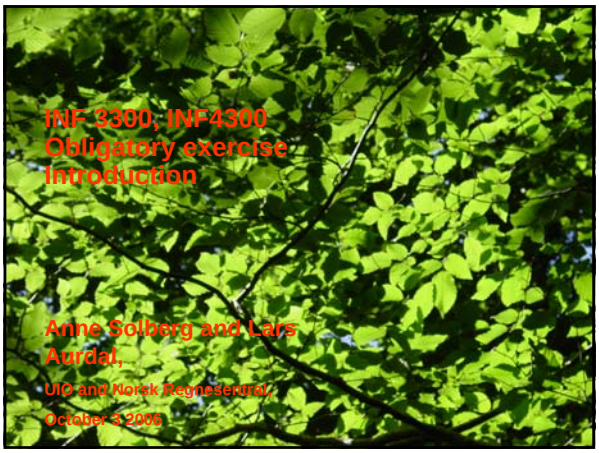


What is phenology

1. Phenology is a branch of science that deals with the relations between climate and periodic biological phenomena such as bird migration or plant flowering.
2. In this exercise we will study plant phenology.

What is phenology

1. Plant phenology has occupied scholars for a long time.
   a. Omens
   b. Religion
   c. Quasi-science
   d. Science
2. Some countries have very extensive (and high quality) data on plant phenology, for instance apple budburst, apple blossoms etc.

Who cares

1. Until recently there was not much interest in research on plant phenology.

2.But then…

Who cares

1. As human beings started worrying about possible climatic effects of their activities, there was a need for indicators of climatic change.

2.Enter plant phenology…

Phenology and climate

1. The phenology of certain plants is an excellent indicator of changing climatic conditions.
2. In the US the average growing season of plants in general is 15 days longer in urban areas than in rural areas.
3. In the period 1951-1997 the first blossom date of *Syringa Vulgaris* has on the average advanced with more than 0.2 days per year (by almost 10 days over the period, this was measured in Central Europe).

Norwegian phenological studies

1. There has never been any official Norwegian effort to gather such data.
2. However there have been (and are) a few private initiatives.
3. The most extensive gathering of such data where carried out by the botanist K.H.O. Printz of the University of Oslo.

Printz's study

1. In 1928, inspired by the phenological networks in Europe, he established an observational network throuhgout Norway.
2. 705 stations were established and various phenological phases were monitored:
   a. First flowering date.
   b. Foliation.
   c. Fruiting.
   d. etc…

Printz's study

1. Results from 260 stations (stage 1) were published by Prof. Printz et. al. in in 1955 and 1959.
2. A new set of 60 stations (only a few in common with stage 1) was started in 1952. A few more stations were added later and observations continued more or less until 1977. Results from stage 2 were published by Prof. Printz et. al. in 1978 and 1990.

Printz's study

1. In his 1959 publication Prof. Printz listed all the tables containing the data.
2. There are 282 tables in all.
3. They were handwritten by one single person with a beautiful handwriting.

The tables, interpretation

Table number

Location

County and municipality number

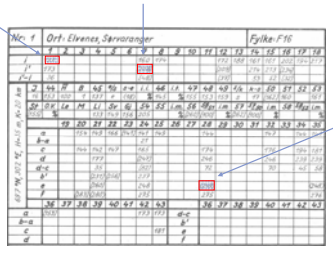
Position, EASL and distance to the ocean

The tables, interpretation

Average Julian date of some event

Average based on only 3 or 4 years

Average based on only 1 or 2 years



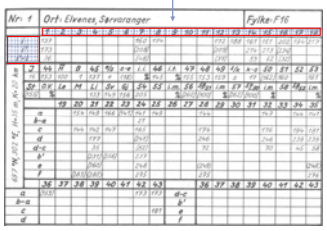
The tables, interpretation

1. The date numbers are simple Julian dates.
2. In an ordinary year the 28th of February is denoted by 59.
3. In leap years the 29th of February is denoted 59x.

The tables, interpretation

Plant type

i=flowering
i'=fruit ripe
i'-i=timespan

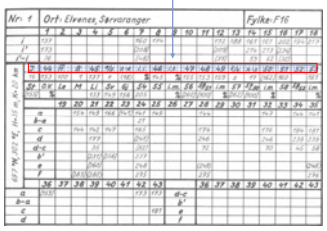


The tables, interpretation

1. Plant types (Norwegian names):
   1. Hestehov
   2. Blåveis
   3. Hvitveis
   4. Rødsildre
   5. Ballblom
   6. Nyresildre
   7. Maria nøklebånd
   8. Ballblom
   9. Liljekonvall
   10. Markjordbær
   11. Gjøksyre
   12. Skogstjerne
   13. Linnea
   14. Blåbær
   15. Multer
   16. Gjeitrams
   17. Mjødurt
   18. Røsslyng

The tables, interpretation

See next slides



The tables, interpretation

1. J: Number of observation years.
2. 44: Budburst at tree limit.
3. H : Average height where 44 is measured.
4. B: Type of tree defining tree limit (in Norwegian):
   1. Bjørk
   2. Bjørk og rogn
   3. Selje
   4. Hegg
   5. Gråor
   6. Rogn og selje
   7. Rogn og hegg
   8. Rogn og gråor
   9. Selje og hegg
   10. Selje og gråor

The tables, interpretation

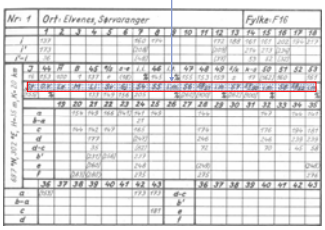
1. 45: Ice break.
2. e/s: River (e) or lake (s) defining ice break. If both are available then the earliest is recorded and the time difference between the events is recorded in s-e
3. i.i.: No ice, the percentage of open water
4. i.t.: No ground frost, the percentage of ground not frozen.
5. 47: First day of plowing

The tables, interpretation

1. 48: Greenup og cultivated fields.
2. 49: Release of domestic animals.
3. s/k: Sheep (s) or cattle (k) defining 49.
4. k-s: Difference in time between release of cattle nad sheep.
5. 50-53: Sow time for different cereals.

The tables, interpretation

See next slides



The tables, interpretation

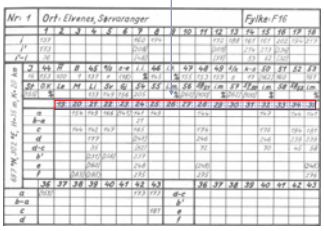
1. St.: First observation of “stær”.
2. O.v.: Percentage of “stær” that did not migrate.
3. Le.: First observation of “lerke”.
4. M.: First observation of “måltrost”.
5. Li.: First observation of “linerle”.
6. Sv.: First observation of “svale”.
7. Gj.: First observation og “gjøk”.

The tables, interpretation

1. 54: Fields ready for harvest.
2. 55-58: Different cereals ready for harvest.
3. i.m.: Percent cereal not ripe at harvest date.
4. 56-51 etc.: Growth period of different cereals.

The tables, interpretation

See next slides

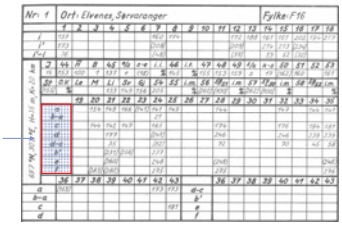


The tables, interpretation

1. Tree types (Norwegian names):
   19. Hassel
   20. Gråor
   21. Selje
   22. Osp
   23. Lavlandsbjørk
   24. Fjellbjørk
   25. Alm
   26. Sommereik
   27. Bøk
   28. Hegg
   29. Slåpetorn
   30. Kirsebær
   31. Eple
   32. Rips
   33. Stikkelsbær
   34. Bringebær
   35. Rogn

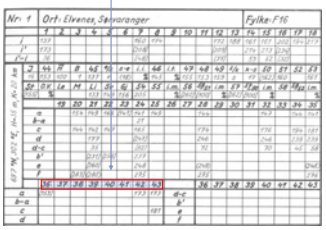
The tables, interpretation

- a=greenup
- b=end greenup
- b'=start senescence
- c=flowering
- d=start ripening
- e=start leaffall
- f=end leaffall



The tables, interpretation

See next slides

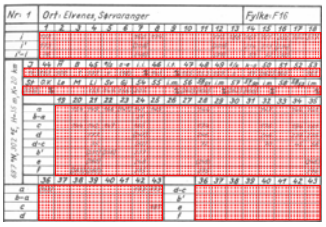


The tables, interpretation

1. Tree types (Norwegian names):
   36. Lønn
   37. Lind
   38. Syrin
   39. Ask
   40. Nyperose
   41. Jasmin
   42. Gran
   43. Furu

The problem

Find the text in the red areas using image processing and pattern recognition for all 282 tables



Practical issues

1. The images are available in the file oblig_h06_data.zip at the course website.
2. The pages of the original book were scanned (graytones, 300dpi).
3. The scans were stored as TIFF images, table_1.tif to table_71.tif.
4. Most table images show 4 tables per image (one image, table_1.tif shows only two).

Practical issues

1. Apart from this NOTHING has been done to the images.
2. The images might have a varying background (smudges etc.)
3. The tables can be somewhat rotated in the images, some distortion may also occur.
4. One table contains some manual notations (table_70.tif).

Practical issues

1. The algorithm can be written in any language.
2. Try not to reinvent the wheel, you may of course write you own routines, but we give **no extra points** for writing existing routines.

Practical issues

1. The is one deliverable:
   a. Preprocessing: Find the regions and symbols to be processed.
   b. Process (recognize) the symbols.
2. Note: It may not be possible to separate the two parts very cleanly, the note will anyhow be based on the totality of what you hand in.

Practical issues

1. Schedule:
   a. October 3, 2006: Exercise and data available.
   b. December 1, 2006: All deliverables due.

Practical issues

1. You must make your own report, you can not hand in a report in collaboration with someone else.
2. Copying other students material will automatically results in an F note, but you can (and should) cite scientific publications.

Practical issues

1. The report must be a **written document**.
2. The report must comprise:
   a. A full description of you algorithm.
   b. Examples of cases where it works and does not work.
   c. An analysis of the cases where it does not work.
   d. A full listing of the programs you have used (excluding libraries).

Practical issues

1. What we are trying to obtain is to force you to make a small scientific report.
2. The logic in the report is therefore important:
   a. You must analyze the problem.
   b. Explain your choice of methods.
   c. Show when it works…
   d. …and when it does not work.
   e. Suggest possible improvements.
   f. Draw a conclusion.
   g. Randomly trying many methods is **not the point.**
   h. Remember, the point is not necessarily the number of images you correct, but the logic behind your choice of method.

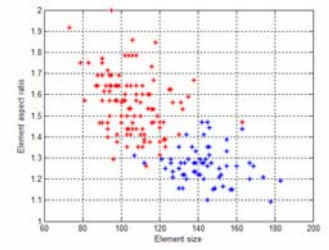
How to get going

1. Start by finding the tables and its orientation.
2. Make list of all possible symbols.
3. Extract the regions containing the symbols.
4. Split symbols and clean them.
5. Extract symbol features.
6. Run the features through a classifier.
7. Compare with manually interpreted tables.
8. What is your best recognition rate?

Reference

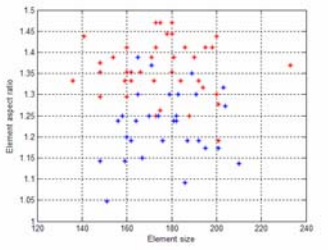
1. Anne and Lars will be solving this to some extent in parallell.
2. Currently we extract tables, symbols and some simple features.

Separability



Symbols **1** (red) and **2** (blue)
element size and aspect ratio

Separability



Symbols **6** (red) and **9** (blue)
element size and aspect ratio