

Aleph

**PV056,
Strojové učení a dobývání znalostí
Jakub Tischler, jaro 2010**

Osnova

System Aleph

Popis

Formát dat

Michalskiho vlaky

Aleph „v praxi“ (dataset z klinické praxe)

Popis dat & cíle

Od dat k učení

Výsledky alephu

System Aleph - Popis

Založen na predikátové logice (prolog)

Možnost vložit do učení expertní (doménovou) znalost

Snadný zápis strukturovaných dat

System Aleph – Popis II

Postup učení:

Vezmi jeden pozitivní příklad

Vytvoř co nejspecifičtější model, pro tento příklad
(bottom clause)

Generalizuj

Vyber nejlepší generalizaci

Odstraň pozitivní příklady, které generalizace pokrývá

Opakuj dokud nedojdou pozitivní příklady

System Aleph – Formát dat

3 soubory:

*.f – pozitivní příklady

*.n – negativní příklady

*.b – vše ostatní

Soubor *.b obsahuje:

Mode & Determination

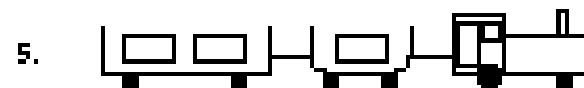
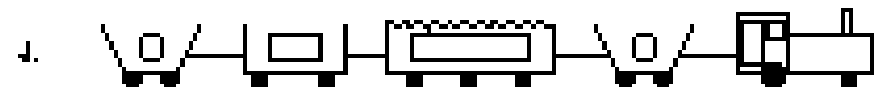
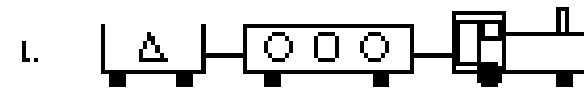
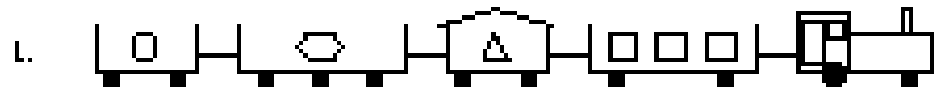
Seznam všech možných hodnot pro použité typy

Data

Doménovou znalost

Nastavení

System Aleph - Vlaky



Praxe



Praxe – Popis dat

Data z klinické praxe

Cca 90 pacientů, atributy:

Věk, pohlaví, délka léčby, dávka léku, ..

Vedlejší příznaky léčby (~20 atributů)

Varianta jednoho genu (A1 / A2 / A3 ...)

Všechny mutace tvořící variantu genu

Rychlost metabolizace (MR) jistého druhu léčiv

Před podstoupením léčby (třídní atribut)

Po podstoupení léčby (třídní atribut)

Cíl: Predikce MR

Praxe – Od dat k učení

Převést data do formátu alephu

Strukturované atributy lze výhodně zapsat

Rozchodit samotné učení (horor)

Rozšířit doménovou znalost

Sepsat / dát dohromady kód provádějící
crossvalidaci

Spustit experimenty.

Praxe – Výsledky alephu

Výsledky ukázaly, že v těchto datech žádné silné závislosti nejsou.

Aleph našel pravidla, která:

- Mohou napovědět, jaké závislosti v datech existují

- Mohou být využita ke konstrukci nových atributů

- Nemohou sloužit jako prediktivní model (nízká accuracy)

Závěrem

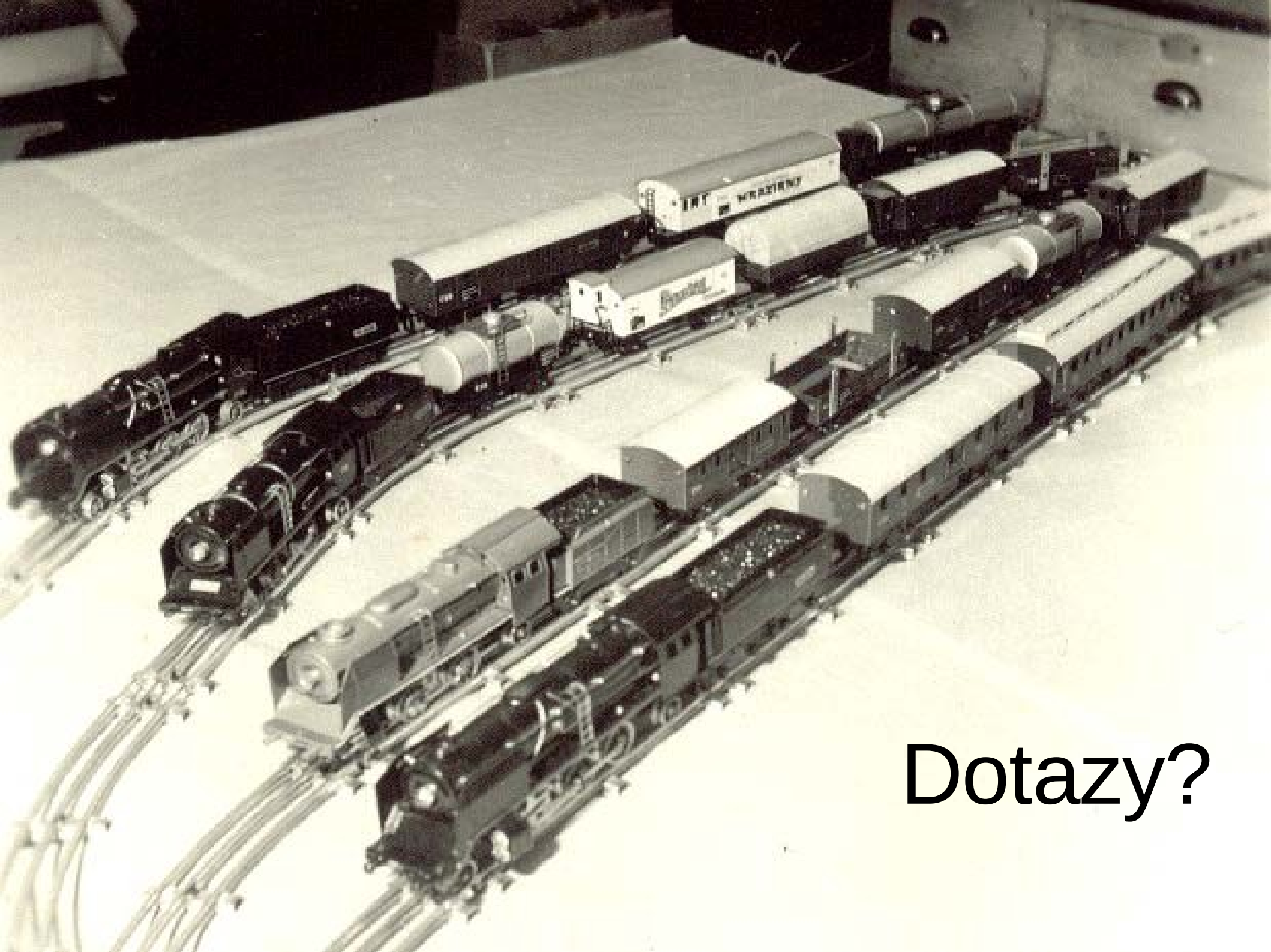
Při práci s alephem pozor na několik věcí:

Deklarace modů bývá problém

Některé drobné chyby zboří celé učení

Aleph nic nepočítá – jen dosazuje fakta

S alephem se pracuje dobře, ale pokud začínáte, konzultujte s někým, kdo už aleph umí. Často. Předejdete hodinám frustrace.



Dotazy?

Zdroje

Aleph manual:

<http://www.comlab.ox.ac.uk/activities/machinelearning/Aleph/aleph.html>

Obrázek z úvodního slidu: <http://www.oakbaypolice.org/cybersafety.html>

Obrázek 4 vláčků: <http://www.merkurklubplzen.cz/plechove-vlacky-merkur.cz>

Obrázek vlaků east- westbound:

<http://www.doc.ic.ac.uk/~sgc/teaching/v231/lecture10.html>

Obrázek Garfielda: <http://cinie.wordpress.com/2009/08/27/look-on-the-bright-side/>