

Central themes and questions in TEK5010 MAS 2023

Theme 0: General concepts in multiagent systems

Game theory topics:

Theme 1: Non-cooperative game theory

Theme 2: Voting

Theme 3: Cooperative game theory

Theme 4: Auctions

Theme 5: Bargaining

Theme 6: Arguing

Swarm intelligence topics:

Theme 7: Classical swarm intelligence

Theme 8: Response threshold

Theme 9: Swarm robotics 1

Theme 10: Swarm robotics 2

Theme 0: General concepts in multiagent systems

Question 1: Could you describe an agent?

En agent er et autonomt vesen (ofte en robot eller lignende) plassert i et miljø, som er i stand til å handle selvstendig for å nå et mål som på forhånd har blitt gitt av noen andre. Målet vil alltid være å gjøre endringer i miljøet på en ønskelig måte.

En intelligent agent:

- Har reactivity: responderer på endringer i miljøet
- Har proactiveness: klarer på selvstendig vis (uten ekstern input) å bestemme seg for handlinger som leder mot å fullføre målet
- Har social ability / er interactive: kan interagere med andre agenter for å nå målene sine

Typer agenter:

- Deductive reasoning agents: Bruker eksplisitte logiske regler for å bestemme og utføre mål
- Reactive agents: agentene reagerer kun på miljøet direkte (uten noen deduksjon) og emergens fremkommer av mange agenter som er reaktive sammen (og begynner å opptre komplekst)
- Hybrid agents: kombinasjon av de to. Reactive og proactive layer.

Question 2: What is a multiagent system?

Et multiagent-system er en sammensetning av mange agenter som handler, og samhandler, i ett og samme miljø. Disse agentene kan, som beskrevet over, være reaktive, interaktive eller en sammensetning av begge. Når disse agentene i fellesskap opptrer på en intelligent måte, kaller vi det Swarm Intelligence (SI).

Question 3: What do we mean by emergence in SI?

Med emergence mener vi at enkle, reaktive agenter som interagerer sammen gjennom primitive handlinger kan handle på en intelligent, mer kompleks måte i fellesskap. Den intelligente oppførselen emergerer altså fra enkel interaksjon mellom agentene.

Question 4: What is stigmergy?

Stigmergi er indirekte kommunikasjon mellom to agenter, gjennom manipulering av miljøet de begge to befinner seg i. Direkte kommunikasjon er for eksempel når mennesker snakker sammen. Indirekte kommunikasjon, eller stigmergy, er for eksempel hvis et insekt markerer et område med feromoner, og et annet insekt senere kommer til samme område og reagerer på feromonene.

Theme 1: Non-cooperative game theory

Question 1: Could you classify the different types of game theory?

- Non-cooperative: selvstendige agenter som er self-interested, altså vil de bare maksimere sin egen utility.
- Cooperative: de selvstendige agentene danner coalitions, hvor koalisjonene opptrer som selvstendige enheter. Disse koalisjonene kan også være self-interested, og man kan formulere spørsmål i non-cooperative GT med koalisjoner i stedet for singulære agenter.
- Evolutionary games: payoffs eller utility er ikke konstant, men avhengig av frekvens.
- Behavioural games: felt som ser på forskjellen mellom teori og virkelighet – antar ikke at agenter alltid opptrer optimalt!

Vi har sett mest på de to første.

Question 2: Explain what we mean by self-interested agents

At en agent er self-interested betyr at i ethvert spill eller game ønsker agenten å oppnå utfallet som gir maksimal utility.

Question 3: How do we model utility?

Utility er et reelt tall, som regnes ut fra en funksjon av utfall eller outcome av et spill. For hvert utfall får man altså en viss utility, og høyere utility er bedre. En utility function kan være unik for en agent. Fordi utility er en funksjon av utfall, kan vi rangere hvor bra et utfall ift. andre utfall basert på hvor mye utility det har. Vi har da en preference ordering av outcomes/utfall.

a) Name a few properties of the preference orderings

Preference orderings har disse egenskapene (som i matte):

- Reflexivity: For alle utfall A har vi at $A \geq A$, dvs. at ethvert utfall rangeres som minst like bra som seg selv.
- Transitivity: Hvis utfall $A > B$ og utfall $B > C$, så er $A > C$.
- Comparability: Ethvert utfall kan rangeres mot andre utfall, dvs. av to utfall er de alltid enten like, eller det ene er bedre enn det andre. Enten $A \geq B$ eller $A \leq B$.

b) Discuss the relation between utility and money

Det er viktig at å forstå at penger ikke har et lineært forhold til utility. For eksempel er det enorm økning i utility for et menneske om man starter med 0 kr, og får 1mrd kr. Derimot er økningen i utility proporsjonalt mindre om man starter med 1mrd kr og øker til 2mrd kr; en milliardær har allerede enorm nytteverdi i 1mrd kr, og nytteverdien av å få 1mrd til er proporsjonalt sett mindre.

Question 4: What do we mean by strategic interaction?

Den grunnleggende ideen bak strategic interaction er at hva som er den optimale handlingen for en agent, kommer an på handlingene til alle de andre agentene.

Dette konseptet er spesielt viktig i spill med simultaneous actions, hvor agentene handler samtidig, og ikke kan se hva de andre agentene vil gjøre før de har valgt en handling selv. I sannsynlighetstrær / utfallstrær markerer man treet med en stiplet linje for å vise at en agent ikke vet hva som skjedde i det forrige leddet av treet. Agenten vet altså ikke hvilken situasjon den er i, av de to situasjonene linjen går mellom.

Question 5: Could you name a few solution concepts for solving a 2x2 game on strategic form?

Et løsningskonsept er en ide om hvilket mål en agent bør prøve å oppnå i et spill når han velger strategien sin i spillet. Eksempler er:

- Social welfare: Velg strategien som leder til utfallet som gir høyest sammenlagt utility for alle agentene i spillet.
- Pareto efficiency: Velg en Pareto-optimal løsning. En løsning er Pareto-optimal når det ikke finnes noen annen løsning som er minst like god for alle agentene, og bedre for minst en agent. Eventuelt: ingen forbedring er mulig for en agent, uten å ødelegge for minst en annen agent. Eventuelt 2: Ingen løsning dominerer løsningen.
- Dominant strategy: En strategi er dominant for agent i, hvis den er den beste strategien mot alle mulige strategier for agent j. En slik strategi finnes ikke alltid.
- Nash equilibrium: To strategier er i et Nash-ekvilibrium hvis de velges slik at dersom den andre agenten får vite på forhånd at denne strategien skal spilles, vil ikke den andre agenten bytte strategi; det er ingenting bedre å velge likevel. Med andre ord: to strategier er i et Nash-ekvilibrium hvis de oppfyller:
 - Hvis agent i spiller strategi si, kan spiller j ikke velge noen bedre strategi enn sj.
 - Hvis agent j spiller strategi sj, kan spiller i ikke velge noen bedre strategi enn si.

Question 6: Could you explain the Prisoner's Dilemma?

To fanger skal forhøres for en forbrytelse. Hver av dem har valgene Cooperate (C), og Defect (D). Samarbeid betyr her at de to fangene samarbeider med hverandre ved å ikke røpe noe til politiet, mens å defektere er å sladre på den andre fangen. Dersom begge fangene sladrer (DD), får begge 2 år i fengsel. Dersom en av dem sladrer (DC eller CD), får sladreren gå fri (0 år), mens den andre får 3 år i fengsel. Om ingen av dem sladrer (CC), får hver av dem 1 år i fengsel.

a) Why is it a dilemma?

PD er et dilemma fordi hver fange må vurdere hvor mye tillit de har til at den andre samarbeider. Sett fra fange 1 sitt perspektiv: tror jeg fange 2 defekterer? I så fall bør jeg også defektere for å ikke ødelegge for meg selv. Hvis jeg tror fange 2 samarbeider, burde jeg da defektere for å minimere min fengselsstraff? Hva om fange 2 også har tenkt på dette, og vurderer å defektere fordi han tror at jeg kommer til å defektere?

Dette kan lede til en uendelig tankerekke.

b) Name a few important real-world PDs

Tragedy of the commons: En slags betegnelse på hva som ofte skjer i situasjoner der mange aktører har tilgang til en begrenset, verdifull ressurs; aktørene overforbruker ressursen, som gjør at ressursen går tom/mister verdi, eller skader samfunnet på andre måter. De gjør det fordi det lønner seg å overforbruke for hver aktør individuelt, dersom man antar at andre aktører ikke gjør det samme. Dette er eksempel på PD med flere enn to aktører involvert.

Eksempler på Tragedy of the commons er: overfiske (fiskebestanden krymper så vi mister mat), oljeutvinning (overforbruk forurenses og ødelegger miljøet), overbeiting (beitet blir oppbrukt), osv.

Question 7: How does repeating the PD game affect the outcome?

Ved gjentakelse av PD blir det mer rasjonelt for agentene å samarbeide, så samarbeid vil skje oftere. Dette er fordi å defektere kan lede til at den andre agenten defekterer senere, som over tid ikke lønner seg fordi utility kan regnes over flere runder. Noen tilfeller:

- Infinite rounds of PD: Samarbeid blir rasjonelt for å maksimere utility over tid.
- Fixed number of rounds PD: Som over, men det blir mer rasjonelt å defektere jo nærmere man er slutten (spesielt siste runde), fordi konsekvensene i fremtiden av å ikke samarbeide forsvinner.
- Non-zero probability of future PD round: Om det er rasjonelt å samarbeide avhenger av sannsynligheten for en ny runde, og forventet utility.

Question 8: What do we mean by competitive games and zero-sum games?

Gitt to utfall er et spill mellom agent i og j strictly competitive hvis og bare hvis utfall 1 er strengt bedre for agent i, og utfall 2 er strengt bedre for agent j. Her er det spesielt vanskelig å skape samarbeid, fordi hver gang spillet spilles vil en av agentene få sitt uønskede utfall.

Et spill er zero-sum hvis summen av utility for agentene involvert i spillet, er 0. For to spillere vil det si at om en spiller vinner (får positiv utility), må den andre tape (få tilsvarende negativ utility).

Theme 2: Voting

Question 1: Explain what we mean by social choice theory

Social choice theory = voting theory. Det handler om å ta gruppebeslutninger, i motsetning til individuelle beslutninger. Hvert agent som stemmer vil ta hensyn til hvordan alle de andre agentene kommer til å stemme, og kan derfor stemme taktisk for å øke sannsynligheten for sitt ønskede utfall. Derfor kan group decisions ses på som spill.

Question 2: Explain Plurality voting, its strengths and weaknesses

Plurality: Hver stemmegiver gir sin preference order, altså rekkefølge av foretrukne alternativer, og alternativet som velges først flest ganger, vinner / blir valgt.

Styrker: Enkelt å forstå og implementere.

Svakheter: svært utsatt for strategic manipulation gjennom tactical voting. Også utsatt for Condorcet's paradox, som omhandler scenarier der uansett hvilket alternativ som velges, ville flertallet av velgerne ha foretrukket et annet valg.

Question 3: Explain Simple majority voting, its strengths and weaknesses

Simple majority voting: plurality med kun 2 utfall.

Styrker: Ikke enkelt å manipulere (taktisk stemming er umulig). Condorcet's paradox inntreffer ikke.

Svakheter: ofte er det mer enn 2 utfall, og da kan den ikke brukes som voting procedure.

Question 4: Explain Sequential majority election, its strengths and weaknesses

Sequential majority election: Mange Plurality elections som gjennomføres sekvensielt. Typisk pleier utfallene å deles opp parvis, og man gjør Simple majority voting mellom parene. Disse er enten lineært sekvensielle, eller strukturert som et balansert binærtre.

Styrker: For hvert individuelle Simple majority-valg sett for seg, får man fordelene av simple majority med at Condorcet's paradox ikke inntreffer.

Svakheter: Rekkefølgen valgene gjennomføres i åpner for at stemmegiverne kan manipulere sine stemmer taktisk. Prosedyren er altså sensitiv for hvilken rekkefølge valgene gjennomføres i, som kan bestemmes både tilfeldig og av valgarrangøren.

Question 5: Explain the Borda count

Borda count er et aggregert tall som regnes per utfall, ved å se på rangeringen til utfallet i den ærlige preference ordenen til hver eneste agent i valget. Jo høyere tall, jo høyere ble utfallet rangert i snitt blant velgerne. Det sosiale utfallet / social outcome er når utfallet med høyest Borda count velges.

Question 6: Explain the majority graph

En majority graph konstrueres fra velgernes preferanser slik at hver node er et utfall i valget, og hver kant er rettet slik at pilen går fra en eventuell vinner til en eventuell taper mellom to noder i et parvis valg.

a) What is a possible winner?

En possible winner er et utfall som har mulighet til å vinne i minst en sequential majority election (en slik sekvens må altså finnes).

b) What is a Condorcet winner?

En Condorcet winner er et utfall som vinner parvis mot alle andre mulige utfall. I grafen er det tegnet som en node med kanter pekt ut mot alle andre noder.

Question 7: Explain the Slater rule

Slater rank: Rank av social preference orders, altså en rekkefølge av utfall etter sosial prioritering. Ranken er et mål på hvor langt den er unna hvordan agentene faktisk stemmer. Gitt en majority graph, hvor mange kanter må bytte retning for at grafen samsvarer med en social order? Lavere tall tilsvarer høyere rank, som er ønskelig (samsvarer mer med den opprinnelige grafen).

Slater rule: Velg social ordenen med lavest rank, siden den minimerer forskjell mellom ordenen og grafen.

Question 8: Explain dictatorship

Social welfare function: En funksjon som tar preference orderene til alle velgerne, og produserer en social preference order basert på dem, for eksempel et aggregat av alle velgernes prioriteringer.

Dictatorship: En social welfare function hvor den produserte social preference ordenen settes lik preference ordenen til en spesifikk velger. Det er altså kun en velger som påvirker hva samfunnet velger (en diktator).

Question 9: What does Arrow's theorem say?

Arrow's Theorem: Gitt at velgere har 3 eller flere alternativer, finnes det ingen rangerende valgsystemer som kan konvertere de rangerte preferansene av individer til en social preference order, og fortsatt innfri følgende krav:

- Unrestricted domain: Alle velgerne har lov til å stemme på alle alternativ.
- Pareto efficiency: Når et utfall velges, skal det ikke finnes noe alternativ som er minst like bra for alle velgerne, og bedre for minst en velger. IKKE tilfredsstilt av Sequential majority election (av systemene vi har sett på).
- Independence of irrelevant alternatives (IIA): Social preference mellom utfall i og j er kun avhengig av velgernes individuelle preferanser mellom i og j. Endring i preferanse av andre utfall skal ikke påvirke hvordan i og j sin social preference mellom hverandre er. BARE innfridd av Dictatorship.
- Non-dictatorship: eneste kravet som Dictatorship IKKE innfrir!

Question 10: What is Condorcet's paradox?

Condorcet's paradox: Det finnes scenarier der uansett hvilket utfall som velges, ville en majoritet av velgerne ha foretrukket et annet utfall. Dette kommer an på preference orders til velgerne.

Question 11: What do we mean by tactical voting and strategic manipulation?

Tactical voting: En velger kan produsere et bedre utfall for seg selv ved å strategisk manipulere representasjonen av sine preferanser.

Strategic manipulation: Egentlig et ord som beskriver det samme mer generelt. En social welfare function er manipulerbar hvis det er mulig for en agent å stemme taktisk (manipulere egen preference order) for å øke sin egen utility.

Question 12: Could you give an interpretation of the Gibbard-Satterthwaite theorem?

Gibbard-Satterthwaite (GS) Theorem: Gitt at velgere har 3 eller flere alternativer, sier teoremet at det IKKE finnes noen stemmeprotokoll som er ikke-manipulerbar,

unntatt diktatur. Det vil si at man i teorien kan stemme taktisk ved ethvert ikke-diktatorisk valg.

Likevel kan dette være vanskelig å regne ut, og informasjonen som trengs for å regne på det kan være vanskelig å skaffe.

Question 13: Why could the Second-order Copeland be an advantageous election scheme in terms of strategic manipulation?

Second-order Copeland er eksempel på en stemmeprotokoll hvor det er vanskelig å regne ut hvordan man kan strategisk manipulere valget, selv om det er teoretisk mulig.

Theme 3: Cooperative game theory

Question 1: Classify the different types of game theory

- Non-cooperative: selvstendige agenter som er self-interested, altså vil de bare maksimere sin egen utility.
- Cooperative: de selvstendige agentene danner coalitions, hvor koalisjonene opptrer som selvstendige enheter. Disse koalisjonene kan også være self-interested, og man kan formulere spørsmål i non-cooperative GT med koalisjoner i stedet for singulære agenter.
- Evolutionary games: payoffs eller utility er ikke konstant, men avhengig av frekvens.
- Behavioural games: felt som ser på forskjellen mellom teori og virkelighet – antar ikke at agenter alltid opptrer optimalt!

Vi har sett mest på de to første.

Question 2: Explain what we mean by cooperative game theory, especially in relation to non-cooperative game theory

Cooperative Game Theory prøver å svare på: hvilke coalitions blir formert av self-interested, rasjonelle agenter, og hvordan fordeles utility mellom dem innenfor en coalition. Generelt svarer dette til å se etter stable coalitions, altså koalisjoner hvor medlemmene ikke vil ønske å bryte ut av koalisjonen eller danne andre koalisjoner.

Question 3: Describe and explain the formal description of a cooperative game

Et cooperative game er et par $G = \langle Ag, v \rangle$, hvor Ag er alle agentene, og v er characteristic function, som gir en numerisk verdi til alle mulige koalisjoner. Denne verdien er total utility for koalisjonen, og må fordeles på alle medlemmene av den. Man antar alltid at den karakteristiske funksjonen er gitt, men at det ikke er gitt hvordan utility fordeles blant koalisjonsmedlemmer.

Question 4: What is the core?

Coalition Core: Kjernen er settet/mengden av utfall for en koalisjon, som ingen annen koalisjon har objection til. Gitt en koalisjon C med utfall x vil det si at om det finnes noe mulig utfall x' for en annen koalisjon C' slik at utility blir strengt høyere for hver eneste agent involvert i x' , vil C' objecte til C .

Stable Core: C sin kjerne er stabil hvis den er ikke-tom. Det vil si at ingen andre potensielle koalisjoner C' kan ha noe bedre utfall for en gruppe av agentene i C , enn utfallene i C sin kjerne.

Question 5: In relation to cooperative games, what is the Shapley value?

Shapley value: den gjennomsnittlige marginal contributionen til en agent i en gitt koalisjon C , over alle mulige rekkefølger av agenter når de legges til i C i en viss rekkefølge. Shapley value er symmetrisk, så to spillere som gir samme contribution til en koalisjon skal få tildelt samme utility. Shapley-verdien er additiv, så agenter som spiller to spill får summen av de to spillene. Shapley-

verdien er også slik at dummy players, spillere som ikke bidrar til koalisjoner, ikke får mer utility enn det de kan lage selv.

Question 6: Explain marginal contribution net

Marginal contribution net: En utvidelse av induced subgraph som gjør at representasjonen av characteristic function blir komplett. Ikke en tegning/graf, men et sett med nedskrevne regler som man kan utlede grafen (og alle marginal contributions) fra.

Question 7: Explain the induced subgraph

Induced subgraph: Graf hvor hver agent er en node, og de urettede kantene mellom dem er verdien de tilføyer i tillegg til egen verdi i en koalisjon hvor de begge inngår. Grafen kan gjøre det lettere å forstå og regne på Shapley-verdier for agentene, men er ikke komplett.

Question 8: What are simple games?

Simple games: Spill med koalisjonsverdier 0 eller 1 (en koalisjon kan enten vinne eller tape). Settet av vinnerkoalisjoner benevnes som W . Noen egenskaper:

- Non-triviality: noen koalisjoner er vinnere (ikke ingen), men ikke alle koalisjoner er det.
- Monotonicity: Hvis C_1 er inneholdt i C_2 , og C_1 er en vinner, så er også C_2 en vinner. Alle supersets av en vinnerkoalisjon C er også vinnere.
- Zero-sum: Hvis koalisjon C vinner og agent i er utenfor C , så vinner ikke agent i .
- Den tomme koalisjonen er ikke en vinnerkoalisjon.
- Grand coalition vinner, altså koalisjonen av alle agentene summert vinner.

Question 9: What are weighted voting games?

Weighted voting games: En enklere representasjon av mange simple games. I stedet for å representere hver agent individuelt ser vi på “blocks” av agenter som stemmer likt, og gir hver agent en vekt w_i . Vi setter en kvote q slik at hvis summen av $w_i > q$ i en koalisjon C , så er det en vinnerkoalisjon W . I simple majority voting (2 alternativer) settes q til å være lik en mer enn halvparten.

Shapley-verdien er NP-hard å regne ut. Kjernen er mye enklere å regne ut, altså se etter ikke-tomme stabile koalisjoner.

Question 10: What is coalitional structure formation?

Coalitional structure formation: Koalisjoner formes sentralt (en designer sitter med makt over alle noder/agenter). Designeren ønsker å maksimere social welfare, altså total/aggregert utility i en koalisjon.

Problemet i Coalitional structure formation er altså å splitte opp alle agentene i koalisjoner som maksimerer total utility skapt av alle agentene. Antallet kombinasjoner av koalisjoner vokser eksponentielt med antallet agenter. Problemet er derfor vanskelig å løse. Søkealgoritmer kan brukes for å sjekke mulige sammensetninger av koalisjoner.

Theme 4: Auctions

Question 1: What are auctions?

Auction: en auksjon er en mekanisme for å allokere sjeldne ressurser, enten en per auksjon eller flere ressurser av gangen, til agenter. Hver ressurs må være sjelden, og ønsket av mer enn en agent (hvis ikke er problemet trivielt).

a) In what way are auctions thought to be efficient?

De antas å fordele godene til de agentene som verdsetter dem høyest.

b) What two main types of auctions are there?

Single item auction: en ressurs selges på auksjon.

Combinatorial auction: flere ressurser selges samtidig, gjerne i grupper/bundles definert av auctioneer/auksjonsholderen.

Question 2: Explain the English auction and motivate optimal strategy

First-price: det høyeste budet er prisen som blir betalt av vinneren.

Open-cry: agentene kan se hverandres bud.

Ascending auction: etter et startbud til en startpris fastsatt av auctioneer, kan de andre agentene gi høyere bud. Agenten som ga det høyeste budet vinner når ingen andre bud legges frem.

English auction: First-price, open-cry, ascending auction.

Den beste strategien gitt rasjonelle agenter er å alltid by så vidt mer enn hva som er det tidligere høyeste budet til de andre agentene, frem til agenten når sin personlige verdsettelse av ressursen som bys på. Hvis budene overstiger personlig valuering, burde agenten trekke seg / ikke by mer.

Question 3: Explain the Dutch auction

Descending auction: Auctioneer gir en høy startpris, og tilbyr så lavere og lavere priser helt til en agent kjøper på den tilbudte prisen. Den første agenten som aksepterer prisen, får ressursen til den prisen.

Dutch auction: first-price, open-cry, descending auction. Agenter ønsker å by så langt under true valuation som mulig, men det kommer an på de andre agentene.

Question 4: Explain the first-price, sealed-bid auction and motivate optimal strategy

One-shot: hver agent legger inn sitt bud på forhånd, og ett bud velges (for eksempel det høyeste, eller nest høyeste).

Sealed-bid: agentene kan ikke se hverandres bud.

First-price, sealed-bid auctions: også one-shot. Ressursen gis til agenten som byr høyest, og agenten må betale prisen han ga i det høyeste budet.

Som med Dutch auction, ønsker agentene å by så langt under true valuation som mulig, men ikke så lavt at andre byr høyere. Kommer også an på de andre agentene.

Question 5: Explain the Vickrey auction and motivate optimal strategy

Second-price: Agenten som vinner (bød mest) må betale prisen som ble gitt i det nest høyeste budet.

Vickrey auction: second-price, sealed-bid, one-shot.

Tanken bak Vickrey er at agenten motiveres til å by sin true valuation. Som i andre tilfeller ønsker agenten ikke å by mer enn true valuation, fordi ressursen ikke er verdifull nok til noen høyere pris. Men hvis agenten byr mindre enn true valuation, øker han risikoen for å ikke få ressursen, selv om hvis han vant auksjonen, er prisen han betaler uavhengig av prisen han bød på uansett.

Question 6: What is winner's curse?

Winner's curse: en vinner som har vunnet en auksjon, og som ikke vet den sanne verdien av ressursen han vant, risikerer å ha overvaluert ressursen med prisen han betaler.

Question 7: What do we mean by combinatorial auctions?

Combinatorial auction: flere ressurser selges samtidig, gjerne i grupper/bundles definert av auctioneer/auksjonsholderen.

a) How do we solve this combinatorial problem?

Vi ønsker å velge kombinasjonen av vinnere som totalt sett verdsetter alle ressursene mest, når alle har blitt bydd på, og at disse byr sin true valuation. Vi ønsker altså å maksimere social welfare. Å representere og kalkulere dette er NP-hardt.

b) How do we represent complex bids?

Vi representerer agents true valuation gjennom atomic bids $\beta = (Z, p)$ der Z er en bundle (subsett av alle ressurser som kan bys på) definert av agenten selv, og p er prisen agenten er villig til å betale for Z . Et supersett Z' av Z vil altså ha samme valuering for en agent som Z alene, med mindre Z' også inneholder andre ressurser (eller kombinasjoner av ressurser) som agenten også verdsetter.

XOR bids: Agenten er kun villig til å betale for maks 1 av bundlesene han har verdsatt, og betaler for den bundlen som har høyest verdi for ham.

OR bids: Agenten er villig til å betale for flere bundles, og betaler den høyeste summen av hans egendefinerte atomic bids som er mulige å by på.

Question 8: Explain the VCG-mechanism

VCG er en generalisering av Vickrey auctions til det kombinatoriske tilfellet, og er incentive compatible, altså gjør den det rasjonelt for agenter å by sin true valuation. Ideen er at hver vinner betaler en kompensasjon tilsvarende “lost utility” av alle de andre spillerne som mistet utility fordi den vinneren spilte (og forhindret dem fra å vinne).

Theme 5: Bargaining

Question 1: Explain what we mean by bargaining

Bargaining / negotiation: Agenter forhandler for å nå avtaler, spesielt angående konflikter rundt mål eller preferanser.

Resource division: en ressurs “kommer fra himmelen”, og agenter skal forhandle om hvordan de fordeler ressursen mellom seg. Her finner man ultimatum game, hvor agentene etter tur foreslår avtaler, og den andre agenten kan akseptere eller avslå, og eventuelt ende i conflict deal, som regnes som det verste utfallet.

Task allocation: agentene har visse oppgaver med en cost function som forteller hvor mye en task koster å gjøre for en gitt agent. De prøver å avtale hvordan å omfordele oppgavene sånn at alle øker utility / minimerer kostnad.

Resource allocation: agentene har fordelte ressurser, med value function som forteller hvor mye en agent verdsetter en gitt ressurs. De foreslår avtaler for å omfordele ressursene, hvor side payments tillates.

Question 2: What is the conflict deal?

Conflict deal: Betegnelse på utfall i forhandlinger der ingen enighet nås. Den regnes som det verste mulige utfallet for alle involverte agenter, også verre enn en avtale hvor man får ingenting.

Question 3: Explain the alternating offer bargaining model, assuming some simplifying assumptions.

Alternating offer bargaining model: anta single issue (kun et gode forhandles om), symmetric (agentene antas å ha like forutsetninger i forhandlingene) og one-to-one (kun to agenter involvert). Da er alternating offer-modellen at agent 1 gir et tilbud, og agent 2 kan godta eller avslå, eventuelt med et mottilbud. Agent 1 kan godta mottilbudet eller gi eget mottilbud, osv.

Question 4: Explain the ultimatum game

Ultimatum game: Et spill som følger alternating offer bargaining model, og er tilfellet single-issue, symmetric, one-to-one. En "pai" skal fordeles, x til spiller 1 og $1-x$ til spiller 2. Spiller 1 foreslår en fordeling, og spiller 2 kan enten godta avtalen, eller avslå, og enten få conflict deal eller by selv avhengig av hvor mange forsøk man har i spillet.

a) What is the Nash equilibrium in one-shot ultimatum?

Nash equilibrium: $(1, 0)$. Siden conflict deal er det verste utfallet mulig for agent 2, som må akseptere eller avslå tilbudet, vil en rasjonell agent 2 alltid godta enhver avtale. Dersom agent 1 vet dette, vil en rasjonell agent 1 da alltid tilby en avtale som maksimerer egen utility / andel av ressursene.

b) What is the Nash equilibrium in two-shot ultimatum?

Nash equilibrium: $(0, 1)$. Tilsvarende som over, men denne gangen gir agent 2 det siste budet, og har derfor makten til å forsere at alle ressursene går til ham.

c) What is the Nash equilibrium in unlimited round ultimatum?

Nash equilibrium: $(1, 0)$, dersom spiller 1 kan overbevise spiller 2 om at han vil avslå alle andre tilbud. Sett fra spiller 2 sitt perspektiv: dersom man virkelig tror på at spiller 1 aldri godtar annet enn $(1, 0)$, kan man like gjerne godta hva enn han tilbyr, siden det er bedre enn conflict deal.

Question 5: What happens when the players get impatient in the ultimatum game?

a) What is the Nash equilibrium in impatient one-shot ultimatum?

Utålmodighet påvirker ikke one-shot equilibrium (forblir $(1, 0)$), siden tålmodigheten ikke rekker å renne ut i et slikt tilfelle.

b) What is the Nash equilibrium in impatient two-shot ultimatum?

Nash equilibrium: $(1 - \delta_2, \delta_2)$. δ_2 er tålmodigheten til spiller 2, så $\delta_2=1$ betyr at spiller 2 er tålmodig, og får alle ressursene. $\delta_2=0$ betyr at spiller 2 er utålmodig og godtar første tilbud uansett hva.

c) What is the Nash equilibrium in impatient unlimited round ultimatum?

Nash equilibrium:

$((1 - \delta_2) / (1 - \delta_1 \delta_2), \delta_2 (1 - \delta_1) / (1 - \delta_1 \delta_2))$.

Nash-equilibrium etableres allerede i første runde.

Question 6: Explain the negotiation decision function as shown in the figure 15.3

Negotiation decision function: Funksjon som modellerer en agents strategi i forhandlinger (hvor mye er jeg villig til å kjøpe/selge for) som en funksjon av tid (tid kan være verdifullt i seg selv).

Boulware: Strategi om at man ikke flytter mye på akseptert pris før tiden nesten er ute, men godtar verre og verre tilbud raskt når tiden snart er ute (kjøper for mer / selger for mindre). Får ofte god pris, men risikerer å vente lenge for det.

Conceder: Agenten tillater verre tilbud raskt, og er allerede nærme laveste aksepterte pris lenge før deadline. Risikerer å få dårlig avtale, men får avtalen gjennom raskere enn Boulware.

Fra figuren med piler nedover ser vi hva Boulware og Conceder er villige til å selge for over tid. Fra figuren med piler oppover ser vi hva de er villige til å kjøpe for over tid. Conceder-Boulware vil nå avtale middels-tidlig med god avtale for Boulware. B-B vil nå avtale sent med lik fordeling. C-C vil nå avtale tidlig, også med lik fordeling.

Question 7: Explain the bargaining for task allocation in relation to figure 15.4

Figuren viser verdien av mulige utfall som en funksjon av utility for agent i på en akse, og agent j på den andre (tenk enhetssirkelen). Kun i øverste høyre kvadrant får både i og j positiv utility fra en ny allokering av tasks, og kanten av sirkelen i dette hjørnet er linjen med Pareto-optimale løsninger, altså løsninger hvor det ikke finnes noe bedre alternativ for den ene agenten uten at det er verre for den andre. Origo tilsvarer conflict deal.

a) Describe the monotonic concession protocol

Monotonic concession protocol: agentene foreslår mulige avtaler samtidig. En avtale inngås når agent i foreslår en avtale som gir agent j høyere utility enn j får i avtalen j selv foreslo, eller motsatt vei. Hvis avtale ikke inngås, foreslår begge agentene ny avtale med krav om at utility for den andre agenten må være minst like høyt som hva de tilbød i sitt forrige bud, og vise versa. Hvis ingen konsesjoner/lettelser tilbys av noen agent, inngås conflict deal.

Dette er verifiserbart og garantert å ta slutt, og kompleksiteten til avtalene øker eksponensielt med antall mulige tasks å allokere.

b) Describe the Zeuthen strategy

Zeuthen: agentenes første forslag bør være den optimale avtalen. Hvilken agent som går med på den andres tilbud er avhengig av risk = utility i mister av å akseptere j sitt tilbud / utility i mister av å forårsake conflict deal. Risk måler en agents risikovilje. Agenten med minst risk etter en runde bør tilby nok til å få den andre agenten til å gå med på avtalen i neste runde.

Zeuthen terminerer, men kan skape conflict deal. Outcome er Pareto-optimalt, men ikke alltid Social optimum. Individuelt rasjonelt hvis enighet nås. Ingen tredjepart trengs for å overvåke forhandlingene. Strategisk utfall er Nash; om en spiller spiller Zeuthen-strategien, bør den andre også spille det.

Question 8: What is the basis for many-to-many negotiations of exchanging already endowed goods?

For resource allocation av allerede utdelte ressurser, kan agenter tilby avtaler med omfordeling av ressurser, potensielt inkludert side payments. Man har typisk definert et termination criterium: gjerne at avtalen må være PO eller SO. En rasjonell agent godtar bare tilbud som gir høyere utility enn hva man hadde før. Protokollen er som følger:

Innledende fordeling defineres som nåværende. Alle agenter kan tilby en ny allokering med side payments. Hvis alle agenter er enige med denne avtalen ifølge individuell rasjonalitet, og avtalen innfrir termination criteria, så implementeres den med side payments som definert. Hvis termination criteria ikke nås, men de er enige, settes nåværende allokering til den foreslåtte, og forhandlingene fortsetter. Hvis en av agentene er uenige, fortsetter forhandlingene med den gamle allokeringen.

Theme 6: Arguing

Question 1: What is arguing?

I multiagentsystemer trenger agenter gjerne å bli enige om en felles oppfatning.

Arguing: prosessen å nå en rationaly justifiable position basert på argumentene

alle agentene setter frem. Man ønsker at alle agentene skal nå samme rationally justifiable meningssett.

Rationally justifiable position: Alle meningene en agent holder, må være consistent (en mening kan ikke motsi en annen).

Question 2: What different modes of arguing exist?

De ulike modusene kan minne om appellformer i retorikk.

Logical mode: Logos, matematisk deduksjon av typen "hvis A, og $A \rightarrow B$, så B". Analytisk arguing.

Emotional mode: Patos, følelsesmessig appell og argumentasjon: "hvordan hadde det følt om det skjedde med deg?".

Visceral mode: fysisk, sosialt aspekt av menneskelig argumentasjon, av typen: trampe i bakken for å vise en sterk følelse.

Kisceral mode: appellere til intuisjon, mystikk, religion, metafysikk etc. Kan være noe overtroisk.

Vi forholder oss til logical mode, siden det er mest systematisk og pålitelig.

Question 3: What types of argumentation system exist?

Abstract argumentation system: Argumenter er udelelige, atomiske ting med visse relasjoner. Systemorientert fokus; hvilke argumenter kan høre sammen? Dung-style er abstrakt.

Deductive argumentation system: Man tar hensyn til meningen av forskjellige argumenters relasjoner til hverandre. Fokuset er på å finne ut hva relasjonene mellom argumentene faktisk er.

Question 4: Could you explain a Dung style argumentation system?

Dung-style system: Man lager en graf med argumenter som noder, og relasjoner mellom dem som rettede kanter. En pil fra A til B betyr "A angriper B", altså at hvis vi aksepterer A må vi avvise B. Om A angriper B og B angriper C, har vi at A forsvaret C.

Question 5: Which arguments to accept in a Dung style argumentation system?

Det er ingen universelt akseptert definisjon av hva som er et akseptabelt meningssett. Men Dung har foreslått preferred og grounded extensions som noen alternativer.

Conflict free: Et sett av argumenter som ikke angriper hverandre.

Mutually defensive: For hvert argument som blir angrepet i et mutually defensive sett, er det minst ett annet argument som forsvaret det argumentet.

Admissible set: Et sett som er både conflict free og mutually consistent, derav internally consistent.

Preferred extension: Et admissible set hvor det ikke er mulig å legge til flere meninger uten at det slutter å være admissible.

Grounded extension: Alternativ definisjon av aksepterbare sett. Begynn med uangrepede meninger (alle er jo enig med dem), og vrak alle meninger som de angriper. Fortsett rekursivt å finne hvilke meninger som er "in" og "out" av en grounded extension, og behold alle som er "in".

Question 6: Which arguments to choose if multiple preferred extensions?

Skeptical acceptance: Hvis et argument er i alle preferred extensions, er det skeptically accepted.

Credulous acceptance: Hvis et argument er i minst en preferred extension, er det credulously accepted.

Det er naturlig å foretrekke argumenter som aksepteres av flest perspektiver.

Question 7: Could you describe deductive argumentation?

Deductive argumentation: Man bygger opp argumenter fra logiske relasjoner slik at a impliserer b, som vil si at med a som premiss er det en sekvens av inferenser som leder til en konklusjon b. Ethvert argument er bygd opp av en Sentence og Grounds, hvor sentence er konklusjonen, og grounds er inferenser som er finnes i en og samme database. Argumenter kan angripe hverandre gjennom denne strukturen.

a) What is a rebuttal and an undercut?

Rebuttal: Argument (sentence1, grounds1) rebutterer (sentence2, grounds2) hvis sentence1 = NOT sentence2, det vil si setningene er i direkte konflikt.

Undercut: Argument (sentence1, grounds1) underkutter (sentence2, grounds2) hvis sentence1 = NOT grounds2, det vil si setningen til det ene argumentet strider med premissene til det andre argumentet.

Begge disse er eksempler på angrep mellom argumenter. For å velge en vinner kan vi gå over til abstract argumentation igjen, siden vi nå har identifisert angrep / relasjoner mellom argumenter.

Theme 7: Classical swarm intelligence

Question 1: Explain the binary bridge experiment

Det går en bro som splitter seg i to grener og senere slår seg sammen igjen. På enden av broen er en ressurs (mat). På den andre enden begynner maur å gå for å komme seg til maten, og slipper feromoner på veien. Etter hvert som det blir flere feromoner på den ene grenen av broen, vil flere og flere maur velge å gå til denne siden.

Question 2: Explain the Ant Colony Optimization metaheuristic algorithm

ACO tar inspirasjon fra bl.a. binary bridge ved at sannsynligheten for at en agent velger en retning, påvirkes av hvor mange andre agenter som har gått der før (hvor mye feromoner som finnes).

a) What type of problems are ACO especially applicable to? Could you name one?

ACO passer godt til diskrete / kombinatoriske problemer som Travelling Salesman Problem eller andre grafproblemer.

Question 3: Explain the pheromone update rule in AS

The pheromone update rule:

$$\tau_{ij} \leftarrow (1 - \rho)\tau_{ij} + \sum_{k=1}^m \Delta\tau_{ij}^k$$

where τ_{ij} is pheromone concentration on edge (i, j)

ρ is evaporation rate

m is number of ants

$\Delta\tau_{ij}^k$ is pheromones laid on edge (i, j) by ant k

Vi ser at hvor mye feromoner som finnes på en kant i en graf i et timestep, påvirkes av hvor mye feromoner som var der fra før av, hvor fort gamle feromoner avtar, og hvor mye feromoner som maur/agenter har lagt på i timesteppet.

The pheromone update rule:

$$\Delta\tau_{ij}^k = \begin{cases} \frac{Q}{L_k} & \text{if ant } k \text{ used edge } (i, j) \text{ on its tour} \\ 0 & \text{otherwise} \end{cases}$$

where Q is a constant

L_k is the length of the tour constructed by ant k

Vi ser også at mengden feromoner lagt på av en maur er avhengig av lengden på touren som mauren lagde. Lengre turer betyr at feromonene er mer spredd, og det blir mindre per edge. Dette fungerer godt for TSP hvor målet er å velge så kort vei som mulig, siden det gjør det mer sannsynlig at flere maur velger korte veier.

Question 4: Explain the transition rule (the probability of going to node j) in AS

The transition rule (the probability of going to city j):

$$p_{ij}^k = \begin{cases} \frac{\tau_{ij}^\alpha \cdot \eta_{ij}^\beta}{\sum_{c_{il} \in N(s^p)} \tau_{il}^\alpha \cdot \eta_{il}^\beta} & \text{if } c_{ij} \in N(s^p) \\ 0 & \text{otherwise} \end{cases}$$

where $N(s^p)$ is the set of feasible components (cities not visited yet)

α, β are nonlinear control parameters

$\eta_{ij} = \frac{1}{d_{ij}}$ is invers distance between city i and j

Vi ser at hvor en maur velger å dra er avhengig av hvor mye feromoner som er på hver mulige sti, og hvor lang stien er: mer feromoner øker sannsynligheten, og høyere distanse minsker sannsynligheten. Sannsynligheten regnes ut som softmax av alle stienes feromoner og avstander.

Question 5: Explain the Canonical Particle Swarm Optimization metaheuristic algorithm

PSO parameters:

t is iterations or time (i.e. one increment in while-loop)

$i = \{1, 2, \dots, N\}$ is particle

$d = \{1, 2, \dots, M\}$ is dimension

v_{id} is velocity of particle i in dimension d

x_{id} is position of particle i in dimension d

p_{id} is the best position of particle i in dimension d

p_{gd} is the best position of all particles in dimension d

UpdateParticleVelocity:

$$v'_{id} = \underbrace{w \cdot v_{id}}_{\text{Inertia term}} + \underbrace{w_1 \varphi_1 (p_{id} - x_{id})}_{\text{Cognition term}} + \underbrace{w_2 \varphi_2 (p_{gd} - x_{id})}_{\text{Social term}}$$

where v'_{id} is updated velocity of particle i in dimension d

φ_1 and φ_2 are uniform random variables $[0,1]$

w, w_1 and w_2 are parameters that need to be tuned

PSO forutsetter at en oppgave skal løses i et kontinuerlig søkefelt, for eksempel finne punktet med høyest respons på en sensormåling, slik at hvert punkt kan tilordnes en verdi. PSO fungerer slik at hver partikkel (agent) sin fart/retning påvirkes av 3 faktorer:

Inertia term: Fartsretningen til partikkelen fra forrige timestep.

Cognition term: partikkelen vil i noen grad bevege seg mot sitt eget beste kjente punkt fra før.

Social term: partikkelen vil i noen grad bevege seg mot det beste kjente punktet blant alle partiklene samlet.

a) What type of problems are PSO especially applicable to?

PSO er best for kontinuerlige søkefelt, i motsetning til ACO som passer bedre til det diskrete domenet.

Question 6: What is taxis?

Taxis er bevegelsen til en organisme som direkte respons av stimuli, for eksempel fluer som flyr mot lyset, generelt i retning mer stimuli.

a) Could you explain how taxis could be an inspiration for swarm robotics?

I et kontinuerlig søkefelt hvor man ikke har informasjon om stimuli i andre områder enn punktet man er på, fungerer taxis (for eksempel chemotaxis, thermotaxis, phototaxis etc.) som en søkealgoritme som ligner på gradient descent, men med mer tilfeldighet og mer exploration.

Man går i en retning i et timestep – dersom stimulusen er høyere enn i forrige timestep, fortsett i omtrent samme retning. Hvis ikke, snu til omtrent motsatt retning. Hvis en sverm gjør dette i et søkefelt, spesielt med passende dekning / avstand mellom agentene, kan dette lede til effektiv søking.

Question 7: What is Artificial Potential Field?

Motivasjonen til APF er å forhindre at agenter kolliderer med hverandre og hindringer i miljøet. Dette gjøres ved å simulere frastøtende krefter fra alle agenter og hindringer (som like poler fra to magneter), slik at agentene automatisk unngår hindringer. Dette kan enten preprogrammeres off-line eller plasseres i miljøet dynamisk, on-line mens agentene utforsker.

Theme 8: Response threshold

Question 1: Could you explain the inspiration behind response thresholds in task allocation?

Generelt ønsker man ofte å oppnå spesialisering av oppgaver blant agenter, for å koordinere oppgaveløsning effektivt.

Response thresholds svarer til sannsynligheten til at en agent responderer på stimuli fra en oppgave. Hvis stimulien passerer en terskel / threshold, vil agenten mest sannsynlig gjøre den oppgaven. Ulike agenter kan ha forskjellige thresholds (lavere eller høyere).

Question 2: How do we model response thresholds in task allocation?

Det finnes flere matematiske modeller for response thresholds. Disse gjelder for både kontinuerlig og diskret tid:

$$T_{\theta}(s) = \frac{s^n}{s^n + \theta^n}$$

where s is stimulus

T is probability of doing task T in response to s

θ is threshold

n is steepness of threshold ($n > 1$)

$$T_{\theta}(s) = 1 - e^{-s/\theta}$$

where s is stimulus

T is probability of doing task T in response to s

θ is threshold

De to gir ganske lignende grafer. Vi ser at om $s \gg \theta$ blir sannsynligheten for at agenten velger å gjøre task T cirka lik 1, mens om $s \ll \theta$ blir den cirka lik 0. Merk at man gjerne pleier å dele inn response threshold-funksjoner per type arbeidere, altså kan en threshold deles av mange arbeidere av samme type.

Question 3: Could you describe and explain the stimulus dynamics (assuming one task)?

Discrete time:

Stimulus of task (discrete-time dynamics)

$$s(t+1) = s(t) + \delta - \alpha \frac{N_{act}}{N}$$

where s is stimulus of task T at time t

δ is the increase in stimulus intensity per unit time

α is a scale factor measuring the efficiency of task performance

N_{act} is number of active individuals

Continuous time:

Stimulus dynamics (continuous-time):

$$\delta_t s = \delta - \alpha \frac{N_1 + N_2}{N} = \delta - \alpha f x_1 - \alpha (1 - f) x_2$$

since $(N_1 + N_2)/N = f x_1 + (1 - f) x_2$.

Fra første ligning ser vi at endring i stimuli fått fra å gjøre task T i et timestep, er avhengig av hva stimulien i forrige timestep var. Den øker gjennom delta, som er "passiv" økning i stimuli bare fra å gjøre oppgaven, og minsker basert på hvor mange agenter som er aktive. Flere aktive agenter forminsker stimulien mer enn om mange agenter er inaktive (dette gir mening fordi det er mer nødvendig at en agent gjør en oppgave dersom få andre agenter gjør egne oppgaver).

Den andre ligningen modellerer det samme som den første, men i kontinuerlig tid. Bokstaven f står for andelen av arbeidere av type i totalt, og x står for andelen av arbeidere av type i som gjør oppgave T .

Question 4: Could you describe and explain the transition probabilities (assuming one task)?

The transition probabilities (discrete-time dynamics):

$$P(X_i = 0 \rightarrow X_i = 1) = T_{\theta_i}(s)$$

$$P(X_i = 1 \rightarrow X_i = 0) = p$$

where i is worker type

$X_i = 0$ means that agent of type i is inactive in task T

$X_i = 1$ means that agent of type i is active in task T

p is probability of an agent of type i gives up task T

$1/p$ is average time spent on task T , independent of s

Transition dynamics (continuous-time):

$$\underbrace{\delta_t x_i}_{\text{Change in active workers}} = \underbrace{T_{\theta_i}(s)(1 - x_i)}_{\text{Inactive workers recruited}} - \underbrace{p x_i}_{\text{Retired workers}}$$

where n_i is number of workers of type i , i.e. $N = \sum n_i$

N_i is number of workes of type i engaged in task T

$x_i = \frac{N_i}{n_i}$ is fraction of workers type i doing task T

$f = \frac{n_i}{N}$ is fraction of workers type i in colony

Vi ser at andelen arbeidere av type i som gjør task T øker når sannsynligheten for type i -arbeidere for å gjøre T øker, og det er mange arbeidere av type i som ikke gjør T allerede. Andelen minsker når det er en stor andell arbeidere av type i som allerede er engasjert i T , og sannsynligheten for å gi opp på oppgaven er høy.

Question 5: How would you alter this transition dynamics (continuous time) to allow for more than one task?

Transition dynamics of several tasks (continuous-time):

$$\delta_t x_{ij} = \frac{s_j^2}{s_j^2 + \theta_{ij}^2} (1 - \sum_{k=1}^m x_{ik}) - p x_{ij}$$

where j is one out of m possible tasks

n_i is number of workers of type i , i.e. $N = \sum n_i$

N_{ij} is number of workes of type i engaged in task T_j

$x_{ij} = \frac{N_{ij}}{n_i}$ is fraction of workers type i doing task T_j

Kort sagt, man må modifisere formelen slik at man tar hensyn til alle arbeiderne av type i som gjør, og ikke gjør, oppgave j til enhver tid, siden disse faktorene påvirker hvor mange agenter som vil begynne og slutte på å gjøre oppgaven. Man må også sjekke stimulus for hver enkelte oppgave, siden det nå er separate stimuli per oppgave.

Det samme prinsippet gjelder også om man skulle modellert stimulus dynamics (endring i stimulus) for flere tasks.

Question 6: How would you alter this transition dynamics (continuous time) to account for specialization through learning?

Man måtte først innført at threshold-verdien også endrer seg kontinuerlig (hittil har den stått stille). Dette modellerer at ettersom arbeidere av en type jobber mer med oppgave T, blir de flinkere på T (de lærer av erfaring):

Reinforcement of response threshold:

$$\theta_{ij} \leftarrow \theta_{ij} - \underbrace{x_{ij}\xi\Delta t}_{\text{Performing task } j} + \underbrace{(1 - x_{ij})\varphi\Delta t}_{\text{Not performing task } j}$$

where Δt is time interval of evaluation

ξ (ksi) is constant decreasing θ_{ij} if task j is performed

φ is constant increasing θ_{ij} if task j is not performed

x_{ij} is fraction of time spent on performing task j

Vi ser at jo større andel av arbeidere av type i som gjør task j , jo raskere faller threshold-verdien, og motsatt; jo mindre andel, jo raskere øker threshold. Vi ser også at det er egne parametre som representerer hvor fort task j læres (ξ), og hvor fort task j glemmes (φ), som bidrar i endringen.

Samme i kontinuerlig tid:

Reinforcement of response threshold:

$$\delta_t \theta_{ij} = [(1 - x_{ij})\varphi - x_{ij}\xi] \Theta(\theta_{ij} - \theta_{min}) \Theta(\theta_{max} - \theta_{ij})$$

where Θ (theta) is a step function for maintaining θ within bounds

Vi får til slutt følgende endring i transition dynamics, som inkluderer læring (gjennom theta):

Transition dynamics of specialization (continuous-time):

$$\delta_t x_{ij} = \frac{s_j^2}{s_j^2 + \theta_{ij}^2} (1 - \sum_{k=1}^m x_{ik}) - p x_{ij} + \psi(i, j, t)$$

where $\psi(i, j, t)$ is a $N(0, \sigma)$ stochastic process simulating that individuals are in different environments.

Theme 9: Swarm robotics 1

Question 1: What is swarm robotics?

Swarm robotics: Swarm Intelligence anvendt på roboter.

Question 2: Could you characterize swarm performance as a function of robots?

a) What processes affect performance in the inference region?

To prosesser påvirker performance:

- Contention / inference: kostnad eller gevinst/utility fra å fordele begrensede ressurser innad i svermen.
- Lack of coherency: Svermen deler ikke informasjon optimalt pga diverse begrensinger i kommunikasjon (fysiske begrensninger, eller de har dårlig strategi i hva de kommuniserer, etc).

b) Could you model this general swarm performance using the USL?

The Universal Scalability Law [Gunther, 1993] of parallel processing systems:

$$R(N) = c \frac{N}{1 + \alpha(N-1) + \beta N(N-1)}$$

where $R(N)$ is performance as a function of N processors

α is degree of contention (inference)

β gives the lack of coherency in the distributed data

C is a scalar

Fra USL har vi at om vi ikke betrakter contention/inference eller lack of coherency, vil performance øke ettersom antallet roboter i svermen øker. Manglende coherency gjør at performance synker kvadratisk, så det er den største bottleneck-en i formelen (representert ved beta, som regel ikke-null positivt tall). Om det å dele ressurser gjøres suboptimalt (positiv alfa / høy grad av contention), vil også contention bidra til å sakke ned økning i performance. Men om alfa er negativ, altså at deling av ressurser gir økt total performance, vil vi få super-linear speed-up, altså bedre enn lineær performance. Dette kan sies å være eksempel på emergens.

Question 3: Could you explain swarm modelling as a series of mappings?

A swarm system of size N in 2D space can be described by the state vector:

$$\gamma = (x_1, x_2, \dots, x_N, v_1, v_2, \dots, v_N, s_1, s_2, \dots, s_N)$$

This system has a configuration space of $\gamma \in \Gamma$

$$\dim(\Gamma) = 2N + 2N + N = 5N$$

Vi ønsker å redusere configuration space uten å miste viktig informasjon, og vil derfor representere swarm modelling som en rekke mappings til lavere dimensjoner:

$$f: \Gamma \rightarrow \varphi$$

where $\varphi \in \Phi$ and $\dim(\varphi) \ll \dim(\Gamma)$

Modelling represented as a series of mappings:

$$\begin{array}{ccc} \boldsymbol{\gamma}_t & \xrightarrow{g} & \boldsymbol{\gamma}_{t+1} \\ f \downarrow & & \downarrow f \\ \boldsymbol{\varphi}_t & \xrightarrow{h} & \boldsymbol{\varphi}_{t+1} \end{array}$$

$$\begin{aligned} g: \Gamma &\rightarrow \Gamma \text{ giving } g(\boldsymbol{\gamma}_t) = \boldsymbol{\gamma}_{t+1} \\ h: \Phi &\rightarrow \Phi \text{ giving } h(\boldsymbol{\varphi}_t) = \boldsymbol{\varphi}_{t+1} \end{aligned}$$

Å regne $h(\varphi)$ er betydelig enklere enn å regne $g(\gamma)$, og vi ønsker derfor å bruke transformen f til et lavdimensjonalt domene.

Question 4: When is it appropriate to use rate equations for modelling swarm dynamics?

Å bruke rate equations er passende for å modellere swarm dynamics når vi er interessert i endringsraten til spesifikke verdier, for eksempel endring i density av roboter over tid etc.

Question 5: Could you explain the Langevin equation?

Generelt for bevegelsesmodeller av en agent med drift:

Model of stochastic motion of one agent with drift:

$$\dot{\mathbf{R}}(t) = \mathbf{F}_t + \mathbf{C}$$

where $\mathbf{R} = (r_x, r_y)$ is the position of an agent in 2D space

$\mathbf{F}_t = (X_t, Y_t)$ is the stochastic term of random motion

$\mathbf{C} = (c_x, c_y)$ is the non-stochastic drift term

Et tilfelle av dette er Langevin-likningen:

The Langevin equation:

$$\dot{\mathbf{R}}(t) = \underbrace{\mathbf{A}(\mathbf{R}(t), t)}_{\text{Non-stochastic drift term}} + \underbrace{B(\mathbf{R}(t), t)\mathbf{F}(t)}_{\text{Stochastic random term}}$$

where $\mathbf{R}(t)$ is the position of an agent in 2D space

$\mathbf{F}(t)$ is the stochastic term of random perturbation

\mathbf{A} is the non-stochastic drift term

B is a scalar for the stochastic term

The Langevin equation, typically:

$$\mathbf{A}(\mathbf{R}(t), t) = \nabla P(\mathbf{R}(t), t)$$

where $\nabla P(\mathbf{R}(t), t)$ is a gradient in a potential field P

Langevin-ligningen modellerer hvordan agenten beveger seg som følge av drift og tilfeldig perturbation (støy).

Question 6: Could you explain the Fokker-Planck equation?

Langevin-likningen modellerer agents bevegelse på "mikronivå" / veldig lavnivå. Vi ønsker å se på koblingen mellom dette og bevegelse på makronivå / mer overordnet.

The Fokker-Planck equation:

$$\frac{\partial \rho(\mathbf{r}, t)}{\partial t} = \underbrace{-\nabla(\mathbf{A}(\mathbf{r}, t)\rho(\mathbf{r}, t))}_{\text{Non-stochastic drift term}} + \underbrace{\frac{1}{2}Q\nabla^2(B^2(\mathbf{r}, t)\rho(\mathbf{r}, t))}_{\text{Stochastic diffusion term}}$$

where ρ is the robot density at position \mathbf{r} and time t

Q is a scalar for the stochastic term

∇ is the nabla operator and ∇^2 is the Laplacian

We can derive the fraction of robots at time t over area W by

$$s(t) = \int_{\mathbf{r} \in W} \rho(\mathbf{r}, t)$$

Vi ser at Fokker-Planck modellerer ikke bevegelsen til en enkelt robot i rommet, men heller tettheten av roboter i et punkt, og hvordan den endrer seg over tid.

Hvis vi vil utlede andelen roboter i et område W over tid, kan vi integrere tetthetsfunksjonen til robotene over alle punktene \mathbf{r} i området W .

Theme 10: Swarm robotics 2

Question 1: Could you explain the voter model for decision-making in a swarm system?

Voter model: en robot i har tilgang til alle sine naboers meninger, velger en nabo j tilfeldig, og bytter til j sin mening. Denne modellen er veldig enkel og har høy accuracy, men konvergerer sakte. Hva accuracy betyr er uklart; kan bety at den initielt dominante meningen pleier å ende opp som meningen det konvergeres mot til slutt.

Question 2: Could you explain the majority rule?

Majority rule: roboten ser på alle sine naboers meninger, og teller hvor mange ganger hver mening forekommer. Roboten velger meningen som forekommer oftest. Denne modellen konvergerer raskere enn voter model, men er mindre accurate (kan oftere skje ved tilfeldighet at den konvergerer mot en annen mening enn den predominante).

Question 3: Could you explain how urn models function?

Urn models: går ut på tilfeldig trekking av agenter i svermen (vi ser på dem som baller med farger). Man bryr seg ikke om hvor i rommet de er plassert, så man får helt tilfeldig trekning. Avhengig av hvilken urnemodell bestemmer de trukkede ballene hvordan andelen av farger på ballene endres (enten bytter de trukkede ballene farge, eller så bytter en annen kule i urnen farge).

Ehrenfest urn model [Ehrenfest & Ehrenfest, 1907]:

$$\Delta B(B(t)) = -2 \frac{B}{N} + 1$$

where N is total number of balls

Eigen urn model [Eigen & Winkler, 1993]:

$$\Delta B(B(t)) = \begin{cases} 2 \frac{B}{N} - 1, & \text{for } B \in [1, N-1] \\ 0, & \text{else} \end{cases}$$

Swarm urn model [Hamann, 2013]:

$$\Delta s(s) = 4 \left(P_{FB}(s) - \frac{1}{2} \right) \left(s - \frac{1}{2} \right)$$

Where Ehrenfest $P_{FB}(s) = 0 \Rightarrow s^* = 0.5$

Eigen $P_{FB}(s) = 1 \Rightarrow s^* = 0 \vee 1$

Swarm $P_{FB}(s) = 0.75 \sin \pi s \Rightarrow s^* = 0.23 \vee 0.77$

Ehrenfest gir negativ feedback (konvergerer mot 50/50), Eigen gir positiv feedback (divergerer mot 100/0), Hamann er midt imellom (divergerer mot ca. 75/25).

Question 4: Could you explain the Hegselmann and Krause decision-making model?

Clustering of opinions by having robots move to the centre of gravity of their neighbourhood:

$$x_i = \frac{1}{|G_i|} \sum_{j \in G_i} x_j + \varepsilon_i$$

where $G_i = \{1 \leq j \leq N: \|x_i - x_j\| \leq 1\}$ and ε_i is a noise term

Hegselmann-Krause: Man flytter robotene til "sentrum" av sitt nabolag, slik at nabolagene blir sentrerte clusters, hvor hvert cluster ender opp med en distinkt mening. Med tillagt støy konvergerer systemet til en mening for alle agentene til slutt.

Question 5: We want to find the warmest, lightest or most radioactive spot in a search area. Could you name a few types of models for solving this problem?

Swarm Adaptive Aggregation: lage et kontrollsystem som samler svermen på et spesifikt punkt avhengig av sensorinput, men som også er fleksibelt til dynamiske endringer i miljøet. Dette er målet vi ønsker å oppnå her.

Noen modeller som kan løse problemet: PSO og taxis kan fungere. Ad-hoc random search er mulig. Gradient ascent og evolutionary optimization også mulig. I tillegg har vi BEECLUST.

Question 6: Could you sketch the Beeclust algorithm?

Hvis vi anser BEECLUST fra perspektivet til en agent:

1. Gå rett fremover
2. Er det hindringer eller roboter i nærheten?
 - a. Hvis hindring: snu vekk, gå til steg 1
 - b. Hvis robot: stopp, ta måling med sensor, vent en stund avhengig av målingen (jo bedre måling jo lenger tid), u-vending, gå til steg 1.

Ideen med dette overordnet er at hver gang to roboter møtes, vil de evaluere om de har funnet et bra område. Om de har det, vil de stå der lenger, som øker sannsynligheten for at andre kommer til samme sted og stopper der de også. Over tid er det sannsynlig at det danner seg et cluster i områder med høye sensorverdier.

Question 7: Could you explain the Langevin equation in relation to the Beeclust algorithm?

BEECLUST svarer til å sette drift-termen i Langevin lik 0, så det bare er tilfeldig bevegelse som styrer en agent (den stokastiske termen blir igjen). Vi har også en mikronivå-modell av ventetiden:

Microscopic model: The waiting time

$$w(\mathbf{R}) = \frac{w_{max}P^2(\mathbf{R})}{P^2(\mathbf{R})+c}$$

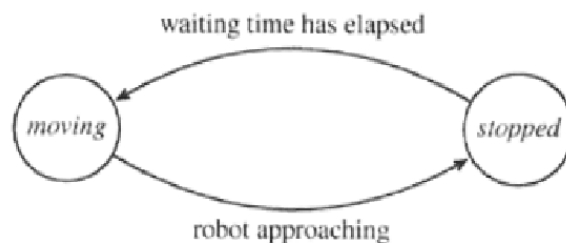
where $P(\mathbf{R})$ is temperature at position \mathbf{R}

w_{max} is maximal waiting time

c is a scaling constant

Man kan også anse den mikroskopiske modellen som en finite state machine:

Microscopic model: Finite state machine



Where

moving is: $\dot{\mathbf{R}}_m(t) = B\mathbf{F}(t)$

stopped is: $\dot{\mathbf{R}}_s(t) = 0$

Slik at når en agent har ventet ferdig, er state = moving og bevegelsen tilfeldig. Når agenten møter en robot, er state = stopped og bevegelsen lik 0.

Question 8: Could you explain the Fokker-Planck equation in relation to the Beeclust algorithm?

Macroscopic model: Fokker-Planck equation (from Langevin)

moving: $\dot{\mathbf{R}}_m(t) = B\mathbf{F}(t) \Rightarrow \frac{\partial \rho_m(\mathbf{r}, t)}{\partial t} = B^2 \nabla^2 \rho_m(\mathbf{r}, t)$

stopped: $\dot{\mathbf{R}}_s(t) = 0 \Rightarrow \frac{\partial \rho_s(\mathbf{r}, t)}{\partial t} = 0$

Where $\rho(\mathbf{r}, t)$ is density of moving or stopped agents

Fokker-Planck svarer også for BEECLUST som i det generelle tilfellet til å modellere tetthet av agenter fremfor bevegelsen til en agent. Det man finner i BEECLUST er at når agenten står stille og venter, så endrer ikke tettheten i punktet seg; den forblir konstant.

Når vi tar hensyn til hvilke agenter som går fra bevegelse til stans, og motsatt, får vi:

Macroscopic model: Fokker-Planck equation (from Langevin)

moving:

$$\frac{\partial \rho_m(\mathbf{r}, t)}{\partial t} = \underbrace{B^2 \nabla^2 \rho_m(\mathbf{r}, t)}_{\text{Diffusion term}} - \underbrace{\rho_m(\mathbf{r}, t) \varphi}_{\text{Flow into stopped}} + \underbrace{\rho_m(\mathbf{r}, t - w(\mathbf{r}))}_{\text{Flow into move}}$$

Diffusion forteller hvordan agentene spres naturlig, mens flow into stopped forteller hvor ofte det stoppes i et område (gitt av en konstant φ), mens flow into move forteller hvordan agenter som har ventet ut sin tid i området begynner å bevege seg igjen. Når agentene står stille har vi i stedet:

Stopped (not necessary to model explicit):

$$\frac{\partial \rho_s(\mathbf{r}, t)}{\partial t} = \underbrace{\rho_m(\mathbf{r}, t) \varphi}_{\text{Flow into stopped}} - \underbrace{\rho_m(\mathbf{r}, t - w(\mathbf{r}))}_{\text{Flow into move}}$$