# Téma Beszámoló

# Útvonal és erőforrás optimalizáló rendszer készítése

### Tagok:

### Holló-Szabó Ákos

### Koppány Bence

### Manninger Miklós

### Réti Marcell

## A Téma leírása:

Célunk egy térképen, vagy ehhez megfelelő gráf struktúrában (ahol a csúcsok a városok, és az élek a városok között futó utak) való összes város, csúcs meglátogatása minél optimálisabb idő alatt. A matematikában, ezt utazó ügynök problémának nevezik.

Az utazó ügynök problémában a bemenetünk egy teljes gráf és a várt eredmény pedig egy lehető legminimálisabb összsúlyú Hamilton körút, azt szimulálva, hogy az ügynökünk minden csúcsot végigjárt egyszer, és ezt megpróbálta a leggyorsabban megtenni. A probléma NP-teljes számítási nehézségű, ami annyit tesz, hogy egyelőre nem találtak rá polinom időben lefutó algoritmust, nem determinisztikusan polinom időben megoldható. Ha egy NP – teljes problémára, (amely minden NP-beli problémánál nehezebb) egy polinom idejű optimális algoritmust találna valaki, az megoltaná a P=NP? híres matematikai kérdést és teljesen megváltoztatná a matematikai hozzáállást jó néhány témakörből.

A mi feladatunk felkutatni és leimplementálni a legjobb approximációs módszereket, amik természetesen nem az optimális megoldást adják, csak egyre jobb lefutási időt vagy egyre jobb becsléseket, közelítést adnak.

Készítenünk kellett egy vizualizációs keretrendszert, ahol a leimplementált algoritmusokat vizsgálhatjuk meg, akár futás közben, illetve egy hozzá tartozó teszt keretrendszert, ahol a futási eredményeket tudjuk kiértékelni.

A probléma bonyolultságát redukálandó, teljes gráfokat használtunk, ahol minden csúcs mindegyik másikkal egyszeresen össze van kötve, illetve betartottuk, hogy a gráf bármely három pontjára igaz a háromszög-egyenlőtlenség tétele. Ezzel a kikötéssel redukáltuk a problémát (ezt euklideszi utazó ügynök problémának nevezik), bár ezzel a bonyolultsága nem csökkent, hiszen már a Hamilton kör keresése is NP – teljes probléma. Az euklideszi tér jellemzőit viszont kihasználhatjuk olyan módon, hogy a gráftérbeli feladat bármikor átültethető egy térképen értelmezett valós problémába.

A feladat komplexitását növeljük azzal, hogy nem egy ügynököt, hanem tetszőleges számú ügynököt indítunk a gráf csúcsainak legkevesebb idő alatti bejárása érdekében, ezzel még jobban megközelíthetünk egy való életbeli problémát. Természetesen az ügynökök és a gráf paramétereit mi generáljuk a rendszerbe.

## A vizualizációs keretrendszer:

A rendszer alapvető elgondolása az volt, hogy átfogó keretet adjon az implementálandó algoritmusoknak. Rendelkezzen egy felülettel, ahol a gráf és ágensek információit lehet bevinni a rendszerbe, illetve ezeket el is lehessen menteni. Továbbá szükséges volt egy grafikus rendszer kialakítása, amely az adott algoritmus szerint jeleníti meg a folyamat lépéseit.

### Felhasználás:

Az adatok beolvasása szöveges állományból történik. A gráf koordinátáit adhatjuk meg az alkalmazás megjelenítő egységének két dimenziós koordinátarendszerében megjelenítve, illetve az ágensek kezdőpozícióját, hogy melyik ügynök mely indexű csúcsból indul. A legtöbb algoritmus egyelőre úgy lett megírva, hogy az első ágens kezdőpozíciójától indul a bejárás minden ágensnek, így a többi kezdő index elhanyagolható.

Az adatok beolvasása után lehetőségünk van elmenteni azokat egy sorosítható összefoglaló konfigurációba, ami tartalmazza a gráf és ágensinformációkat egyaránt. Ez az objektum később vissza is kérhető, újra betölthető.

A konfiguráció kiválasztása után ki kell választani a futtatandó algoritmust, majd elindítani azt. Az algoritmusok lefutási, számítási ideje nagyban függ a gráf pontok és ügynökök számától, így elképzelhető, hogy még a kezdeti inicializációs lépés feldolgozása is hosszadalmas lehet.

Ha az algoritmus elindult, akkor lehetőségünk van lépésenként futtatni és kiértékelni a megoldást, vagy végigfuttatni az algoritmust az algoritmus kilépési feltételéig. A végigfuttatjuk a programot az algoritmuson, akkor lehetséges, hogy az egyes lépések vizualizációját nem látjuk rendesen, ha gyorsan vált a rendszer a lépések között. Az algoritmus tetszés szerint újraindítható, illetve bármikor válthatunk az algoritmusok között.

Az 1. ábrán látható a felhasználói felület a fentebb felsorolt funkcionalitásoknak megfelelően. A bal felső csoport felelős a gráf információk és ágensinformációk fájlból való betöltéséért és konfigurációként való elmentéséért. Az alatta lévő csoport keretein belül választhatjuk ki a kívánt konfigurációt és a rajta futtatni kívánt algoritmust. A *Run Algorithm* gomb segítségével inicializálhatjuk és elindíthatjuk a folyamatot. A felhasználói felületen ekkor megjelenik a gráf és az inicializációs lépés eredménye. Ezek után a *Next Move* gomb segítségével léptethetjük az algoritmust, vagy a *Run through* gombbal végigfuttathatjuk azt. Az aktuális legjobb eredményt az *Actual result* mező értékeként láthatjuk.



1. Ábra: Grafikus keretrendszer felhasználói felülete

### Implementáció:

Tervezési szempontból az alkalmazás elkülöníthető rétegeket valósít meg, amelyek fejlesztését külön is lehet végezni. Az implementáció Visual Stúdióban készült a C# nyelv és a WinForms keretrendszer segítségével.

A program tartalmaz, olyan alap osztályokat melyekhez mindegyik réteg hozzáférhet. A *Vertex* és *Edge* osztályok egy gráf csúcsait és éleit reprezentálják, a *Coordinate* osztály a való életbeli koordinátázáshoz szükséges. Ezen alap osztályok felhasználásával készült az *AbstractGraph* és ebből leszármazó egyszerű gráf, *SimpleGraph*, és egyszerű teljes gráf, *CompleteGraph*. A gráf osztályok implementálásakor a gráffal végzett műveletek során (például csúcs hozzáadás, él elvétel stb.), a gráf nem lép ki a saját típusából (például a teljes gráf egy csúcs hozzáadása után is teljes gráf marad). Az *Agent* és *AgentManager* osztály tárolja az ágensinformációkat. A konfigurációt, ami a gráf és ágens információkat tartalmazza, a szerializálható *Configuration* osztály tartalmazza.

A felhasználói felület, az aktuális algoritmus és gráf állapot kirajzolásáért a fő ablak felel. Az ablak meghívható függvényeiben megjelennek az előzőleg említett alap osztályok.

Az Algoritmusok ősosztálya az *Algorithm*, amely egy egységes interfészt biztosít az összes leszármaztatott algoritmus, és az őket meghívó folyamatok számára. Minden algoritmus külön osztályba lett kiszervezve, így, ha új algoritmust szeretnénk felvenni, azt leszármazott osztály szinten kell megkódolni. Az implementált algoritmusok osztályai rendre: *BruteForce*, *Christofides*, *GenetcAlgorithm*, *GreedySearch*.

A felhasználó felület és az algoritmusok közötti kapcsolatot a *Coordinator* osztály teremti meg, amely referenciát tárol a megnyitott ablak osztályáról és az aktuálisan futtatni kívánt algoritmusról. A felhasználó általi hívások az ablak interakciójából a *Coordinator* osztályba futnak be, amely meghívja az adott algoritmus megfelelő függvényeit. Miután az algoritmus szakasz lefutott, a Coordinator kinyeri a változásokat és eredményeket, majd meghívja az ablakot frissítő metódusokat.

A fájl kezeléssel kapcsolatos műveleteket a *FileManager* osztály végzi. Feladata a gráf és ágens adatok kinyerése szöveges erőforrásból, illetve a konfigurációk szerializálása és deszerializálása.

## Algoritmusok:

### Brute Force

A Brute Force (nyers erő) algoritmus a legegyszerűbb algoritmusok közé tartozik, hiszen a legtriviálisabban oldja meg az ügynök problémát, amit azt jelenti, hogy az összes lehetséges útvonalat kiszámítja és közülük a legrövidebbet visszaadja egy adott gráfban, azaz számunkra az optimális útvonalat. Esetünkben teljes gráfokat vizsgálunk, ami azt biztosítja, hogy bármely két tetszőleges csúcs között biztosan fut egy él. Ezt kihasználva, ha a gráf csúcsait egymás után rakjuk valamilyen sorrendben, akkor azon végig menve egy Hamilton utat/kőrt kapunk, attól függőben, hogy vissza szeretnénk-e térni az indulási pozícióba. Ezt a gondolatmenetet folytatva, ha egy gráf csúcsait permutáljuk, akkor az összes lehetséges bejárási sorrendet megkapjuk, amelyekre igaz, hogy minden csúcsot pontosan egyszer érintettünk. Egy n csúcsú gráf esetén az n! darab permutációt jelent. A brute force algoritmus egyik nagy problémája, hogy mivel az összes lehetséges esetet megvizsgálja, így nagyobb gráfok esetén ez a megoldás szinte kivitelezhetetlen a hosszú futási idő miatt. Azt is tudjuk, hogy egy teljes gráfban (n-1)!/2 darab különböző Hamilton-kőr van. Ebből jól látszik, hogy az algoritmus ugyan azt a kőrt többször is megtalálja és kiszámolja rá az út értékét, tehát nem csak lassú, de rengeteg ismétlést végez feleslegesen. Ezek után jogosan merül fel a kérdés: Miért implementáltuk a brute fore algoritmust?  
Annak érdekében, hogy a többi implementált algoritmust (lásd lentebb) által adott eredmények jóságát tudjuk mihez mérni, szükségünk van az optimális megoldásra is. Ez a későbbiekben jó viszonyítási alapot, támpontot ad ahhoz, hogy össze tudjuk hasonlítani a többi algoritmus által adott eredményekkel (pl futási idő, megoldás pontossága).

Az eddigi leírt megoldás egyelőre csak egy ágens létével foglalkozik. Annak érdekében, hogy egyszerre több ágenst is be tudjunk vetni, a fenti megvalósítás némi módosítást igényel. Ebben az esetben meg kell határozni, hogy melyik ágens melyik csúcsokat fogja bejárni és milyen sorrendben. Ennek szemléltetését az alábbi példa mutatja.

Csúcsok száma: 8.

Ágensek száma: 2.

A csúcsokat számokkal különböztetjük meg: 1 2 3 4 5 6 7 8

Az ágenseket az egyszerűség kedvévért betűvel különböztetjük meg: a és b.

Egy lehetséges csúcs sorrend bejárás: 1 3 4 2 6 5 8 7.

Az ágensek egy lehetséges hozzárendelése: a a b b a a b a.

Összegezve:

a ágens által bejárt csúcsok: 1 3 6 5 7.

b ágens által bejárt csúcsok: 4 2 8.

Az összes ágens-csúcs kisosztást úgy kapjuk meg, ha vesszük az ágensek ismétléses permutációját a gráf csúcsszámának függvényében. Ez 2^n darab megoldás, ahol n a csúcsok számát jelöli. A brute force algoritmus több ágens esetén tehát egy permutációból és egy ismétléses permutáció egymásba ágyazásából áll. Fontos még megjegyezni, hogy több ágens esetén Hamilton utakat keresünk, vagyis nem kötelező a kezdőpontba való visszatérés, továbbá az ágensek kezdőpozíciója megegyezik, így más-más pontokból indítva őket más-más eredményt fogunk kapni.

### Christofides

A Cristofides algoritmus volt az első algoritmus, amellyel foglalkoztunk a projekt folyamán. Ugyan még csak az egy ügynök problémára jelentett megoldást, de mivel az egy ügynök problémának a több ügynök probléma speciális esete, jelentősen hozzá járult a probléma feltérképezésében, megértésében, és a csapat összehangolódásában. Fontos szerepet játszott abban is, hogy elkészüljön a keresztrendszer, amiben a további algoritmusokat futtattuk, és teszteltük.

A Christofides ugyan csak egy közelítő módszer, de jól megírva rendkívül gyorsan ad páratlanul jó közelítéseket az egy ügynök problémára. Az alap ötlete az, hogy ha veszünk egy minimális súlyú feszítő fát, és azt a lehető legkisebb súlyú élekkel eulerkörré alakítjuk, akkor egy olyan élhalmazt kapunk, ami a legnagyobb éleket nem tartalmazza, és amelyet könnyű olyan hemiltonúttá alakítani, melynek élei a legrövidebbek közül valóak. Ez a hemiltonút jelentette a teljes gráfban az egy ügynök probléma megoldását.

Az algoritmus által specifikált lépések összefoglaló jelleggel:

1. Először keressünk egy minimális összsúlyú feszítő fát a térképet reprezentáló teljes gráfban
2. Keressük ki a fa páratlan fokszámú éleit, és készítsünk egy teljes részgráfot belőlük és a köztük futó élekből.
3. A részteljes gráfban keressünk minden csúcsot lefedő, minimális összsúlyú független élhalmazt.
4. A független élhalmaz csúcsait a fában a megfelelő csúcsokkal megfeleltetve fésüljük össze a két gráfot. (A mindkettőben szereplő éleket itt duplikálni kell)
5. Ekkor egy olyan gráfot kaptunk, aminek van Euler köre, mivel minden csúcsának páros a fokszáma. Ennek az az oka, hogy a független élhalmaz élei a fa páratlan fokszámú csúcsainak fokszámát eggyel növelték, a párosoknak pedig egyetlen nem üres részhalmazát se fedik. Keressük meg ezt az Euler kört!
6. Az Euler körből hagyjuk el az ismétlődő csúcsokat úgy hogy minden csúcs pontosan egyszer szerepeljen végül. Ekkor egy hemilton kört kapunk, amiből ha elhagyjuk az egyik az ügynök központra illeszkedő élet, akkor meg is kapjuk a keresett hemilton utat.

A feladat megoldásának pontossága leginkább ebben az utolső lépésben dől el, mivel a többi lépésben többnyire jól ismert algoritmusokat kellett alkalmazni, melyeknek közel egyértelmű az eredménye. Ennek a lépésnek viszont számos megoldása van. Érdekesség képp azt is megemlítem, hogy bizonyítható, hogy a helyes megoldás is kihozható még ekkor; a probléma csak az, hogy exponenciális futási idővel. Mi ezt a lépést nem optimalizáltuk le teljesítményre, hogy időt nyerjünk a valódi feladatunk megvalósításához.

A lépések és az azokra alkalmazott algoritmusok részletes kifejtése:

### Mohó algoritmus

A programozásban megismert mohó algoritmusok ismertetőjele, hogy tulajdonképp gondolkozás nélkül, Trial & Error módon, véletlenszerűen próbálkozva próbálják elérni az optimális, vagy az optimálishoz minél közelebbi megoldást. Ez jelen esetben sincsen másképp.

Az algoritmus bemenetei a következőek:

* A gráf: ezen keressük a lehető legjobb megoldást. Jelen esetben úgy vettük, hogy a megoldás jóságát mérő szám, az az, hogy mely ágens járta be a leghosszabb utat. Mivel egy egység megtétele egyenlő a rá fordított idővel (egy egység megtételéhez szükséges idő konstans egy egység), így az adott megoldás megfogalmazása a következőképpen is történhetne: a megoldás jóságát az összes csúcs bejárási ideje adja, ezt szeretnénk minimalizálni. Mivel az ágensek parallel futnak, így az adott megoldás értéke a leghosszabb ideig futó ágens futási időtartama.
* Az ágensek száma: hány ágens fogja megpróbálni egyidejűleg bejárni a gráfot
* Az ágensek kezdőpontja: Minden ágens egy kezdőpontból indul, és az algoritmus nem várja el a kezdőpontba való visszatérést, tehát diszjunkt Hamilton utakról beszélhetünk, minden ügynök esetében.
* PATIENCE\_PARAMETER: Ez a bemeneti paraméter mondja meg, hogy egy adott lokális minimumba való ragadás során hányszor próbálkozzon a rendszer az onnan való kilépésből.
* NUMBER\_OF\_RUNS: Egyszerű generáció szám, hányszor próbáljon az algoritmus új legjobb megoldást keresni az előző legjobb alapján.
* MAX\_ROUTE\_LENGTH\_PER\_AGENT: Egy adott ágens maximálisan befutható csúcsait korlátozza. Ezen paraméter megválasztásakor ügyelni kell, hogy a feladat megoldható maradjon!

Az algoritmus kezdeti lépése, hogy létrehoz, egy teljesen randomizált megoldást. A megoldás struktúrája az ügynökök szerint van felbontva. A megoldás minden ügynökre tárol egy tömböt, amiben a gráf csúcsainak indexei vannak. Az ágens tömbjében lévő csúcsindexek azt jelképezik, hogy az adott ügynök abban a sorrendben bejárja az adott csúcsokat. Így elképzelhető az is, hogy egy ügynök el sem indul.

Az algoritmus mindig számon tartja a globális legjobb megoldást és az adott generáció legjobb lokális megoldását. Miután az inicializálás megtörtént megkezdődik egy újabb generáció legyártása, ahol egy az eddigi legjobbnál jobb eredményt szeretnénk legenerálni.

A lokális legjobb megoldást véletlenszerűen kell legenerálni, majd ennek egy úgynevezett szomszédját kell létrehozni. A szomszéd létrehozásának öt metodikája van, amiből véletlenszerűen kell egyet kiválasztani.

Szomszéd generálási lehetőségek:

* Úton belüli inverzió: Egy véletlenszerű úton egy rész utat invertálunk.   
  2 3 4 5 6 -> 2 6 5 4 3
* Úton belüli csere: Egy véletlenszerű úton két rész utat felcserélünk egymással.  
  2 3 4 5 6 -> 2 4 5 2 3
* Úton belüli beszúrás: Egy véletlenszerű úton egy rész utat máshova szúrunk be az úton belül.  
  2 3 4 5 6 -> 2 4 5 3 6
* Utak közötti csere: Két véletlenszerű út egy-egy szakaszát átcseréljük.  
  2 3 4 5 6 -> 2 7 6  
  7 -> 3 4 5
* Utak közötti transzfer: Két véletlenszerű utat választunk. Az elsőből egy véletlenszerű szakaszt kivágunk és átmásoljuk a másikba, arra a helyre ahonnan az eredetiből indult.  
  2 3 4 5 6 -> 2 6  
  7 -> 7 3 4 5

Az adott szomszéd generálása után megnézzük, hogy a kijött megoldás jobb-e mint a lokális legjobb, ha igen akkor a lokális megoldást felül írjuk vele. Ha az eddigi legjobb globális megoldásnál is jobb azzal is ugyanezt tesszük. Ezek után a szomszédgenerálás újra indul.

Ha a szomszédok egy idő után nem generálnak jobb eredményt, mint a lokális legjobb akkor egy lokális minimumba érkeztünk. A PATIENCE\_PARAMETER szabja meg, hogy hányszor próbálkozzon kilépni belőle. Ha nem sikerül a generáció legyártása a végéhez ért.

A generációk számát a fent említett NUMBER\_OF\_RUNS szabályozza.

Az algoritmus tehát a leírtak alapján láthatóan nem gondolkozik, csak véletlenszerű cserékkel, inverziókkal, beszúrásokkal próbál utat módosítani és ezzel jobb megoldásokat találni. Nagy előnye, hogy elég gyorsan lefut. Az algoritmus hatékonysága növelhető lehetne szimulált lehűtés bevezetésével.

### Genetikus algoritmus

A genetikus algoritmus nem egy konkrétan egy feladatra alkotott algoritmus mint a Christofides, hanem egy általános módszertan problémák közelítéses megoldására. Előszeretettel használják np teljes problémák esetén, mivel rendkívül kedvező futási idővel ad nagyon jó közelítést, és nem is tartozik a legbonyolultabbak közé(, még ha jobb teljesítményt hozó variánsairól és más algoritmusokkal alkotott hibridjeiről ez nem is mondható el feltétlen). Nevét azért kapta, mert az evolúció modeljének mintájára alkották meg, felfogható egy fajta nemesítés ként is. Itt a fajunk nem más mint a konkrét feladat kellő mértékű általánosításának(enyhítésének) megoldáhalmazának részhalmaza. A cél pedig hogy a megoldáshoz legközelebb állók életben maradjanak, és nagyobb valószínüséggel szaporodjanak, mint rosszabbul teljesítő táraik.

A genetikus algoritmusok során be kell vezetni néhány alapfogalmat, amit a további leírásban és az algoritmusban is használni fogunk.

* Kromoszóma: A feladat enyhítésének egy megoldása. Jelen esetünkben ez az enyhített feladat a következő:
  + az ügynökök mindegyike járjon be egy csúcsot legalább
  + minden csúcsot csak egy ügynök érintsen
  + minden csúcsot érintsenek az ügynökök együtt véve.
* Allél: A kromoszomáinkat megkülönböztető elemi attribútumok. Jelen esetben ez az hogy melyik ügynökök, melyik csúcsokat, milyen sorrendben járják be.
* Populáció: Kromoszómáinknak halmaza. Ők azok, akik a nemesítésben részt vesznek. Létszámukat érdemes konstans értéken tárolni a konstans memóriaigény biztosításának érdekében
* Fitness: Annak mértéke, hogy egy adott kromoszóma milyen közel áll az eredeti probléma megoldásához. Ez alapján válogatjuk őket. Jelen esetben ez a leghosszabb út, amit egy ügynök megtesz, hszen az MTSP-ben ennek kell a lehető legrövidebbnek lennie, mást nem is általánosítottunk.
* Generáció: A nemesítés több iteráción keresztül tart. Az egy iteráció végére megmaradt populációt szokták az iteráció sorszámával is jellemezni, és rajtuk keresztül a genetikus algoritmus hatékonyságát.

Egy általános genetikus algoritmus három főbb lépésből áll, melyek során generációról generációra lépve újabb populációt állít össze. Az Inicializáció csak az algoritmus előtt fut le. Azt követően a másik három lépést írtam le abban a sorrendben, ahogy az iterációban is követik egymást:

* Inicializáció: Létre kell hoznunk egy kezdeti populációt, ami a nemesítés kezdeti alanyául szolgál. Mivel nem ismerjük a megoldást, ekkor még nem tudjuk milyen allélekből áll, de ahhoz hogy az algoritmus jól működjön, elengedhetetlen, hogy legtöbb allélje kellő mennyiségben jelen legyen a kezdetleges populációban. Ezt csak valószínűségi alapon tudjuk biztosítani azzal hogy nagyra vesszük a populáció méretét, és biztosítjuk sokszínűségét.
* Szelekció: Minden populációbeli egyednek megvizsgáljuk a fitness értékét és kiválasztjuk a legjobb, legéletrevalóbb kromoszómákat. A kiválasztottakból származtatjuk alapjáraton az új generáció maradékát, de ekkor még fent áll annak az esélye, hogy benne ragadunk egy helyi minimumban. Ez alatt azt kell érteni, hogy ha egyik körben az egyik szükséges, de nem jelentőségteljes allél nem jelenik meg a legjobbaknál, akkor elveszhet, és ugyan nagyon közel kerülünk a célunkhoz, de nem tudjuk elérni. Ennek orvoslására a legjobbak közé még beválogatunk pár rosszabbat is véletlen szerűen, így adva több esélyt a szükséges allélok fent maradására.
* Keresztezés: A kiválasztott elemeket megtartva, és a többit elvetve megkaptuk azt a halmazt, amiből az új kromoszómákat származtatni szeretnénk. Ebben a szakaszban történik meg maga a származtatás. Valamilyen elv szerint összepárosítjuk a kiválasztott kromoszómákat, és páronként egy vagy több új kromoszómát származtatunk belőlük. Az újonnan keletkezett, és a megtartott kromoszómák fogják alkotni az új populációt.
* Mutáció: Hogy növeljük az kromoszómák sokszínűségét, a keresztezéssel keletkezett kromoszómák egy részén még variálunk is. Például egy allélt átírunk, vagy két azonos típusú allélt felcserélünk úgy, hogy még mindig megoldása maradjon az általánosított feladatnak. Ha egy a megoldásban szereplő allél elveszik a szelekció miatt, vagy sose létezett, akkor ez az egyetlen esélyünk arra, hogy újra bekerüljön a populáció bármelyik kromoszómájába.

A konkrét megvalósítás:

Az első, amire ki kell térnünk az a kromoszóma adatszerkezete, mivel ezt is sok féle képpen meg lehet valósítani. Jelen esetben a teljes gráf csúcsait számokkal azonosítottuk, és az alléleket két tömbben tároltuk. Az első tömb a csúcsok egy permutációját tárolta, amiben nem szerepelt a központ, a második tömb pedig azt hogy melyik ügynök mennyit jár be belőlök. Például ha a hat csúcsról van szó és a permutációnk a {1,4,2,0,3,5}, és két ügynökünk van, melyekre a tömb az {2,4}, akkor az első ügynök az 1,4 csúcsokat járja be ebben a sorrendben, a másik pedig a 2,0,3,5 csúcsokat. Az allél fittségét is maga a kromoszóma tárolta

A populációt mi random generáltuk, mivel a szükséges méretű populációval egy közeli valószínűsége volt annak, hogy minden szükséges allél elégszer szerepel, és kellően gyorsan lefuttatható volt. Az allélek fittségét minden kromoszómára a legenerálás legvégén számítottuk ki. Ezt követően kiszelektáltuk az elemek felét. Ennek gyorsítása érdekében először fittség szerinti növekvő sorrendbe rendeztük a kromoszómákat. A kiválasztottakat teljesen véletlenszerűen szerveztük párba a keresztezéshez.(Itt látszik, hogy mivel a populáció felének elemei párbaállíthatóak, a populáció méretének oszthatónak kell lennie néggyel)

A keresztezés talán az egész algoritmus legbonyolultabb és legfontosabb része. Ugyanis úgy kell új elemet generálnunk kettő másikból, hogy az minél nagyobb valószínüséggel járhasson javítással, és ne veszítsék el az elemek pozitív tulajdonságaikat. A két keresztezendő elemet nevezzük apának és anyának. A módszerünk az, hogy megtartjuk az anya csócskiosztását az ügynökökre, és bizonyos elemeket rögzítünk a permutációjában, majd az instabil elemeket olyan sorrendbe rendezzük, ahogy az apában is vannak. Ekkor ha az anya elég jó megoldás, akkor kedvező megtartani az ügynökök csúcs kiosztását. Ahhoz hogy a permutáció pozitív tulajdonságaiból is megtartsunk érdemes biztosítani hogy szomszédos csúcsokat is rögzítsünk így esélyt adva olyan élek megtartására, ami a cél megoldásban is szerepel, vagy köze van hozzá. Ez után származtassunk a párosított két elemből egy kromoszómát fordított szerepkiosztással is.

Jól látható ebből, hogy a keresztezés nem variálja az ügynökök csúcs kiosztását. Nagyon fontos hogy nem is lenne ez célszerű. Ugyanis keresztezésnél egy olyan kromoszómát akarunk kapni, ami a szüleitől nem esik túl messze, hogy egy hangolás lehessen rájuk nézve. Azért jó ez nekünk, mert mivel a szülők közelebb állnak a megoldáshoz mint az előző generáció átlaga, nagyobb valószínüséggel találunk javításra a közelükben A csúcs kiosztást viszont, ha variáljuk, akkor a szülőktől nagyon messze eső, tőlük teljesen idegen megoldást kapunk, ami nagyon kis valószínüséggel lesz náluk jobb.

A mutáció során két csúcsot cserélünk fel a permutációban, de itt sem célszerű a csúcsok kiosztásához nyúlni a feljebb említett indokkal.

Genetikus algoritmus paraméterei, optimalizációja:

A genetikus algoritmus hatékonysága, avagy sebessége, memóriaigénye, és a megoldás pontossága több paramétertől függ, és ezeket nagyon jól kell hangolni ahhoz, hogy a kívánt működést érjük el. A paraméterek az alábbiak:

* A csúcsok száma: Magától értetődően minél több csúcsunk van, annál több helyet foglal egy kromoszóma permutációja, és persze annál több permutáció lehetséges, így annál nagyobb populáció kell ahhoz, hogy a keresett megoldás allélei kellő valószínűséggel megjelenjenek és túléljenek.
* Az ügynökök száma: Ennek is egyértelmű, hogy miért jár több memória igénnyel, és számítással, de talán még fontosabb, hogy a csúcsokat akkor lehet a legtöbb féle képpen kiosztani az ügynökök között, amikor azok száma legközelebb ál a központon kívüli csúcsok számának gyökéhez. Ennek egy elég hosszas számolás a levezetés, de a lényeg az, hogy ismétléses kombináció darab kiosztás létezik, de két megoldás ekvivalens, ha csak a kiosztás sorrendjében különbözik, és a kiosztottaknak megfelelően keverednek a csúcsok is a csúcs permutációban, avagy az ügynökök által bejárt utak páronként megfeleltethetőek. A sima kombináció akkor maximális, ha k=n/2, de ez már az ismétléseshez nem mondható el. Így hát akkor jártunk jobban futási időre és pontosságra, ha nagyon sok vagy nagyon kevés ügynökünk van.
* A populáció mérete: A populáció mérete egy nagyon két oldalú dolog, hiszen minél nagyobb a populáció, annál kevesebb generáció alatt, és annál pontosabb eredményt kapunk, de annál több helyet foglal, és jelentősen több időt emészt fel. Elvégre az elemeket mindig először rendezni kell fittség szerint, ami O(n\*log(n)) művelet, és utána a kiszelektálás O(n),a keresztezés is O(n) az ügynökök számára, a mutáció is várhatóan O(n). Így hát végső soron n\*log(n) arányos vele a számítási idő egy iterációra.
* A mutáció valószínűsége: Minél nagyobb, annál többször kell elvégezni a mutációt, de végső soron elveszik konstans szorzó ként. A memória igénye sem túl kiemelkedő, mivel az általa módosított kromoszómát írja csak át. Sokkal nagyob hatással van viszont a pontosságra, mivel ha túl magas, akkor túl nagyra növeli a szülők és a gyerekek közötti különbséget, ami már mint korábban is említettem, nem túl jó dolog. Ha viszont nagyon alacsony, akkor nagyon magasra ugrik a fonto allélek kihalási esélye.
* Rosszabbul teljesítő kiválasztottak aránya: Ez a paraméter se a memória igényt, se az egy iteráció számítási idejét nem növeli, de ha túl magas, akkor a pontosság nagyon leromlik. Általában nagyon alacsony kell legyen, de elengedhetetlen paraméter.
* Megengedett generációk száma: Ezzel egyenesen arányosan nő a számítási igény, de e nélkül lehetet

A genetikus algoritmusnál először is nagyon fontos megjegyezni, hogy a populáció mérete, a csúcsok és az ügynökök száma nem változik, így a legstatikusabb tárolók is tökéletesen megfelelnek. A nem kiválasztott elemeket pedig nem szabad törölni, mivel ugyan annyi új elem fog keletkezni. Jobban járunk, ha beléjük mentjük el sorra az új elemeket. Az efféle dolgokkal ugyan egy konstans szorzó erejéig, de rengeteg számítási időt lehet spórolni.

Mik a bevett paraméterezések?

Először is a genetikus algroitmusnak ahhoz, hogy egyáltalán működjön, már szükséges a jó beállítás. Éppen ezért vannak nagyon standard beállítások, amikre mindenképp működni fog az algoritmus:

### Dejong Settings (From [DeJong and Spears, 1990]):

Dejong's beállításai a fő standard, ha GA-ról van szó. Dejong bebizonyította, hogy ritk a jól működnek akármely problémára GA kompatibilis problémára

* + Populáció mérete: 50
  + Generációk száma: 1,000
  + Mutáció típusa : bit negálás
  + Mutáció esélye: 0.1% minden kromoszóma minden bitjére a kromoszómát leíró adatnak
  + A rosszabbul teljesítők arányára nem tér ki

### Grefenstette Settings (From [Grefenstette, 1986]):

Tipikusan akkor használják, amikor a számítógép nagyobb populációja nem enged meg túl nagy populációt.

* + Populáció mérete 30
  + Generációk számára nem tér ki
  + táció típusa : bit negálás
  + Mutáció esélye: 1% minden kromoszóma minden bitjére a kromoszómát leíró adatnak
  + A rosszabbul teljesítők arányára nem tér ki

### MicroGA Settings From [David L. Carroll](http://uxh.cso.uiuc.edu/~carroll/)

Nem széles körben elterjedt, de használói négyszer kevesebb evolúció igényről számolnak be.

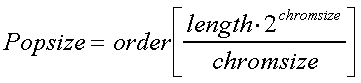
* + Populáció mérete: 5
  + Generációk száma: 100
  + Ceresztezés típusa „uniform”
  + Mutáció típusa: „ugrás és csúszás”
  + Mutáció valószínűsége= 2% and 4%

Notes:

De David teljesen más féle algoritmust használt mint mi. Ő a mutációt mindig újrakezdte random kromoszómákkal, amikor konvergenciát érzékelt, és elmentette a legjobb verziót.

Mik legyenek a mi paramétereink?:

Először is már a feljebb említett algoritmusoknál is megjelent az hogy a kromoszómák által tárolt adat bit értéke szerepet játszik a paraméterek kiszámításában. Hány bitben kifejezhető egy kromoszómánk információ tartalma? Először is a csúcs! permutáció lehetséges és (csúcs szám+ügynök szám-1)alatt az (ügynük szám) nál kevesebb súly kiosztás. Így elég jó felülbecslés az l=felfelekerekít(log(n!))+felfelekerekít((n-m+1)alatt az(m)), ahol n a csúcsok száma és m az ügynökök száma. Mivel a mi állapotterünk túl nagy, nem lesz elég a standard populáció mérete, így használjuk a következő közelítést:



Ahol a Popsize a populáció mérete, a length az egy kromoszóma mérete bitekben, és a chromosomsize pedig félrevezetően egy tulajdonság átlagos mérete avagy (felfelekerekít(log(n!))+felfelekerekít((n-m+1)alatt az(m)))/2.

A mutációról nagyon kevés irodalmat találtunk, de itt is megfogalmazódott egy logika, amit többen is kifejtettek. Nagyobb mutációs esélyre van szükség olyan példa esetén, ahol a fontos allélok könnyen elveszhetnek, és kisebbre ott ahol nagy valószínűséggel azért túl élnek. Mi esetünkben erre egy jó példa a csúcsok kiosztása az ügynökök között. Mi úgy generáltuk az első populációnál, hogy az első ügynöknek sorsoltunk valamennyit úgy hogy a többinek is juthasson legalább egy csúcshoz, majd a másodiknak is, és az az utáninak is. Ennek az a következménye hogy mivel nem biztosítottunk azonos valószínűséget az egyes megoldás osztályoknak, lesznek köztük valószínűbbek és kevésbébb valószínűek. Például hat csúcs és 3 ügynök esetén az {1,1,4}={1,4,1}={4,1,1} valószínűsége 1/4\*1/4+1/4\*1/4+1/4=0.375.., míg a {2,2,2}-nek 1/4\*1/3=0,08333... és az {1,2,3}={1,3,2}={2,1,3}={2,3,1}={3,1,2}={3,2,1}-nek 1/4\*1/4+1/4\*1/4+1/4\*1/3+1/4\*1/3+1/4\*1/2+1/4\*1/2=0.541666 Ez azt jelenti hogy azok a kiosztások például, ahol az összes ügynökre ugyanannyi csúcs jut, nagyon ritkák lesznek az első populációban. Másik példa képp vegyünk 10 csúcsot négy ügynökre it az {1,1,8}={1,8,1}={8,1,1} valószínűsége 1/8\*1/8+1/8\*1/8+1/8=0.15625, míg a {3,3,4}={3,4,3}={4,3,3}-nek 1/8\*1/6+1/8\*1/6+1/8=0.166.... Ebből is látható hogy minél nagyobb a kiosztás értékeinek a szórása, annál kisebb lesz a valószínűsége. Ahhoz viszont hogy tudjuk a pontos valószínűséget már ismernünk kéne a megoldást. Közelíteni ugyan még tudnánk, például ha a városok nagy része közel van a központtól, és alig pár darab van távolabb, de azok is más-más irányban, akkor várhatóan nagy lesz a végső megoldásban a csúcsok kiosztásának a szórása, de ez rend kívül bonyolult matematikai feladat. Ezért mi úgy döntöttünk, hogy a mutáció legyen egy általános közép érték, ami a legvalószínűbb esetekben működik a legjobban. Ez az érték a 30%/kromoszóma

## Tesztkeretrendszer: