

# Oppgaver INF 2820 V2011

## Obligatorisk innleveringsoppgave 1, deloppgave A

*Frist 7.2 – strikt – se beskrivelse på semestersiden*

### Oppgave 1: klokkelsettuttrykk

Gjør Oppgave 2.4 (s. 43 i brun bok) i J&M. Siden det viser seg at denne av en eller annen grunn ikke står i blå bok gjentar vi den her:

**2.4** Design an FSA that recognizes simple date expressions like *March 15, the 22nd of November, Christmas*. You should try to include all such “absolute” dates, (e.g. not “deictic” ones relative to the current day like *the day before yesterday*). Each edge of the graph should have a word or a set of words on it. You should use some sort of shorthand for classes of words to avoid drawing too many arcs (e.g., furniture → desk, chair, table).

Du kan velge om du vil gjøre den for norske eller engelske uttrykk. Du kan skrive forkortelser i et eget dokument, og så bare bruke navnet på forkortelsene i nettverket som du implementerer i JFLAP.

Innlevering: Dokumentet med forkortelser og JFLAP-nettverket.

### Oppgave 2: norske stavelser

- Noen bokstav- (og lyd-)sekvenser er norske ord, eks.: *skrått*
- Noen bokstav- (og lyd-)sekvenser er ikke norske ord, og kunne heller ikke vært det, eks.: *ktår*
- Mens noen bokstav- (og lyd-)sekvenser som ikke er norske ord kunne vært det, men er det ikke, eks.: *krått*

Skal vi lage en stavesjekker som bare godkjenner ekte norske ord, trenger vi en ordliste. Men uten en ordliste, kan vi bruke en FSA (eller et regulært) uttrykk for i hvert fall å skille mellom (b) på den ene siden og (a) og (c) på den andre siden.

Et ord består av en eller flere stavelser. Hver stavelse har en **kjerne**, som vi her kan regne med er en vokal eller diftong. Før kjernen kan det komme 0 eller flere konsonanter, og det samme er tilfelle etter kjernen. Eksempel: **da-ta-ling-vi-stikk**

Noen ganger kan det være vanskelig å avgjøre hvor en stavelse slutter og neste begynner, men det betyr ikke noe for denne oppgaven.

Oppgaven består i å lage en NFA for bokstavsekvenser som er eller kunne vært et norsk ord. Vi bruker bare små bokstaver. Oppgaven skal løses i JFLAP og testes på eksemplene:

skrått

krått

datalingvistikk

saueøye

falskmyntner

superfriskt

\*ktår

\*dataIngvistikk

\*superfistk

Hint:

Tenk igjennom hvilke konsonantsekvenser som kan innlede en stavelse; hvilke to bokstavssekvenser, hvilke trebokstavssekvenser, osv. Hva er de lengst kononantsekvensen som kan innlede en stavelse? Se hvordan du kan gruppere sammen tilfeller. Observerer at en kan referere til grupper i JFLAP som, eks, [j-n] for {j, k, l, m, n}.

Tenk tilsvarende for konsonantsekvensene som kan komme etter kjernen. For at det ikke skal bli for omfattende er det tilstrekkelig å ta med konsonantsekvenser av lengde 3 etter kjernen.

Oppgavene levers til Lars-Erik på e-post: Lars-Erik Bruce [larsereb@student.matnat.uio.no](mailto:larsereb@student.matnat.uio.no)

Skriv: 'INF2820 innlev 1a' i emnefeltet

Fint om besvarelsene pakkes sammen til et arkiv og kalles: <ditt navn>\_1a

- SLUTT