

Årlig programevaluering Informatikk: Språkteknologi, torsdag 7. februar 2019

Tid: 09:15 – 16:00 sted: Ole-Johan Dahls hus

Tilstede: Joakim Nivre, Lilja Øvrelid (LTG, vara), Jan Tore Lønning (programrådsleder), Stephan Oepen (LTG, referent), Erik Velldal (LTG, referent), Nikolas Martin (studentrepresentant -bachelor), Helge Lødrup (ILN, HF), Ragnhild K. Runde (undervisningsleder), Esra Boncuk (fra studieadministrasjonen, referent)

Ekstern programrådsgiver: Joakim Nivre (Uppsala Universitet)

Det jobbes med å finne en studentrepresentant fra masterprogrammet, samt en som kan ta over for Nikolas etter hvert.

Sak 1: Bachelorprogrammet

Historikk og oppbygning

Jan Tore innleder med forhistorien til programmet, beskriver og forklarer utviklingen til dags dato. Kort om stadiene: EDB for humanister (semesteremne, fra ca 1980), SLI (cand.mag og cand. philol, fra 1989), ITSLP (bachelor + master, fra 2003, etter bologna-/kvalitetsreformen), ISK (BSc + MSc, fra 2010, etter revisjon av alle IFI-programmer), 'Informatikk: språkteknologi' (fra 2017, etter revisjon av alle IFI-programmer).

Joakims tilbakemeldinger tyder på at bachelorprogrammet i språkteknologi har fulgt utviklingen i fagfeltet på lik linje med utenlandske læresteder. Endringen av navnet på programmet til språkteknologi er også slik fagfeltet omtales internasjonalt.

Føringer på utforming av programmet

- Krav fra IFI:
 - o Obligatoriske emner på alle IFI-programmer:
 - IN1020 – Introduksjon til datateknologi
 - IN1030 – Systemer, krav og konsekvenser 1
 - o Programmeringsundervisningen begynner med Python, så Java
- Krav fra MN:
 - o Minst 30 poeng skal være frie
 - o 5. semester = utviklingssemester, skal kun ha frie og ikke obligatoriske emner.
- Krav fra UiO: EXPHIL03 i alle bachelorprogrammer

Ikke krav om matematikkbakgrunn. Betyr at det regnes som et tverrfaglig studium og deretter at maksimalt halvparten av de obligatoriske emnene kan være informatikk/realfaglige.

Oppbygging og gjennomføring av bachelorprogrammet

- 110 av 180 poeng tas av obligatoriske emner.
 - 90 poeng av disse er IFI-emner:
 - IN1000 – Introduksjon i objektorientert programmering
 - IN1010 – Objektorientert programmering
 - IN1020 – Introduksjon til datateknologi
 - IN1030 – Systemer, krav og konsekvenser
 - IN1140 – Introduksjon til språkteknologi
 - IN1150 – Logiske metoder
 - IN2010 – Algoritmer og datastrukturer eller IN2040 – Funksjonell programmering
 - IN3120 – Søketeknologi eller IN3050 – Introduksjon til kunstig intelligens og maskinlæring
 - IN2110 – Språkteknologiske metoder
 - I tillegg kommer:
 - EXPHIL03 – Examen philosophicum
 - LING1112 – Morfologi og syntaks 1
- Siden programmet er tverrfaglig skal man også velge en emnegruppe på tilsammen 40 studiepoeng innen lingvistikk eller et språkfag.
- Da gjenstår 30 poeng frie emner.

Oppbygningen av bachelorprogrammet sammenlignes med Universitetet i Uppsala. Uppsala tilbyr matematikk til sine studenter i språkteknologi sammen med programmering. I løpet av andre halvdel av første semester, fortsetter de med en introduksjon til språkteknologi. Matematikkemnet tilbys også masterstudenter.

Det stilles spørsmål om man skulle tilby flere emner innen språkteknologi, men det ville resultert i flere små-emner. For å være en forskningsgruppe har språkteknologi-gruppen høy undervisningsbelastning. Man har valgt å fortsette undervisningen i IN2040 Funksjonell programmering da man har fått gode tilbakemeldinger, og det ligger et underliggende forsøk på å vekke interesse for språkteknologi blant andre studenter som følger IN2040.

Andre kommentarer til bachelorprogrammet

Fra Joakim:

- Det høres ut som en god struktur som samsvarer godt med tilsvarende programmer andre steder. Ikke stort innslag av obligatoriske emner i lingvistikk; gjenspeiler trenden i fagfeltet ellers.

- Ved Uppsala er det større variasjon i størrelsen på emner (f.eks 5, 7.5, 10, 15 poengs). Har eksamen flere ganger 1. semesteret. Gir mer fleksibilitet men også vanskeligere å få poengene til å 'gå opp'.
- Kreves heller ikke matematikk i Uppsala, men det gis en introduksjon til relevant matematikk i 1. semester, sammen med programmering. Andre halvdel av 1. semester brukes så på språkteknologi. anbefaler at vi fortsetter å gi språktek. fra 1. semester.
- Flere spesialiserte og valgfrie språkteknologiemner kunne vært fint men naturligvis vanskelig med tanke på ressurser.
- Ang. relasjon mellom BSc og MSc-programmene. Uppsala: vanskelig å rekruttere nok til masterprogrammet dersom man krever spesifikke bakgrunnskunnskaper i språkteknologi. Ikke minst vanskelig med internasjonal rekruttering. Kandidatprogrammet i Uppsala er nok smalere; i større grad en spesialisering i språkteknologi. Ved IFI gir vi relativt sett færre emner i språkteknologi.
- Jan Tore: Mål med BSc-programmet: kanskje først og fremst å være yrkeskvalifiserende som programmerer, ikke som språkteknolog. Alternativt å kvalifisere til å gå videre til MSc-programmet.
- Stephan: Ved å undervise IN2040 bruker gruppa også undervisningsressurser på ikke-språkteknologiske fag; bevisst valg som gir gruppa synlighet.
- Ang. rekkefølge på emner. Nå får studentene språkteknologi i første semester. Venter med logikk til andre semester. Utfordring å gi IN1140 parallelt med IN1000. Programmeringskonsepter i IN1000 introduseres for sent til at vi kan bruke det i IN1140. Nikolas: opplever det som positivt at studentene får språkteknologi fra 1. semester og da umiddelbart får kontakt med fagfeltet de har søkt seg til, men det kan være vanskelig for dem som ikke har noe bakgrunn i programmering fra tidligere. Også fint med bredde tidlig i graden for å få en oversikt over hva fagfeltet dreier seg om.
- Helge: merker økt tilfang av IFI-studenter ved andre lingvistikkemner utover det obligatoriske LING1112, noe som oppleves som svært positivt.

Sak 2: Gjennomgang av de programspesifikke emnene

Felles for alle de programspesifikke emnene er en todeling mellom praktisk programmering (typisk dekket gjennom obligatoriske oppgaver og gruppeundervisning) og teoretiske konsepter. En utfordring her er å unngå at studentene opplever at det ikke er samsvar mellom endelig eksamen og hva de testes i underveis.

IN1140 Introduksjon til språk teknologi

- Presentert av Lilja
- Emnet er en videreutvikling av INF1820 Introduksjon til språk- og kommunikasjonsteknologi. IN1140 er en Introduksjon til språk teknologi og forberedelse til lingvistikk (exfac-aktig som man har på HF).
- Hva som dekkes: lingvistiske nivåer, morfologi, regulære uttrykk og tilstandsmaskiner, tokenisering, sannsynligheter og språkmodeller, ordklasser og (regelbasert) ordklassetagging, syntaks og kontekstfrie grammatikker, leksikalsk semantikk, språk teknologiske applikasjoner.
- Ca 60 som tok eksamen H18.
- Vurdering: 5 obligatoriske innleveringer som må bestås, 4 timers skriftlig endelig eksamen.

Kommentarer fra Joakim: som introduksjonsbok på tilsvarende emne i Uppsala brukes 'Language and Computers' (2012), av Chris Brew, Detmar Meurers, and Markus Dickinson. Mulig alternativ til Jurafsky & Martin hos oss? I intro-emnene i Uppsala er det også mer fokus på å bruke eksisterende systemer (snarere enn å lage selv) som så trenes og evalueres. + Annotering. Merk at tilsvarende emne ved Uppsala gis som masteremne da de har valgt å rekruttere bredt og ikke nødvendigvis studenter med en bachelor innen språk teknologi.

Stephan Oepen, Erik Veldal presenterer IN2110 Språk teknologiske metoder

Undervises for første gang denne våren (2019). Vi snakker litt om tidligere emner (INF2820 og INF4820) som delvis videreføres.

IN2110 gir en myk introduksjon til maskinlæring (og noe oppfriskning av relevant matematikk) som kommer senere i studiet. Pensum og tema og innhold i forelesninger presenteres.

Interne diskusjoner angående innholdet i emnet presenteres. Tematisk progresjon: dokumenter, ord, setninger.

Emnet bygger delvis videre på IN1140, men man må likevel finne en balanse i undervisningen for de som har IN1140 og de som ikke har det da man tar opp studenter tilhørende forskjellig program.

Obliger er praktiske, mens eksamen er teoretisk. Studentene ønsker generelt sett mer praktisk eksamen. Men man har begynt å legge inn mer teoretiske oppgaver i obliger for å forberede dem bedre til eksamen. Studentene må bestå obliger for å kunne ta eksamen, mens karakteren gis på grunnlag av eksamensbesvarelsen.

Uppsala har gått for en strengere linje, der de krever forkunnskaper til sine emner. Mens ifi har valgt å bruke sine emner som en form for rekruttering til sitt masterprogram og har derfor emnene sine åpne.

Kostnader rundt evaluering av eksamen presenteres (flere må vurdere eksamensbesvarelsen osv.).

Uppsala – få skriftlige eksamener. Flere praktiske oppgaver, eller andre vurderingsformer (deltakelse o.l.) teller med.

Nirve sammenligner med Uppsala, og stiller seg positiv til ifis komplementerende emner (IN1140 og IN2110). Har positive erfaringer med å la studentene utvikle en enkel overgangsbasert parser (ved Marco Kuhlmann).

Jan Tore Lønning presenterer IN3120 Søketeologi

Undervises av Aleksander Øhrn; videreføring av tidligere INF3800. Forelesningene og pensum presenteres.

Burde maskinlæringen i IN3120 revurderes siden det er andre emner på ifi innen maskinlæring?

Jan Tore Lønning presenterer IN3050 Introduksjon til kunstig intelligens og maskinlæring

Forventet å bli et populært emne. Utvikles sammen med Robotikk (viderefører deler av INF3490 Biologically inspired computing).

Pensum er ikke avklart. Innføring med ikke altfor mye teori. Det er en utfordring å finne egnet pensum og lærebok. Jan Tore kritisk mot Marshland. Joakim kjenner igjen problemstillingen og forteller om hvordan Uppsala har løst dette; passe gode erfaringer med boka til Hal Daume III.

Både på UiO/ifi og Uppsala diskuteres Maskinlæring og kunstig intelligens – det er populært.

Sak 3 Masterprogram

Opptakskravene og oppbyggingen presenteres:

Valg av veileder og masteroppgave velges innen desember.

Andre semester skrives et essay som inngår som en del av masteroppgaven, dette blir gjerne brukt som bakgrunnsmateriale for masteroppgaven.

Joakim er enig i at det er gunstig å komme tidlig i gang med oppgaven.

Spør om programmering pluss noe lingvistikk kunne vært tilstrekkelig opptakskrav.

Kommenterer at Uppsala har i praksis lagt ned bachelorprogrammet i språkteknologi de siste årene, ønsker nå å rekruttere bredt og internasjonalt til mastergrad. Ellers ingen kommentarer til oppbyggingen.

Jan Tore Lønning presenterer IN4080 Natural Language Processing

Samme utfordring kan man se her: ikke alle studenter har lik bakgrunn/ballast (ikke alle har IN2110). Men høsten 2020 skal studentene ha bestått maskinlæring og IN2110 og da må man se på innholdet i IN4080 og IN5550.

Joakim: Man kan vurdere egne studenters behov og andre studenters behov i forhold til IN4080 fra høsten 2020 og fremover.

IN4080 i 2018 var tilpasset det gamle bachelorprogrammet og bakgrunnen studentene har fra det gamle bachelorprogrammet. Det forventes en gradvis overføring av stoff nedover, dvs. til IN2110 og IN3050.

Uppsala, har lavere forventninger til forkunnskaper til masterprogrammene. Det første semesteret bruker de på innføring. Etter jul får de maskinlæring og deep learning.

Eksamen – vi ser på et eksamenssett med løsningsforslag.

Joakim: skal masterprogrammet være en faktisk fordypning av det man har lært i bachelor eller skal man åpne for andre uten bakgrunn i språkteknologi? Viktig å være tydelig på hva man ønsker her. Det er ikke mange med en bachelorgrad i språkteknologi internasjonalt.

Stephan Oepen og Erik Velldal presenterer IN5550 / INF5820 Advanced Topics in Natural Language Processing / Language technological applications

Pensum gjennomgås.

Tidligere INF5820 brukt som pilot høst 2018

IN5550 vår. Ny vurderingsordning kun bestått/ikke bestått. Eksamen er prosjektoppgaver.

Joakim: tydelig sammenheng mellom IN4080 og IN5550. Han viser til Uppsalas prosjektemne (Language Technology and Research Development) på 15 poeng som går på diskusjon av forskning og jobbe i grupper på prosjekter. Uppsala har hatt publikasjoner gjennom dette emnet.

Jan Tore sender Joakim noen eksamensbesvarelser fra samtlige emner til gjennomgang.

Annet: Internasjonalisering

Stephan jobber med å se på potensielle avtalepartnere.

Vi har hatt to utvekslingsavtaler relevante for språkteknologi-studenter i Tyskland, men disse har ikke vært populære blant studentene.

Hittil, har ifi-studenter ikke vist stor interesse for utveksling, men det kan tenkes en endring da utviklingssemesteret gjør et utvekslingsopphold enklere. Utvekslingsstudentene som kommer til oss, har vi fine erfaringer med.

Gode utvekslingsavtaler anses som en mulighet for rekruttering.

Stephan har sett på to mulige utvekslingsdestinasjoner for språkteknologistudenter:

Universitetet i Uppsala
University of Melbourne

Begge lærestedene har gode fagmiljøer relevant for språkteknologistudenter, og UiO har allerede avtaler med begge læresteder.

Vi ser etter læresteder som er engelskspråklig, compatible semesterlengde, faglig meningsfylt. Stephan ønsker å se mer på emner evt. pakker eller kunne anbefale emner. Skal utveksling først og fremst være faglig motivert kan det være mest relevant på master.

Uppsala har lik erfaring som oss. (Tuebingen og Heidelberg i Tyskland). De har ikke fått sine studenter ut. Men fått ett par fra Tyskland inn.

Joakim: Uppsala og Oslo burde kunne tilby hverandre relevante emner som kan inngå i respektive studieprogram. Et annet spørsmål er om det er interessant å reise til Sverige for våre studenter. Søketeknologi og maskinlæring ved Uppsala er interessant for ifi-studenter.

Ifi nøler med å sende studenter ut på tredje semester. Men er åpen for å godkjenne en erstatning for IN5550 hos Uppsala.

Dersom masterprogrammet i språkteknologi åpnet for kort masteroppgave, kunne utveksling tredje semester vært et alternativ.

Videre diskuteres UiOs søknadsfrister for utveksling. Søknadsfristen 15.september passer dårlig for førsteårs masterstudenter som vil reise sitt andre semester. Esra hører med fakultetet om muligheter for å tilrettelegge fristen for masterstudenter.