAU）标准星座可使用附件中的“asu_constellations.tsv”（Davenhall 等，星座边界数据，1989 年，下载自 <https://cdsarc.cds.unistra.fr/viz-bin/cat/VI/49>）；其字段“_RAJ2000”和“_DEJ2000”提供星座边界的“角点”坐标，“cst”字段以标准三字母代码表示星座名称；完整星座名称可参考 <https://cdsarc.u-strasbg.fr/ftp/cati/vizier/constellations.htx>；注意蛇夫座为双部星座，分别编码为“SER1”和“SER2”。

2. 提出一种将恒星分组（即“新星座”）的模型，要求这些分组应具备可识别性、相互分离性、便于肉眼实际观测，并可兼顾其他便利考量。划分标准需明确且可量化，提出的划分方案在该标准下应尽可能最优。

可将“新星座”定义为可见恒星的集合，或天球的分区。较暗的恒星可从分析中略去，惟完成所有恒星的分类更佳。

需注意，量化球面上点对邻近度的常用方法包括“大圆距离”和连接两点的弦长。两者都可用经度与纬度（此处即赤经与赤纬）较为简洁地表达。

3. 请为最终得到的“新星座”创建命名体系，命名需与每个星座星群图案中“可见”的意象相关联。由于该任务本质上是涉及极广可能性的图像识别问题，可同时处理图形图像与文本的人工智能（AI）系统或可作为合适的辅助工具。从星群模式中识别“有趣”目标的能力，可反向融入前述任务中的星座形成标准，作为其考量因素之一。

提交

你的团队所提交的论文应包含 1 页摘要，其正文不可超过 20 页，包括摘要则最多不超过 21 页。附录和参考文献应置于正文之后，不计入 21 页之限。

特别说明

IMMC“诚信赛制”的原则同样适用于大语言模型或生成式 AI 工具的使用。如果团队在完成建模任务过程中有使用任何大语言模型或生成式 AI 工具，从问题研究、模型开发、程序编写到论文写作等建模工作的方方面面，团队必须诚实和公开透明地做披露，包括做出文内标注和在“参考文献”部分详细列出相关内容。应认识到大语言模型或生成式 AI 虽具有生产力工具的优势，亦具有明显的不足，并对使用者构成风险（例如 AI 生成的内容存在 AI 幻觉，也可能构成剽窃）。团队使用或不使用大语言模型或生成式 AI 工具本身，对评审没有影响；评委严肃地提醒团队，若果有任何 AI 工具的使用，都应当是正确、诚实和公开透明的使用。



IMMC 2026 中華賽 A 題（秋季賽）（English 簡體 繁體）

新星座

問題與背景

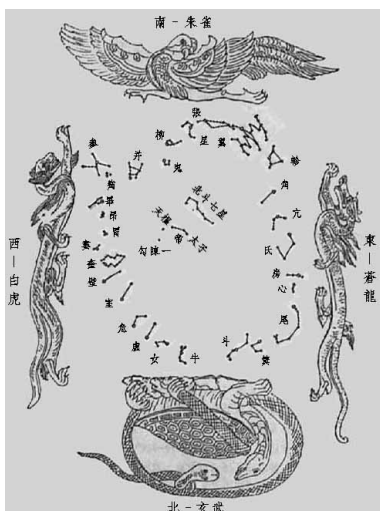


图1 中國二十八星宿圖

（感謝中國台灣成功大學物理學系天文學實驗室的圖片）

牛女二星河左右；
參商兩曜斗西東。

——《笠翁對韻》

自古以來，人們觀測星空時發現群星（及其他天體）會構成易於尋找和辨識的結構——星座。不同文化發展出各自劃分星空的星座體系：希臘-羅馬體系（國際天文學聯合會採用的現代星座體系即源於此）、中華體系、印度體系等等。

在現代天文學中，被普遍認可的 88 個星座被定義為天球區域，而非簡單的星群組合——這種劃分方式排除了不屬於任何星座的“空白”區域。然而，區分星座的目的仍是為了方便通過肉眼直觀辨識天球區域。同時由於歷史原因，星座的形狀往往不規則，對應著各種意象——神話英雄、動物以及器物等。



图2 約翰內斯·赫維留的 *Prodromus Astronomiae* 星表中的獵戶座插圖，1690 年（圖片來自互聯網）

我們是否能用更便捷的方式重新定義星座，同時保留其肉眼可辨識的特性以及便於記憶命名的“情感化”或“象徵性”描述？

任務

1. 請研究題目附件數據集——摘自《耶魯亮星目錄》（Hoffleit 等，1991 年，第五次修訂版；可下載自：

<https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=V/50>）。該數據集包含 9,110 顆肉眼可見的恆星。構建一個天球可視化圖，要求正確利用恆星的赤經（RA）和赤緯（DE）作為恆星在天球上的基本球面座標，並確保該可視化圖可用來完成標準天文觀測操作——通過北鬥七星（大熊座）的“鬥邊”找到北極星。

重要補充資訊：

- a. 數據集當中的 `asu.tsv` 是主目錄檔，為分號（;）分隔的文本檔，以 # 開頭的行是註釋行；重要欄位包括：

“_RAJ2000”：J2000.0 歷元的十進位赤經

“_DEJ2000”：J2000.0 曆元的十進位赤緯

“HR”：亮星編號

“Vmag”：目視星等

有少量天體的赤經和赤緯欄位為空；此類天體應被排除。

- b. 目視星等是相對於織女星（Vega）亮度的逆對數量度：

$$m \approx -2.5 \log_{10} \left(\frac{F}{F_{Vega}} \right), \text{ 其中 } F \text{ 為可見亮度}$$

因此， $m_{Vega} = +0.03$ ， $m_{Sirius} = -1.46$ ， $m_{Polaris} = +2.02$ 。 m 數值越大表示恆星越暗： m 星等相差 1 級對應約 2.512 倍的亮度差異。通常認為 $m = +6$ 或 $m = +6.5$ 是肉眼可見的星等極限。

- c. 識別著名恆星可使用附件中的 `asu_names.tsv`，該檔提供 HR 編號與星名的對應關係。關鍵參數包括：北極星 $HR_{Polaris}=424$ ；天樞星 $HR_{Dubhe}=4301$ ；開陽星 $HR_{Merak}=4295$ ；織女星 $HR_{Vega}=7001$ ；天狼星 $HR_{Sirius}=2491$ 。

- d. 識別國際天文學聯合會（IAU）標準星座可使用附件中的 `asu_constellations.tsv`（Davenhall 等，星座邊界數據，1989 年，下載自 <https://cdsarc.cds.unistra.fr/viz-bin/cat/VI/49>）；其欄位“_RAJ2000”和“_DEJ2000”提供星座邊界的“角點”座標，“cst”欄位以標準三字母代碼表示星座名稱；完整星座名稱可參考 <https://cdsarc.u-strasbg.fr/ftp/cati/vizier/constellations.htx>；注意蛇夫座為雙部星座，分別編碼為“SER1”和“SER2”。

2. 提出一種將恆星分組（即「新星座」）的模型，要求這些分組應具備可識別性、相互分離性、便於肉眼實際觀測，並可兼顧其他便利考量。劃分標準需明確且可量化，提出的劃分方案在該標準下應盡可能最優。

可將「新星座」定義為可見恆星的集合，或天球的分區。較暗的恆星可從分析中略去，惟完成所有恆星的分類更佳。

需注意，量化球面上點對鄰近度的常用方法包括「大圓距離」和連接兩點的弦長。兩者都可用經度與緯度（此處即赤經與赤緯）較為簡潔地表達。

3. 請為最終得到的「新星座」創建命名體系，命名需與每個星座星群圖案中「可見」的意象相關聯。由於該任務本質上是涉及極廣可能性的圖像識別問題，可同時處理圖形圖像與文本的人工智慧（AI）系統或可作為合適的輔助工具。從星群模式中識別「有趣」目標的能力，可反向融入前述任務中的星座形成標準，作為其考量因素之一。

提交

你的團隊所提交的論文應包含 1 頁摘要，其正文不可超過 20 頁，包括摘要則最多不超過 21 頁。附錄和參考文獻應置於正文之後，不計入 21 頁之限。

特別說明

IMMC “誠信賽制”的原則同樣適用於大語言模型或生成式 AI 工具的使用。如果團隊在完成建模任務過程中有使用任何大語言模型或生成式 AI 工具，從問題研究、模型開發、程序編寫到論文寫作等建模工作的方方面面，團隊必須誠實和公開透明地做披露，包括做出文內標註和在“參考文獻”部分詳細列出相關內容。應認識到大語言模型或生成式 AI 虽具有生產力工具的優勢，亦具有明顯的不足，並對使用者構成風險（例如 AI 生成的內容存在 AI 幻覺，也可能構成剽竊）。團隊使用或不使用大語言模型或生成式 AI 工具本身，對評審沒有影響；評委嚴肅地提醒團隊，若果有任何 AI 工具的使用，都應當是正確、誠實和公開透明的使用。