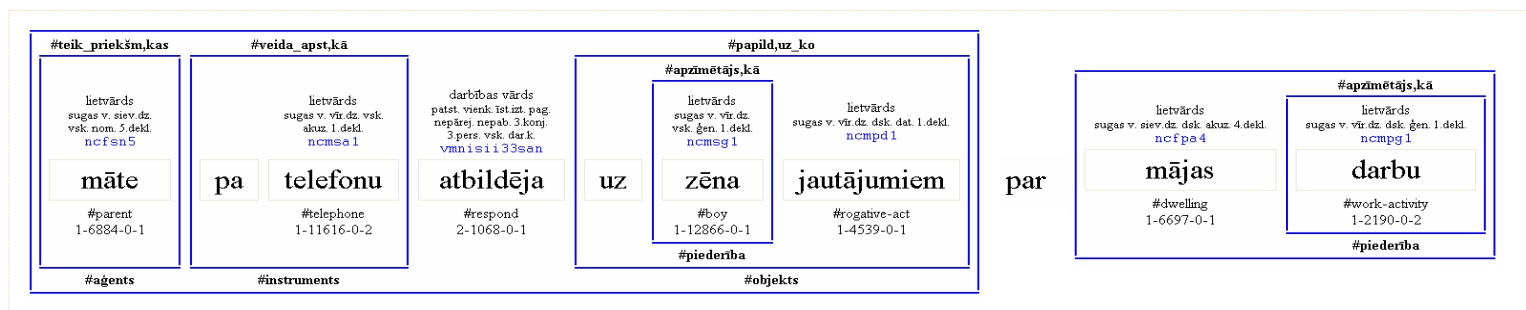


# **Semantiskā tīmekļa un datorlingvistikas aplikācijas Grid tīklā**

***G.Bārzdīņš, J.Džeriņš, M.Freivalds, I.Poikāns***

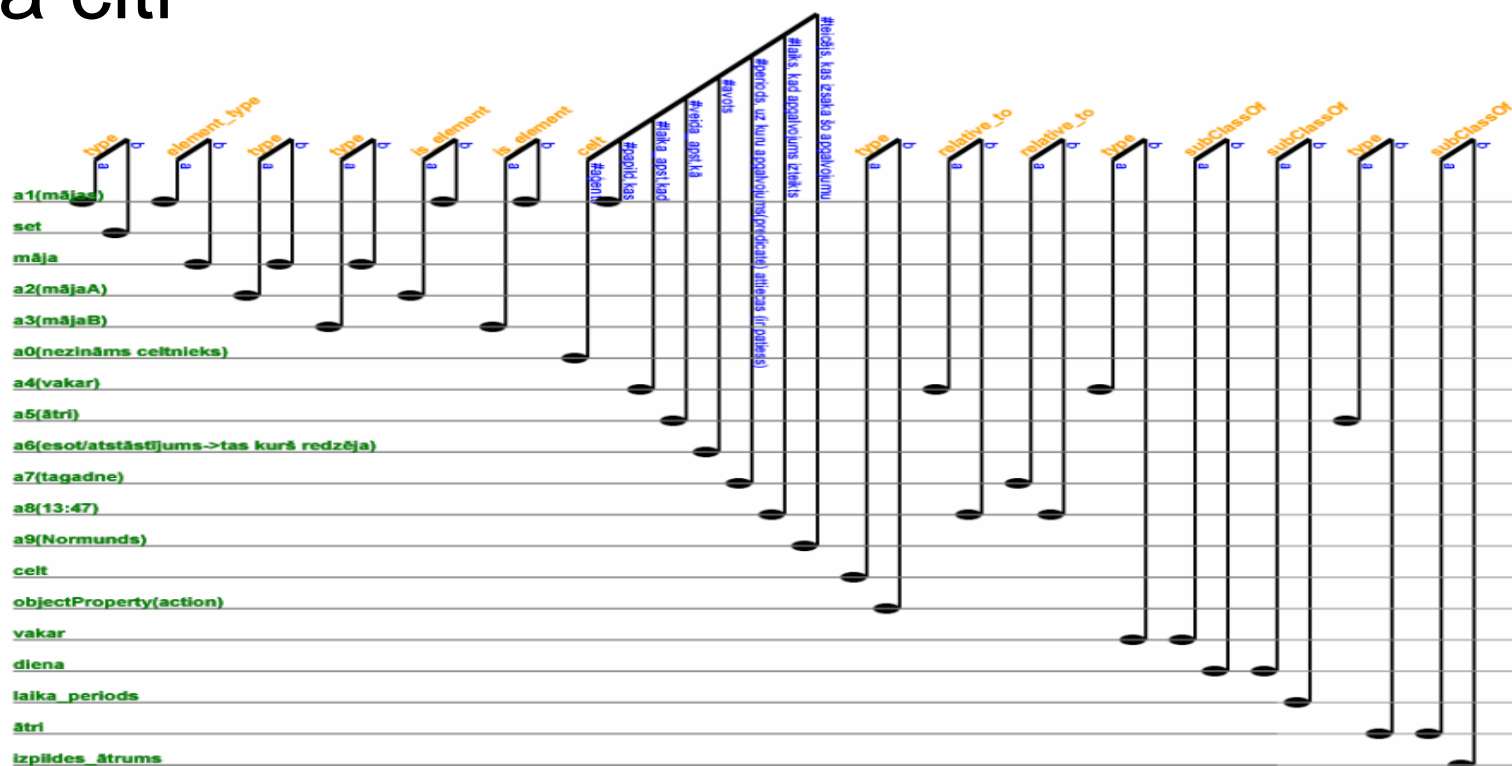
# SemTi-Kamols

- Mūsu aplikācija atpazīst ierobežotu skaitu vārdu un teikumu konstrukciju
- Piemēram, teikums “Māte pa telefonu atbildēja uz zēna jautājumiem par mājas darbu”



# Saistība ar semantisko tīmekli

- Jēdzieni tiek piesaistīti definīcijām ontoloģijā
- Rīkus un tehnoloģijas darbam ar ontoloģijām taīsa citi



# Datoru resursu nepieciešamība

- Mūsu aplikācija atpazīst tikai ierobežotu skaitu vārdu un teikumu konstrukciju
- Atpazīstot vienu konstrukciju, var gadīties, ka tādējādi esam izslēguši citas konstrukcijas atpazīšanu
- Tāpēc nepieciešams apstrādāt visus iespējamus teikuma apakšteikumus, lai atrastu visgarāko atpazīstamo

# Datoru resursu nepieciešamība

- Viegli izsecināt, ka šāda veida teksta apstrādei piemīt eksponenciāla sarežģītība atkarībā no apstrādājamā teksta garuma
- Toties:
  - Varam izmantot nepabeigtu programmu
  - Varam noteikt, ko mūsu programma nespēj, un attiecīgi to papildināt

# Mūsu mērķis

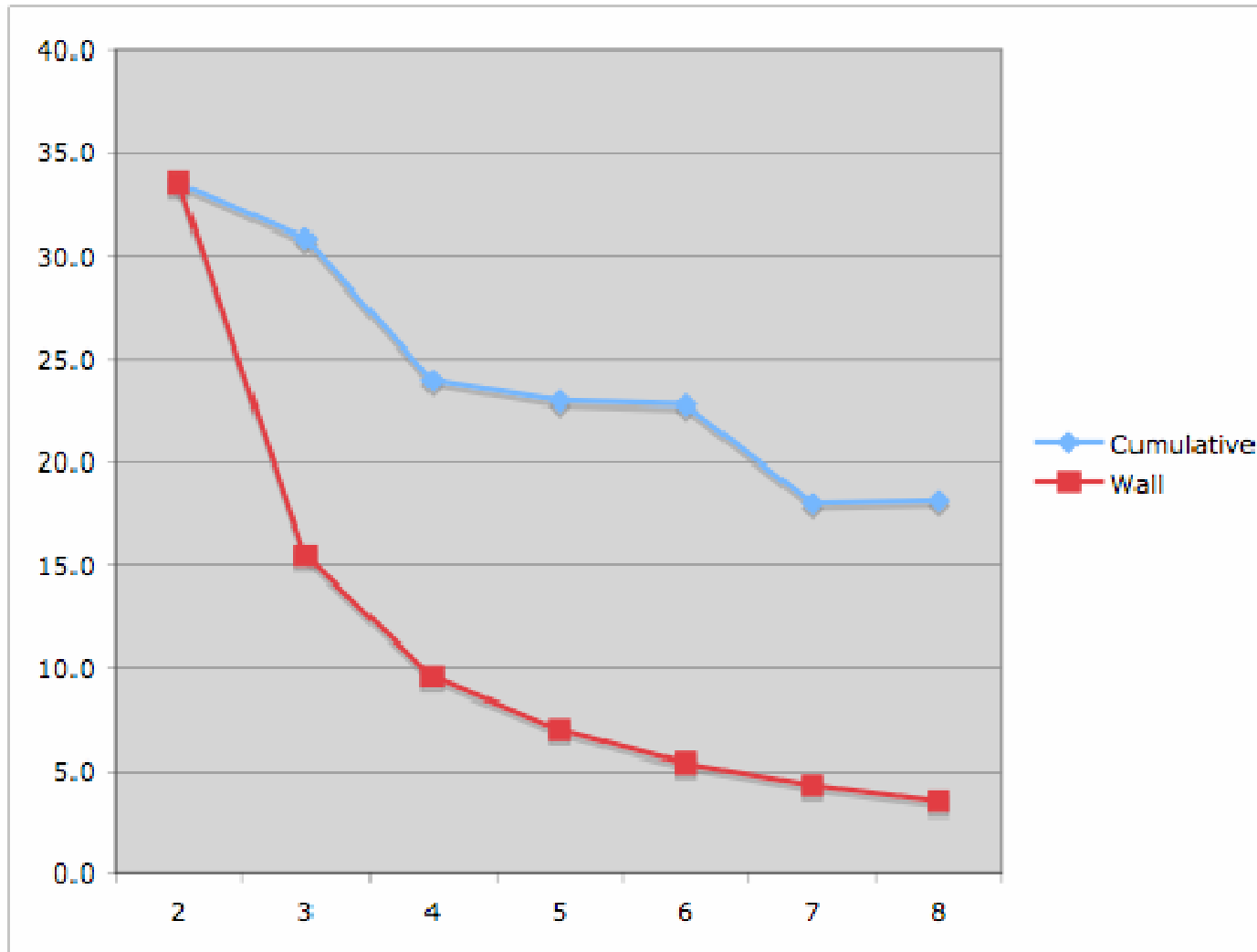
- Mērķis ir apstrādāt lielus teksta korpusus, kas ļaus mums:
  - Statistiski vērtēt programmas rezultātus (piem. vidējo sekmīgi atpazītā apakšsteikuma garumu)
  - Izgūt tekstu semantiku [vēlāk]

	A	B	C	D	E	F	G	H
1	Fails	Laiks	Ilgums	Faili	Atpazīti	Neatpazīti	Atpazīts	Neatpazīt
2	D:\Webarhiivs\www.saeima.lv\	03:58:15	00:12:18	1338	7,503,218	3,019,201	71%	29%
3	D:\Webarhiivs\www.antivirus.lv\	23:59:18	00:05:18	1330	732,975	445,774	62%	38%
4	D:\Webarhiivs\www.likumi.lv\	02:20:42	00:04:27	1761	3,980,398	1,356,279	75%	25%
5	D:\Webarhiivs\www.kapaok.lv\	01:49:20	00:02:22	167	289,836	578,834	33%	67%
6	D:\Webarhiivs\www.mikrotik.com\	02:52:57	00:02:18	89	112,081	333,223	25%	75%
7	D:\Webarhiivs\www.historia.lv\	01:26:56	00:01:56	1862	1,684,670	822,651	67%	33%
8	D:\Webarhiivs\www.narcomania.lv\	03:02:25	00:01:54	166	320,342	77,557	81%	19%
9	D:\Webarhiivs\www.liepajniekiem.lv\	02:14:30	00:01:43	2607	1,801,941	976,596	65%	35%
10	D:\Webarhiivs\vip.latnet.lv\	23:31:31	00:01:34	1066	1,262,445	496,112	72%	28%
11	D:\Webarhiivs\www.alberts.lv\	23:49:44	00:01:30	649	2,773,071	1,135,457	71%	29%
12	D:\Webarhiivs\robsardze.gov.lv\	23:17:23	00:01:25	2221	3,168,787	1,945,344	62%	38%
13	D:\Webarhiivs\www.lza.lv\	02:37:22	00:01:24	419	1,151,086	430,477	73%	27%

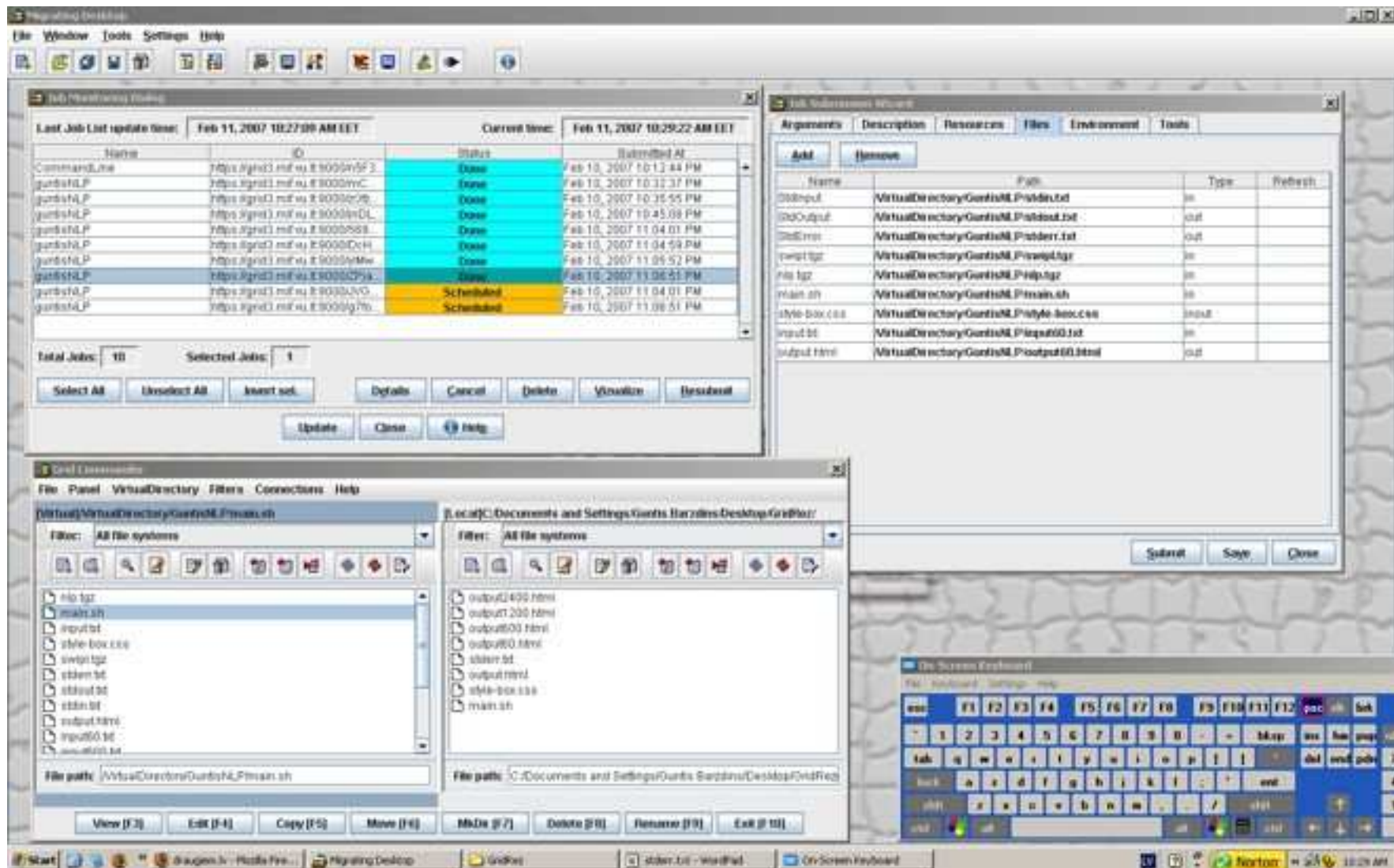
# Programmas darbināšana izmantojot MPI

- MPI ir viena no izplatītākajām bibliotēkām paralēlu programmu darbināšanai
- Mūsu aplikācijai tika uzrakstīts papildinājums SWI Prolog implementācijai, kas ļauj lietot MPI Prolog programmās
- Plašāka informācija: <http://apps.lumii.lv>

# Paralēlās darbības mērījumi



# Ar poļu palīdzību programma darbojas Migrating Desktop vidē



# Turpmākie darbi

- Integrācija ar OCM-G, kas ļautu programmu darbināt Gridā interaktīvi