

Gesamtpunktzahl: 30

Abgabe der Lösungen bis zum 11.01.2016

Anhand der folgenden Aufgaben sollen Sie lernen, eine etwas komplexere Aufgabenstellung mit den Mitteln der Logikprogrammierung zu lösen. Begründen und dokumentieren Sie in jedem Fall Ihre Annahmen und Designentscheidungen und diskutieren Sie die Vor- und Nachteile Ihrer Lösung im Vergleich zu den ggf. vorhandenen Implementationsalternativen.

**Hinweis:** Wenn Prolog die Ausgabe von Listen mit ...abkürzt, können Sie durch drücken von `w` und `p` bei der Wahl einer Alternative die vollständige Ausgabe ein- und ausschalten. Dasselbe erreichen Sie durch die Eingabe `set_prolog_flag(answer_write_options, [quoted(true), spacing(next_argument)])`.

**Aufgabe 1:** Datenimport und Formatierung

6 Punkte

maximale Bearbeitungszeit: 45 Minuten

Auf [www.energy-charts.de](http://www.energy-charts.de) stellt die Fraunhofer Gesellschaft Statistiken zur Stromproduktion in Deutschland bereit; die Daten sind im JSON-Format gespeichert.<sup>1</sup>

1. Interpretieren Sie die in `month_2014_01.json` enthaltene Information, sortieren Sie danach, welche fachlich relevant sind und welche der Darstellung auf der Webseite dienen. Wo und in welcher Einheit ist die Information zur elektrischen Energie gespeichert, wo und in welcher Einheit die Zeitinformation?
2. Implementieren Sie ein Prädikat, das eine JSON-Datei in das Format überführt, das in `2014_01_facts.pl` beschrieben ist und verwendet wird.

**Hinweise:** Das Prolog-Modul zur Verarbeitung von JSON laden Sie mit `:-use_module(library(http/json))`. Nutzen Sie `open/3` zum Öffnen der Datei und `json_read/2` aus dem JSON-Modul zum Einlesen. Analysieren Sie dann die Ausgabe von `json_read` und bringen Sie die Daten in das geforderte Format.

Bonus (3 Punkte): Erstellen Sie ein Prädikat, welches die Daten direkt aus dem Internet herunterlädt und geeignet formatiert.

<sup>1</sup> Beispielsweise finden sich die Dateien für die monatsweisen Erzeugungsdaten unter <https://www.energy-charts.de/power/month-JAHR.MONAT.json>, wobei JAHR und MONAT entsprechende Zahlenangaben sind.

## Aufgabe 2: Zeitreihenverarbeitung

24 Punkte

maximale Bearbeitungszeit: 180 Minuten

1. Die Zeitreihen in `2014_01_facts.pl` enthalten stündliche Werte (eine Granularität von  $1/(3600 \text{ Sekunden})$ ); für einen besseren Überblick wäre eine tageweise Darstellung (also Werte alle  $24 \times 3600 = 86400 \text{ Sekunden}$ ) nützlich.
  - a) Implementieren Sie ein Prädikat, welches die Granularität der Daten reduziert indem aufeinanderfolgende Werte (24 Werte für Stunden  $\rightarrow$  Tage) aggregiert werden. Ihr Prädikat soll eine `energy_time_series/4` sowie die neue Granularität (z.B. 86400 für einen Tag) entgegennehmen und die resultierende `energy_time_series/4` produzieren.
  - b) Reduzieren Sie die Granularität der Wind- und der Steinkohlezeitreihen auf Tage und überzeugen Sie sich anhand der Daten von Januar 2014, dass am Wochenende weniger Steinkohle verstromt wird, wohingegen der Wochentag keinen Effekt auf den Windstrom hat. (Wochenendtage im Januar 2014: 4., 5., 11., 12., 18., 19., 25., 26.)
2. Die bisher genutzte Zeitreihenrepräsentation ist ‘zeilenbasiert’ für jede Energieart getrennt. Der Vergleich von Erzeugungsarten zum gleichen Zeitpunkt ist allerdings in einer ‘spaltenbasierten’ Darstellung deutlich einfacher.
  - a) Implementieren Sie ein Prädikat zur Matrixumformung, das eine M-elementige Liste von N-elementigen Listen in eine N-elementige Liste von M-elementigen Listen überführt:

$$\begin{bmatrix} [1, & 2, & 3, & 4, & 5] \\ [5, & 4, & 3, & 2, & 1] \\ [4, & 4, & 4, & 4, & 4] \end{bmatrix} \rightarrow \begin{bmatrix} [1,] \\ [5,] \\ [4,] \end{bmatrix} \begin{bmatrix} [2,] \\ [4,] \\ [4,] \end{bmatrix} \begin{bmatrix} [3,] \\ [3,] \\ [4,] \end{bmatrix} \begin{bmatrix} [4,] \\ [2,] \\ [4,] \end{bmatrix} \begin{bmatrix} [5,] \\ [1,] \\ [4,] \end{bmatrix}$$

Verzichten Sie auf die Verwendung eingebauter Prädikate in Ihrer Prädikatsdefinition! Beschreiben und testen Sie Ihre Implementierung, auch im Vergleich zum eingebauten Prädikat `transpose/2`.

- b) Der Gesamtstromverbrauch zu jedem Zeitpunkt lässt sich berechnen als die Summe der einzelnen Erzeugungsarten und der Im-/Exportsalden. Erstellen Sie ein Prädikat das den Gesamtstromverbrauch als Zeitreihe (mit Bezeichnung ‘Total’) liefert. Nutzen Sie `findall` um die Einzelzeitreihen aus der Datenbasis in eine Liste von Listen zu lesen, und die Lösung aus 2a (oder `transpose`), summieren Sie dann die einzelnen Spalten.

3. Die Korrelation beschreibt die Ähnlichkeit zweier Zeitreihen. Der Pearson'sche Korrelationskoeffizient nimmt Werte zwischen  $-1$  und  $+1$  an, wobei Werte nahe  $0$  für keinen Zusammenhang, Werte nahe  $+1$  für eine starke Übereinstimmung und Werte nahe  $-1$  für einen starken Gegensatz stehen. Für zwei Vektoren  $\vec{x}$  und  $\vec{y}$  der Länge  $n$  berechnet er sich wie folgt<sup>2</sup>:

$$r = r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}$$

Die Korrelationsanalyse zwischen einzelnen Zeitreihen ist zum Beispiel nützlich um festzustellen, ob Erzeugungsarten dem Lastgang folgen, also immer dann besonders viel Strom erzeugen, wenn besonders viel Strom benötigt wird.

- a) Implementieren Sie ein Prädikat zur Berechnung des Pearson'schen Korrelationskoeffizienten für zwei gleichlange Zeitreihen.
  - b) Mit welcher Anfrage finden Sie die Stromerzeugungsart die am stärksten mit dem Gesamtstromverbrauch korreliert?
4. Das *Lastprofil* der Stromerzeugung bezeichnet den gemittelten täglichen (wöchentlichen, jährlichen) Verlauf. Aus dem Lastprofil ist erkennbar, dass nachts weniger Strom gebraucht wird als tagsüber, sonntags weniger als unter der Woche, und so weiter. Für das Tageslastprofil werden jeweils die Werte der Tagestunden  $0, 1, 2, \dots, 24$  über den Monat gemittelt.
- a) Implementieren Sie ein Prädikat, das das Lastprofil einer Datenreihe für eine gegebene Periode (Tag, Woche, Jahr) berechnet. Teilen Sie dazu zunächst wie in 1a die Daten in die einzelnen Abschnitte, nutzen Sie dann die Matrixumformung aus 2a und bilden Sie schließlich die Mittelwerte über die einzelnen Teillisten.
  - b) Die Datei `lostdate_facts.pl` enthält die stündlichen Daten zum Gesamtstrombedarf eines unbekannten 4-Wochen-Zeitraums. Finden Sie heraus, mit welchem Wochentag der Zeitraum beginnt. Reduzieren Sie dafür die Granularität auf Tage und bilden Sie das Wochenlastprofil.

---

<sup>2</sup>Siehe [http://en.wikipedia.org/wiki/Pearson\\_product-moment\\_correlation\\_coefficient#For\\_a\\_sample](http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient#For_a_sample).