# Designovervejelser:

**Selektion**

**Roulette metoden**

Hvis alle individer er tildelt et stykke på en roulette og størrelsen på stykket er proportional med individets fitness. Rouletten bliver spinnet n antal gange, hvilket vil vælge forældrene til den næste generation. Under hvert spin bliver individet under roulettens markør valgt til, at være en del af en gruppe af forældre til den næste generation. En kandidat kan godt blive valgt til, at være forældre flere gange, dette er tilladt, da forælderen vælges til næste generation og ikke selve individerne i generationen. Formålet med denne metode er, at få valgt de forældre med den største fitness til næste generation, da de har større sandsynlig for, at skabe individer med større fitness. Problemet med denne metode er dog, at den genetiske algoritme hurtigt vil stå fast i den ene del af fitness rummet. Dette gør det muligt at vælge den samme forælder flere gange og derved skabe en meget ensartet population, hvilket betyder, at bestemte områder vil udforskes af rummet i stedet for at udforske hele rummet.

Rank metoden

Rank metoden har ligheder med roulette metoden, men i stedet for at være proportionel med den absolutte fitness, er den proportionel med den relative fitness. Der er altså ligegyldigt om den fitteste har 10 gange højere fitness en den næste i rangen, eller om den har 0.0001% højere fitness. I begge situationer vil sandsynligheden for den ’fitteste’ være den samme.

Tournament metoden

To tilfældige individer bliver valgt fra populationen. En tilfældig værdi fra 0-1 genereres, der sammenligner den med valgte sandsynligheds værdi. Hvis værdien er mindre eller lig med sandsynlighedsværdien, bliver det individ med højst fitness valgt, ellers bliver individet med den lavere fitness valgt. Sandsynlighedsværdien bliver altid sat højere end 0.5 for at favorisere individet med den højeste fitness.

”Search space” referer til en gruppe af kandidat løsning til et problem, hvor der er en ”distance” i mellem kandidaterne. Tags eksempelvis et vigtigt problem indenfor bioengeering, hvordan et protein designes. Antaget at der skal søges efter et protein, som er en sekvens af aminosyre, der kan blive bruges til at bekæmpe en virus. ”Search space” vil være en kollektion af alle mulige proteiner. Dette vil give uendelige mange muligheder, derfor begrænses længden af proteinet til længden 50 som stadig vil være i et stort ”search space” siden der er 20 mulige aminosyre i hver position i proteinet. Hvis aminosyrerne repræsenteres i form af alfabetet vil et muligt protein se således ud.

ASDKEGHB…. Distancen defineres mellem proteinerne som forskellen i alfabetet på den tilsvarende position i et andet protein fx ASDKEGHB og BSDKEGHB er distance 1 og distance mellem ASDKEGHB og GCCHAKAA er 8.