

ÓBUDA EGYETEM  
ÓBUDA UNIVERSITY

NEUMANN JÁNOS  
INFORMATIKAI KAR



# SZAKDOLGOZAT

OE-NIK

2022

Hallgató neve:

Hallgató törzskönyvi száma:

Szarka Levente

T/006386/FI12904/N

Óbudai Egyetem  
Neumann János Informatikai Kar  
Szoftvertervezés és -fejlesztés Intézet

## SZAKDOLGOZAT FELADATLAP

Hallgató neve:  
Törzskönyvi száma:  
Neptun kódja:

**Szarka Levente**  
T/006386/FI12904/N  
SK7EGS

A dolgozat címe:

**Biokémiai modellek implementációja ágens alapú környezetben**  
**Implementation of biochemical reactions in an agent-based framework**

Intézményi konzulens: Kiss Dániel  
Külső konzulens:

Beadási határidő: 2021. december 15.

A záróvizsga tárgyai: Számítógép architektúrák  
Szoftvertervezés és –fejlesztés  
specializáció

## A feladat

Sejtbiológiai jelenségek vizsgálata során a számítógépes modellek használata az utóbbi évtizedekben egyre általánosabbá válik, mivel a szimulációs eredmények elősegíthetik a kutatókat a folyamatok mélyebb megértésében, illetve olyan összefüggésekre mutathatnak rá, amelyek más módszerekkel csak nagyon körülmenyesen vizsgálhatók. A dolgozat célja egy ágens-alapú számítógépes modell megtervezése, implementálása és tesztelése, amelyben a résztvevő ágensek (sejtek) képesek egymásra minden mechanikai, minden kémiai kölcsönhatásokon keresztül hatni. A modellben legyen lehetőség biokémiai reakcióhálózatokat az egyedeik szintjén szimulálni, valamint a reakciókban szereplő komponensek mennyisége legyen képes befolyásolni az ágens viselkedését. Az implementáció és a tesztelés során térjen ki az optimalizációs lehetőségek vizsgálatára is.

### A dolgozatnak tartalmaznia kell:

- az ágens alapú biológiai modellezés főbb irányainak rövid összefoglalóját, a biokémiai reakciók számítógépes szimulációs módszereinek ismertetését, valamint hasonló célú rendszerek bemutatását és elemzését,
- egy választott probléma ágens-alapú modelljének megtervezését,
- a modell szoftveres implementációjának részletes ismertetését,
- a szimuláció gyorsítására használható optimalizációs technikák bemutatását,
- az elkészült szimulációs szoftver tesztelését, valamint az eredmények elemzését,
- továbbfejlesztési lehetőségek ismertetését.



Vámosy  
Dr. Vámossy Zoltán  
intézetigazgató

A szakdolgozat elvégzésének határideje: **2023. december 15.**  
(OE TVSz 55.§ szerint)

A dolgozatot beadásra alkalmasnak tartom:

.....  
külső konzulens

.....  
intézményi konzulens



## HALLGATÓI NYILATKOZAT

Alulírott hallgató kijelentem, hogy a szakdolgozat / diplomamunka saját munkám eredménye, a felhasznált szakirodalmat és eszközöket azonosíthatóan közöltem. Az elkészült szakdolgozatomban / diplomamunkámban található eredményeket az egyetem és a feladatot kiíró intézmény saját céljára térítés nélkül felhasználhatja.

Budapest, 20..21.12.15.

S. En. K. t.  
.....  
hallgató aláírása



## KONZULTÁCIÓS NAPLÓ

Hallgató neve: Neptun Kód: Tagozat:  
**Szarka Levente** **SK7EGS** **Nappali**  
Telefon: Levezési cím (pl.: lakcím):  
**+36 30 709 0613** **2022 Tahitótfalu, Nefelejcs utca 70.**

Szakdolgozat / Diplomamunka<sup>1</sup> címe magyarul:

**Biokémiai modellek implementációja ágens alapú környezetben**

Szakdolgozat / Diplomamunka<sup>2</sup> címe angolul:

**Implementation of biochemical reactions in an agent-based framework**

Intézményi konzulens: Külső konzulens:

**Kiss Dániel**

Kérjük, hogy az adatokat nyomtatott nagy betűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	2021.02.22.	Ágens alapú modellek jellemzői	Kiss Dániel
2.	2021.03.12.	Biokémiai reakcióhálózatok modellezése	Kiss Dániel
3.	2021.04.09.	Numerikus közelítő módszerek	Kiss Dániel
4.	2021.04.30.	Feladat specifikációja	Kiss Dániel

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplomamunka 2 / Diplomamunka 3 / Diplomamunka 4<sup>3</sup> tantárgy követelményét teljesítette, beszámolóra / védésre <sup>4</sup>bocsátható.

Kiss Dániel

Intézményi konzulens

Budapest, 2021. május 14.

<sup>1</sup> Megfelelő aláhúzandó!

<sup>2</sup> Megfelelő aláhúzandó!

<sup>3</sup> Megfelelő aláhúzandó!

<sup>4</sup> Megfelelő aláhúzandó!



## KONZULTÁCIÓS NAPLÓ

Hallgató neve: **Szarka Levente** Neptun Kód: **SK7EGS** Tagozat: **Nappali**  
Telefon: **+36 30 709 0613** Levelezési cím (pl.: lakcím): **2022 Tahítótfalu, Nefelejcs utca 70.**

Szakdolgozat / Diplomamunka<sup>1</sup> címe magyarul:

**Biokémiai modellek implementációja ágens alapú környezetben**

Szakdolgozat / Diplomamunka<sup>2</sup> címe angolul:

**Implementation of biochemical reactions in an agent-based framework**

Intézményi konzulens: **Kiss Dániel** Külső konzulens:

Kérjük, hogy az adatokat nyomtatott nagy betűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	<b>2021.09.16.</b>	<b>Metabolikus hálózat modellje</b>	<i>Kiss Dániel</i>
2.	<b>2021.10.13.</b>	<b>Paraméterillesztési eljárások</b>	<i>Kiss Dániel</i>
3.	<b>2021.11.02.</b>	<b>Szimulációs eredmények értékelése</b>	<i>Kiss Dániel</i>
4.	<b>2021.12.14.</b>	<b>Dolgozat szövegének véglegesítése</b>	<i>Kiss Dániel</i>

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplomamunka 2 / Diplomamunka 3 / Diplomamunka 4<sup>3</sup> tantárgy követelményét teljesítette, beszámolóra / védésre<sup>4</sup>bocsátható.

*Kiss Dániel*

Intézményi konzulens

Budapest, 2021. december 14.

<sup>1</sup> Megfelelő aláhúzandó!

<sup>2</sup> Megfelelő aláhúzandó!

<sup>3</sup> Megfelelő aláhúzandó!

<sup>4</sup> Megfelelő aláhúzandó!

# Tartalomjegyzék

<b>1. Bevezetés</b>	<b>3</b>
<b>2. Irodalomkutatás</b>	<b>5</b>
2.1. Az ágens-alapú modellek általános jellemzői . . . . .	5
2.2. Az ágensek mechanikai modellezése . . . . .	6
2.2.1. Molekuláris dinamikai módszerek . . . . .	6
2.2.2. Monte Carlo-eljárások . . . . .	8
2.2.3. Rácsalapú modellek . . . . .	9
2.2.4. Megállapítások . . . . .	10
2.3. Belső állapot jellemzők modellezése . . . . .	11
2.3.1. Pár szóban az anyagcsere modellezéséről . . . . .	11
2.3.2. A bakteriális kemotaxis és az anyagcsere modellezése . . . . .	11
2.4. Külső környezet modellezése . . . . .	12
<b>3. A modell felépítése</b>	<b>15</b>
3.1. A sejtek mechanikai modellje és mozgása . . . . .	15
3.2. Sejtosztódás . . . . .	16
3.3. Anyagtranszport a környezetben . . . . .	17
3.4. Anyagcsere a sejten belül . . . . .	18
3.4.1. A tápanyag közvetlen felhasználását feltételező megközelítés .	18
3.4.2. A tápanyag felvételét feltételező megközelítés . . . . .	19
3.4.3. Szabályozott felvétel negatív visszacsatolással . . . . .	19
<b>4. Szoftveres megvalósítás</b>	<b>20</b>
4.1. Az elkészített programok . . . . .	20
4.1.1. Az optimalizáló és a megjelenítő komponens . . . . .	20
4.1.2. A szimulátor komponens . . . . .	21
4.1.3. Implementáció . . . . .	22
4.1.4. Gyorsítás GPU-val . . . . .	22
<b>5. Paraméterillesztés és szimulációs eredmények</b>	<b>24</b>
5.1. Kísérleti adatok . . . . .	24
5.1.1. A kísérleti elrendezés . . . . .	24
5.1.2. Live cell imaging felvételek kiértékelése . . . . .	24
5.2. A modell paramétereinek becslése . . . . .	25
5.2.1. A paraméterbecslés menete . . . . .	25
5.2.2. Helyettesítő modell használata . . . . .	26
5.2.3. Finomítás az ágensalapú modellre . . . . .	28
5.3. Néhány szimulációs eredmény . . . . .	28

Összegzés, továbbfejlesztési lehetőségek	31
Kivonat	32
Abstract	33
Köszönetnyilvánítás	34
Irodalomjegyzék	35

# 1. fejezet

## Bevezetés

A kísérletek a különböző tudományterületek legfontosabb aspektusait képviselik. Arra hivatottak, hogy a különböző megfigyelhető jelenségeket képesek legyünk megtervezett módon és ellenőrzött körülmények között megismételni és levonni a megfelelő következtetéseket. A kísérletek célja lehet a tudomány és a technológia fejlődése, továbbá bizonyos jelenségek bemutatása oktatási vagy ismeretterjesztő célkitűzéssel.

Sajnos a tudományos életben számos olyan jelenség merülhet fel, amelyet a gyakorlatban nem, vagy csak nagy nehézségek árán tudunk kísérletekkel vizsgálni. Tipikusan ilyen területnek számít az asztronómia, a részecskefizika vagy a mikrobiológia, ahol a kísérletek rendkívül drágák és időigényesek lehetnek, továbbá nehezen rekonstruálhatók.

A számítógép megjelenésével lehetőség nyílt arra, hogy a megfelelő matematikai modellek implementálásával, amelyek képesek leírni egy-egy természeti jelenséget, kísérlet nélkül kapunk a valóságoshoz közel eredményeket. Ez a vizsgálati módszer a számítógépes szimuláció.

A modellek számos előnyös tulajdonsággal rendelkeznek. Célszerű úgy implementálni, hogy csak a számunkra legfontosabb komponenseket tartalmazza, ezáltal kizárhatsuk különböző zavaró tényezőket, továbbá egyszerűsíthetjük magát a rendszert. Szabadon skálázhatók, vagyis, ha nem kapunk megfelelő eredményt, akkor a kezdeti feltételek könnyen módosíthatók és újabb szimulációkat végezhetünk. A kapott eredmények ezután iránymutatásként szolgálhatnak például egy adott kísérlet tervezésénél.

A számítógépek fejlődésével egyre pontosabb, és gyorsabb szimulációk készítése válik lehetséges, azonban fontos megemlíteni, hogy sok előnye ellenére nem képesek helyettesíteni a tényleges kísérleteket. Lehet bármennyire jó a modell és annak az implementációja, az összes lehetséges tényezőt nem képes figyelembe venni, ezért kompromisszumokat kell kötnünk és ezen kompromisszumok a valós kísérletek eredményeihez képest csak egy jó közelítéssel szolgálhatnak. Azonban ezen eredmények nagyban hozzájárulhatnak a hatékonyabb kísérletek elvégzéséhez.

A dolgozatom célja egy olyan ágens alapú modell felépítése, amely egy adott biológiai sejtvonal növekedési dinamikáját képes megbecsülni a rendelkezésre álló tápanyag függvényében. Különböző részmodellek együttműködése révén az ágensek sejtszerű viselkedést produkálnak, ezáltal a szimulációk futtatásának eredményeként megfigyelhetjük a populáció növekedését.

A modell építésével párhuzamosan az A549-es sejtvonallal végzett kísérlet történt. Ebben a sejtek különböző koncentrációjú tápanyagkezelést kaptak, majd ezen

kísérletek elemzését követően a kapott eredményekből számított értékekkel paramétereztem fel a modellt, majd validáltam az eredményeket.

## 2. fejezet

### Irodalomkutatás

#### 2.1. Az ágens-alapú modellek általános jellemzői

Az ágens alapú modellezés eredete egészen Neumann János idejéig nyúlik vissza. Ő és tudóstársai azt a kérdést vizsgálták, vajon beprogramozható-e egy gép arra, hogy másolatot készítsen saját magáról. Sajnos a technológia fejletlensége miatt erről és a hasonló kérdésekről sokáig csak papíron tudtak értekezni, viszont idővel kiderült, hogy a feltevés lehetséges és megszületett az első sejtautomata. John Conway ez alapján alkotta meg híres kétdimenziós, rács alapú szimulációját az Életjátékot [1].

A sejtautomaták megjelenésével és a technológia fejlődésével egyre komplexebb tulajdonságokkal lehetett felruházni a szimuláció „sejtjeit”, ami végül a ma ismert ágens alapú modellek megjelenéséhez vezetett.

A modell olyan entitásokból áll, amelyek valamilyen előre meghatározott szabályrendszer szerint, a szimuláció során képesek autonóm cselekvésre, továbbá más entitásokkal és a környezettel való interakcióra. Ezeket az entitásokat nevezzük ágensnek [2].

Nincs egyértelmű definíció arra miképp is kell egy ágensnek viselkednie. Bizonyos nézőpontok szerint ágensnek tekinthetjük a rendszer független komponenseit, és ezek viselkedése a primitívtől az egészen komplexig terjedhet, más meglátás szerint mindenki adaptívnak kell lennie, hogy alkalmazkodhasson a környezetéhez. Bár számos elképzelés létezik a megfelelő implementációra, abban mindegyik egyetért, hogy egy ágens képes kell legyen önálló cselekvésre [1].

A modellek felépítése az adott problémától függően változhat. Az ágensek darabszáma egy szimulációban néhány darabtól egészen a milliós nagyságrendig terjedhet. A szimuláció során kezdetben elhelyezzük az ágenseket valamilyen módon abban a virtuális térben, amelyet definiáltunk, majd hagyjuk, őket a saját szabályaik szerint „létezni”. Definiálnunk kell továbbá egy időléptéket, amely a szimulációban egy „pillanatnak” felel meg. Ezen pillanatokban változtatják az ágensek az állapotukat. Gyakorlatilag minden időpillanatban minden egyes ágenst ér valamilyen inger, és az előre meghatározott szabályok szerint az adott entitás erre valamilyen módon reagálni fog. Ezek az ingerek is az adott problémától függnek, viszont esetek többségében ágens-ágens interakcióról van szó. Az interakciók eredményeit minden ciklusban valamelyen alkalmas numerikus módszerrel számítjuk ki minden egyes ágensre. Ezeknek a ciklusoknak az összessége, tehát egy szimulációs futtatás, ami gyakorlatilag egy lehetséges forgatókönyvet állít elő a vizsgált problémára [2].

Ezt a modellezési módszert már sikерrel alkalmazták méhkolóniák szerveződé-

sének vizsgálatára [3], az immunrendszer szimulálására [4], de találhatunk olyan megvalósítást is, ahol tőzsdeárfolyam-volumen viszonyt vizsgáltak [5].

Mint láthatjuk meglehetősen különböznek az alkalmazási területek, viszont ami minden esetben megegyezik az az, hogy a különböző rendszerekben az ágensek egy bizonyos szabályrendszer szerint autonóm működnek bármiféle külső beavatkozás nélkül [2].

Fontos megemlíteni, hogy a modellt a lehető legegyszerűbb módon kell leképezni, nem vehetünk figyelembe minden információt, csak, amire a legnagyobb szükségünk van. Ezért elmondható, hogy például méhek esetén nem szükséges az ágenseknek méh alakzatot tulajdonítani, bőven elég egy a méh méretéhez közelítő alakzat, ugyanis ez nagyságrendekkel egyszerűsít a modellekben, továbbá a szimulációit nem befolyásolja számottevően és legtöbb esetben az eredményül kapott viselkedés ugyanúgy felismerhető. Tipikusan kör vagy gömb alakra szokták leképezni a különböző vizsgált objektumokat[6].

## 2.2. Az ágensek mechanikai modellezése

Elmondható, hogy minden ágensnek van valamilyen, az adott probléma szempontjából fontos tulajdonsága, pontosan specifikált viselkedési szabályokkal rendelkezik, van memóriája és képes döntéseket hozni [1].

A modellben definiálnunk szükséges a környezetet és abban az adott szereplők mechanikáját. Ehhez minden ágenst felruházunk egy pozícióval, sebesség tényezővel és a feladat elvégzéséhez szükséges legfontosabb tulajdonságokkal pl. méretének megadásával [6].

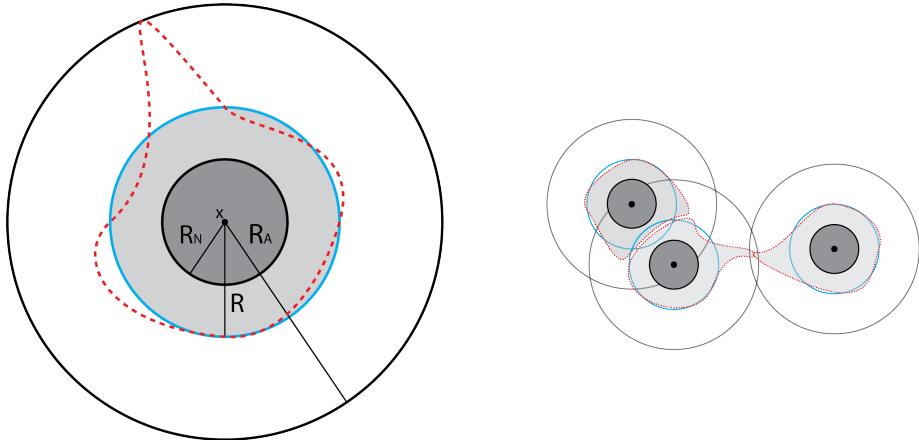
Két ágens között akkor jöhetsz létre valamilyen interakció, ha egymás közelében vannak, ütköznek, keresztezik egymás útját stb. Gyakorlatilag elmondható, hogy abban az esetben lépnek egymással kölcsönhatásba, ha képesek felismerni egy másik ágenst és ennek az ágensnek a közelébe kerülnek. Ha ezen események valamelyike bekövetkezik, akkor az előre meghatározott szabályoknak megfelelően fog reagálni [1].

Ha nem csak ágens-ágens interakciókra vagyunk kíváncsiak, akkor a modellen definiált térben létrehozhatunk olyan elemeket, amelyekkel az ágensek képesek együttműködni, akár táplálékként felhasználni egy-egy ilyen elemet, vagy akár potenciális veszélyforrásként is tekinthetnek rá, ami elől menekülniük kell. Ilyen esetben fel kell készíteni az ágenseket többféle esemény bekövetkeztére. Például egy sejt minden egyes ciklusban megvizsgálhatja, hogy a rendelkezésére álló tápanyag az éppen csak a saját életben maradásához elég, vagy esetleg van megfelelő mennyiségű többlettápanyag ahhoz, hogy osztódjon. Az ehhez hasonló vizsgálatokhoz azonban szükséges, hogy a teret is megfelelő módon definiáljuk és kezeljük [7].

Három modellezési lehetőséget vizsgáltam meg amelyek a biológiai sejtek közötti interakciókat képesek reprodukálni. Az elkövetkezőkben ezt a három stratégiát kívánom bemutatni.

### 2.2.1. Molekuláris dinamikai módszerek

Mivel a sejteket a modellekben célszerű gömbszimmetrikus alakzatokként definiálni ezért tekinthetünk rájuk erőhatások szempontjából pontszerű testekként. Ekkor, ha a rendszerünkben két ágens között csak egyszerű mechanikai hatások érvényesülnek



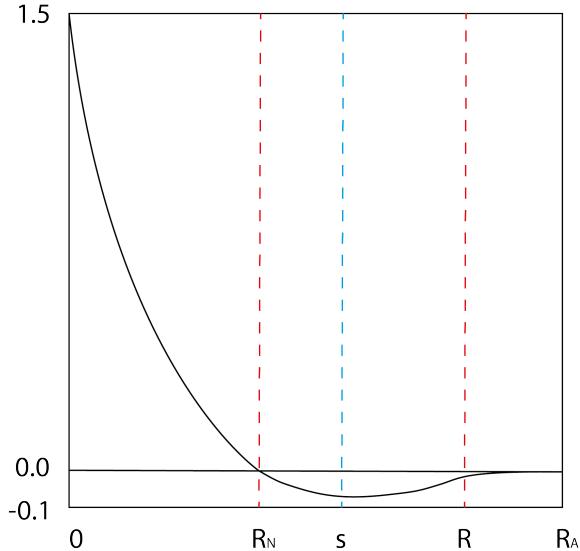
**2.1. ábra.** A képen a Macklin-féle modellben implementált sejtek sematikus ábrája látható. A baloldali ábra a sejt gömbökre történő felosztását mutatja be.  $R_N$  jelenti a sejtmagot reprezentáló gömb sugarát,  $R$  a sejt tényleges mérete,  $R_A$  pedig az adhéziós távolság, amin belül képes összetapadni két sejt. A piros szaggatott vonal szemlélteti egy valós sejt szabálytalan morfológióját. Látható, hogy a kék kör ezt a méretet próbálja közelíteni. Jobbra látható, hogyan is képezték le az adhéziós tulajdonságot. Azzal, hogy az  $R_N$  távolságon belül találkoznak azt modellezik ahogyan a való életben nyúlnának egymás felé bizonyos sejtek.

felmerül az úgynevezett n-test probléma. Az n-test probléma lényege, hogy úgy kell meghatározzuk n db test mozgását, hogy rájuk csak a Newton-féle kölcsönös gravitációs vonzóerők hatnak. Természetesen sejtek esetén csak az egymáshoz ténylegesen közel lévő ágensek lesznek képesek kölcsönhatásba lépni egy másik ágenssel ezért a probléma az interakciók számát illetően lényegesen redukálódni fog [8].

Paul Macklin munkájában a sejteket három különböző sugarú gömbbel modellezte. Ezek közül a legkisebb reprezentálja a sejtmagot, ami áthatolhatatlannak számít és két sejt esetében nem érhet össze. A középső a sejt tényleges méretét jelenti, a legnagyobb pedig az adhéziós területet, amin belül a sejtek képesek lehetnek egymást „elérni”. Gyakorlatilag ez tekinthető egy olyan határnak, amit a való életben egy sejt képes lenne elérni úgy, hogy morfológiaileg rövid időre megváltozik. A sejtmodell a 2.1 ábrán látható. [6].

A sejtek közti interakciók lehetőségeit egy potenciálfüggvénnyel írták fel, ami meghatározza két sejt esetén, hogy a köztük lévő távolság alapján makkora a helyzeti energia nagysága. A függvényt a 2.2 ábrán láthatjuk. Abban az esetben, ha nagyon közel kerülnek egymáshoz a függvény alapján az energia nagyon magasra ugrik. Ebben az esetben a sejtek tasztítani fogják egymást. Ha nagy a távolság két sejt között akkor ugyanezen analógiá mentén a függvény alapján ez az energia 0 lesz és semmi nem fog történni. Azonban, ha az adhéziós távolságaikon belül találkoznak akkor vonzó erővel fognak hatni egymásra. A lépések követően kialakult vonzó és tasztító események váltakozása idővel stabilizálódik és azt a hatást kelti, hogy a két sejt összetapadt [6].

A modellben a vonzó és tasztító tulajdonságokon túl a sejtek rendelkeznek életciklussal. Ez azt jelenti, hogy öregszeneik, tápanyag hiányában elpusztulnak, megfelelő mennyiségű és minőségű erőforrás birtokában pedig képesek az osztódásra. [6].



**2.2. ábra.** A Macklin modellben szereplő interakciós függvény. Látható, hogyha  $R_N$  távolságon belül kerülnek egymáshoz a sejtek, akkor az energia megugrik és taszítani fogják egymást. Azonban ahogy távolodnak az energiaszint egy idő után 0 lesz.  $R_N$  és  $R_A$  között történik meg a vonzás. Az  $s$  egy vonzás és tasztás közti egyensúlyi állapotot jelöl, ezen a ponton tekinthetünk két sejyet összetapadtnak.

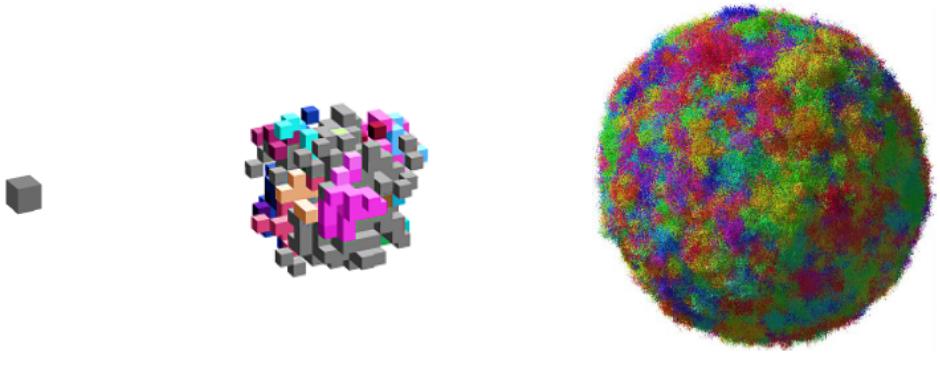
## 2.2.2. Monte Carlo-eljárások

Ha a makroszkopikus világot szeretnénk vizsgálni, célszerű lehet elvonatkoztatnunk a klasszikus mechanikai interakcióktól. Bár sejtek esetén a mozgássegyenletek felállítása nem lehetetlen feladat, mégis nagy bonyolultságú és nagy számítási teljesítményt igénylő műveletsorozatok alakulhatnak ki egy nagyobb populáció szimulációjákor. Ebből kiindulva egy járható út lehet, a különböző random kiválasztáson alapuló módszerek használata. Ez az irány a valószínűségszámításon alapul. Alap felvetése, hogy a teljes rendszer energiáinak változását vizsgáljuk, véletlenszerűen kiválasztott ágensek elmozdítását követően és ezen változásokból következtetünk a tényleges fizikai eredményekre [9].

Ha az ágens alapú modellt a klasszikus fizikai interakciók helyett valószínűségi alapokon szeretnénk kezelni, akkor egy járható út lehet, ha Monte Carlo eljárásokat alkalmazunk. Akkor használjuk a Monte Carlo szimuláció elnevezést, ha a folyamat során véletlen értékeket használunk. Ezeket a véletlen értékeket természetesen egy a feladat szempontjából alkalmas valószínűségi eloszlás alapján határozzuk meg. A sok véletlen kezdőértékkel rendelkező futtatás eredményét a szimulációk végén összesítik és általános következtéseket próbálnak levonni az adott problémával kapcsolatban. Fontos megjegyezni azonban, hogy a számítógéppel előállított véletlen számok csak nagyon jó közelítések, tökéletes véletlenszám generátor nem létezik [9].

D. Drasdo és csapata ezen megközelítéssel alkotta meg saját modelljét és végzett szimulációkat. Modellükben a sejtek között nem mechanikai interakciókat, hanem a teljes rendszer energiájának változását vizsgálták az ágensek elmozdításának függvényében [10].

A szimuláció során először meghatározzák a sejtek helyzeti energiáját, ezeknek összege adja a rendszer teljes energiaszintjét. Választanak egy sejyet, akár véletlenszerűen, akár valamilyen metódus szerint, majd ezt a sejyet elmozdítják. Az elmozdítást követően változik az egész rendszer energiája, tehát újra számolják az ágensek energiáit, és újra megvizsgálják a rendszer energiaszintjét. Ha az elvárásoknak megfelelően változott az eredmény, akkor tovább lépünk a következő szimulációs ciklusba, ha nem, akkor semmisnek tekintik az előbbi elmozdítást, visszaállítják a rendszert és újabb elmozdítással próbálkoznak. A rendszer az energiaminimum elve szerint működik, továbbá a meghatározott potenciálzabályok alapján akkor változik



**2.3. ábra.** Rácsalapú modell, amelyben apró kockák szimbolizálják a sejteket. Balra egyetlen sejtet látunk, középen egy pár száz elemet tartalmazó populációt, jobbra pedig milliós nagyságrenddel szerepelnek az ágensek. Alacsony számú ágens esetén nem túl szemléletes a modell, azonban a milliós nagyságrendnél már egy sokkal szebb végeredményt kapunk. Forrás: [11]

jó mértékben a teljes energiaszint, ha az ágensek közelednek egymáshoz [10].

Az ehhez hasonló szimulációknál nem lehetséges általános definíciót adni arra, hogy egy iterációs ciklus mégis mennyi időt jelent a valóságban. Ezt általában tapasztalati úton határozzák meg rengeteg szimuláció futtatásával, amit a valós kísérletek eredményeivel vetnek össze. Ebből kifolyólag jónéhány aspektusa ennek a modellnek nem tervezhető meg előre [10].

### 2.2.3. Rácsalapú modellek

Ahhoz, hogy a biológiai modellek minél realisztikusabbak legyenek, a lehetőségekhez mérten a leg pontosabb szabályok szerint kell őket megalkotnunk, viszont emiatt például egy ágens alapú modellnél minden plusz vizsgálandó tényezővel nagy mértékben nő a szimulációk számításigénye. Ez egyaránt jelent többletmemóriaköltséget és plusz számítási időt [11].

Ennek megoldásaként felállíthatók olyan modellek, amelyekben egy rácssonkerkezet implementálásával számítási teljesítményt spórolhatunk meg. Ennek a rácsnak a felhasználása többféle módon történhet.

Az egyik ilyen felhasználásban az Eden féle növekedési modellt vették alapul. Ez a növekedési modell a baktériumtelepekhez hasonló szerveződések növekedését írja le. Gyakorlatilag a telepek széleinél kerül sor új sejtek létrejöttére. Ez azért fontos, mert a sejtek így csak a közvetlen a szomszédokkal kerülhet interakcióba [11].

Bartłomiej Waclaw egy olyan modellt alkotott, amelyben a tér egy háromdimenziós rácsként jelenik meg és egy kocka, amit így kapunk reprezentál egy sejtet. Ezzel a megoldással az ágens a rendszer egy rendkívül egyszerű elemévé válik. Nem képes mozogni és csak a szomszédos cellákkal kerülhet interakcióba, azonban az Eden féle növekedési modell szimulációjához ez bőven elegendő. Ez a fajta megközelítés nagyon jól képes szimulálni például egy rosszindulatú sejtburjánzást. A 2.3 ábrán a Waclaw féle modellt láthatjuk különböző felbontásokban. [11]

Ezt az implementációt sejtpopulációk növekedésének vagy mutációjának vizsgálatára lehet a leginkább célszerű felhasználni. Ezzel a módszerrel sok számítási erőforrást spórolhatunk, mert az interakciók nagy része a szerveződések szélén történik. Továbbá mivel a terünk egy rácsszerkezetnek felel meg, ezért minden csak a környező rácsokból kell interakcióra számítanunk. Mivel a modell mondhatni elnagyolt, olyan tekintetben, hogy egy kitöltött cella szimbolizál egy sejtet, a szimuláció során nagyobb hangsúlyt tudunk fektetni más aspektusokra, például a mutációra vagy az osztódás folyamatára [11].

Más megvalósításban is láthatunk rácsszerű felbontást. Például Heather Kelly szakdolgozatában a teret ugyancsak feldarabolta. Az általa vizsgált részecskék ezekben a kockákban voltak jelen és azt a szabályt hozta, hogy két részecske csak akkor hatthat egymásra, ha legalább a szomszédos cellákban helyezkednek el. Mivel ebben az esetben kizártatók az egymástól túl távol elhelyezkedő rézsejtek által okozott hatások, ezért a számítási költségek csökkenhetnek [12].

#### 2.2.4. Megállapítások

A fent ismertetett három implementációs technika bár minden ágens alapú megközelítést használ, teljesen más elképzeléssel rendelkeznek a sejtek definícióját és interakcióinak vizsgálatát illetően.

A mechanikai megközelítéssel egy átlátható, klasszikus fizikai elképzeléseken alapuló modellt kapunk, ahol a bonyolultságot a rendszer elemeire ható rengeteg erő hatékony kiszámítása adja. Ha eléggé leegyszerűsítjük a modellt akkor ezen számítások költségei nagy mértékben csökkenhetnek [1].

A random kiválasztáson alapuló módszer és a matematikai valószínűség bevezetésével láttunk olyan implementációt, amely mechanikai tulajdonságok középpontba helyezése helyett potenciális energiákkal ruházta fel az ágenseket, és a rendszer teljes energiáját vizsgálva hozott döntést arról, hogy bizonyos változtatásokat elfogadunk, vagy sem [10].

A harmadik elképzelés, hogy valamilyen rácsszerkezzel felbontjuk a vizsgált teret és egy cella kitöltött állapotban reprezentál egy sejtet. Ez rendkívül leegyszerűsíti a legtöbb számítást, mert csak a szomszédos cellákat kell konkrét interakció szempontjából figyelembe venni, továbbá alkalmazhatóak olyan halmozódásra hasonlító eljárások, mint pl. a fent említett Eden féle növekedési modell [11].

Minden modellnek van előnye és hátránya egyaránt, továbbá nem alkalmazható mindenkor modell ugyanolyan hatékonysággal ugyanarra a problémára. Vegyük például a rács alapú megközelítést, ahol a sejt egy négyzet vagy kocka szintjére egyszerűsödött le, viszont pont emiatt definiálhatók voltak olyan plusz tulajdonságok, amelyek alapvetően nem függnek a sejt alakjától, tehát elmondható, hogy a probléma szempontjából érdektelen információk redukálásával más fontosabb dolgokra fektethették a hangsúlyt. Azonban bizonyos esetekben elengedhetetlen az ennél pontosabb morfológiai tulajdonságokkal felruházni az ágenseket [7].

Mivel a modellem szempontjából fontos, hogy a sejtek interakcióba léphessenek egymással, továbbá képesek legyenek a mozgásra, időben jól skálázható legyen, könnyen kiegészíthető és kerekítési hibáktól minél inkább mentes ezért dolgozatomban a Macklin féle molekuláris dinamikai modellt valósítottam meg.

## 2.3. Belső állapot jellemzők modellezése

### 2.3.1. Pár szóban az anyagcsere modellezéséről

Ahhoz, hogy a tápanyag függvényében megfigyelhető növekedési dinamikát modellezni tudjuk be kell építenünk az ágensekbe egy komponenst, ami elvégzi a különböző agyagcserefolyamatokat, továbbá figyelembe kell venni a kemotaxis jelenségét.

Az anyagcsere modellezésére két elterjedt megközelítés létezik. Az egyik egy afféle táblázatos módszer, amiben mátrixok segítségével írhatjuk fel a különféle reakciókat. Ebben az esetben gyakorlatilag a mátrix egy adott sora egy adott reakciót ír le. A másik elképzelés pedig a kinetikus modellezési módszer, amelyben differenciálegyenleteket írhatunk fel az anyagcserefolyamatban szereplő anyagok koncentrációjának időbeli változására [13].

A táblázatos módszert akkor célszerű használni, ha egyszerre sok reakciót dolgozunk és ezen reakciók jól meghatározható paraméterekkel rendelkeznek. Gyakorlatilag az egyes reakciók, mint állapotok jelennie meg a rendszerben. A kinetikus modell esetén szükségünk van a reakció sztöchiometrikus egyenletére, amivel megmondjuk, hogy adott anyagból milyen egyéb anyag, vagy anyagok keletkeznek a rendszerben. Ennek tudatában felírható az adott anyag koncentrációjának változása az idő függvényében és ez egy alkalmas numerikus módszerrel ki is értékelhető [13].

Mivel az én esetben nincs szükség csak néhány reakcióra az anyagcsere modellezéséhez, ezért számomra megfelelőbb a kinetikai modell használata. Matthew Egbert kutatásában megalkotott egy egyszerű mégis szemléletes kinetikus modellt az anyagcsere reprodukálására, továbbá kitért a kemotaxis modellezési lehetőségeire.

### 2.3.2. A bakteriális kemotaxis és az anyagcsere modellezése

Egy biológiai sejt parányi mérete ellenére egy rendkívül összetett szerveződés. Képes felismerni, hogy a környezetében lévő anyagok számára kedvező, vagy éppen ellenkezőleg, számára káros tulajdonságokkal bírnak, továbbá képes felismerni ezeknek az anyagoknak a koncentrációkülönbségét és mozgásuk abba az irányba orientálódik, amerre fisiológiai szempontból számára kedvezőbb koncentrációk vannak jelen. Ezt a viselkedést, amikor egy baktérium, vagy sejt, számára kedvezőbb kémiai anyagok felé vagy rá károsan ható anyaguktól távolodva mozdul el kemotaxisnak nevezzük [14].

Ahhoz, hogy a kemotaxis jelenségét felismerhessük a szimulációban az ágenseket fel kell ruházni bizonyos érzékelési képességekkel, hogy képesek legyenek felnérni a környezetükben lévő anyagok koncentrációgradiensét, továbbá, hogy ez az anyag számára kedvező, vagy káros. Emellett szükség van a sejt anyagcseréjének modellezésére, ami alapján képes a tápanyagok hasznosítására, energiatermelésre az osztódáshoz, mozgáshoz és az életben maradáshoz [14].

A kemotaxist alapvetően anyagcserefüggetlen folyamatként tartják számon. Ez azt jelenti, hogy az ágens elmozdulását csak a koncentrációgradiens mértéke és a kémiai anyag típusa határozza meg, az, hogy a metabolikus folyamatok milyen állapotban vannak a sejten belül nem. Azonban több olyan esettel is szembe találták magukat a biológusok amikor olyan mintákat fedeztek fel kísérletek során, amelyek az anyagcserefüggőség jeléül szolgálhatnak. Ilyen minta például, hogy ha a térben homogén módon eloszlik egy a sejt számára kedvező anyag, és egy forrásból beáramlik egy másik számára kedvező anyag, nem indul el az új anyag nagyobb gradiense

felé [14].

Matthew Egbert modellje megalkotásakor továbbáment és egy olyan elképzést állított fel miszerint a kemotaxis modelljéből elhagyja a szenzoros tulajdonságokat, magát a kemotaxist az ágensek anyagcserefolyamatainak eredményeként kívánta bemutatni. A 2.4 ábra szemlélteti a kemotaxis modellezési lehetőségeit az anyagcsere szempontjából [14].

Egberték célja az volt, hogy a szimulációk futtatását követően az ismert kemotaktikus minták felfedezhetők legyenek az eredményekben és ez meg is valósult. Felállítottak öt forgatókönyvet, amit egy valós biológiai kísérlet is eredményül szolgáltatna és ezeknek eleget is tettek a szimulációik [14].

Ahhoz, hogy a kemotaxist ily módon modellezni lehessen anyag és energiaáramláusra van szükség, továbbá egy szerveződésre, ami elegendő erőforrás tartós hiányában megszűnik. Egberték modelljében ezt a szerveződést baktériumok alkotják. Implementálták a baktériumokra jellemző „úszó – bukdácsoló” mozgástípusokat, és ezeket az anyagcserefolyamatok eredményei befolyásolták [14].

Az anyagcsere egy absztrakt megközelítése a 2.1-es reakcióegyenlettel írható fel, ahol  $M$  az anyag,  $E$  az anyag feldolgozásából felszabaduló energia,  $C$  a keletkező katalizátor  $W$  pedig veszteségi energia. Továbbá a rendszer autokatalikus, tehát a katalizátor koncentrációjának függvényében visszahat a rendszerre és gyorsítja az anyagcserefolyamatokat. A rendszerbe továbbá felvettek egy  $S$  paramétert, ami anyagcseregátló anyagot reprezentál. Ez, ha reakcióba kerül  $E$ -vel, akkor a keletkezett  $SE$  toxin az ágens szempontjából káros anyagot fog jelenteni. Ez a 2.2-es reakcióegyenletben látható. [14].



A különböző reakcióegyenleteket numerikus módszerrel közelíthető differenciál-egyenletekké alakítják. Ezek az egyenletek megadják a különböző anyagok koncentrációját iterációként [14].

Ez a fajta megközelítés egyrészt egy érdekes kérdést vet fel a kemotaxis és az anyagcsere kapcsolatáról még biológiai vonatkozásban is, másrészt pedig elhagyhatónak tekinti a szenzoros komponenseket, ezáltal egyszerűbb és szemléletesebb modellt kapunk.

Mivel az Egbert által felírt anyagcsere egy egyszerű absztrakt megközelítés, ezért nem csak baktériumok, hanem sejtek esetén is használható, ezért dolgozatomban ezt az irányt választottam.

## 2.4. Külső környezet modellezése

Annak érdekében, hogy a kemotaxis modell teljes legyen, definiálnunk kell egy teret, amelyben az ágensek létezhetnek, és ebben a térben egy vagy több anyag diffúzióját kell leírnunk. Ezt gyakorlatilag egy külön rétegként képzelhetjük el, ahol megtörténik az anyagok áramlása és ezt a réteget fogják az ágensek a pozíciójuk alapján vizsgálni majd további lépni a számukra megfelelő koncentrációk felé [14].

A diffúziót kétdimenzióban a 2.3-es folytonos egyenlet írja le, amit Laplace egyenletnek nevezünk.  $U$  az anyag koncentrációja,  $D$  a diffúziós együttható,  $t$  az adott időpillanat,  $x$  és  $y$  pedig az adott koordinátákat jelentik [15].

$$\frac{\partial U}{\partial t} = D \left( \frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} \right) \quad (2.3)$$

Az egyenlet analitikus megoldásával képesek lennénk pontosan meghatározni a tér bármely pontjában az adott koncentrációt, azonban az esetek többségében nem lehetséges megfelelő módon kivitelezni számítógép segítségével az analitikus számításokat, ezért valamelyen módon célszerű diszkretizálni a problémát. Ennek legegyszerűbb módja, ha fix nagyságú rácsot alakítunk ki a térben. Ekkor a különböző rácspontok koncentrációjukat a szomszédos négy rácspont koncentrációja alapján határozzák meg [15].

Ezt követően az egyenletet egy alkalmas numerikus sémával kiértékelik minden egyes rácspontra. Mivel célszerű minél sűrűbb rácsot választani a pontosság növelésének érdekében, ezért az iterációkénti számítások ideje jelentősen megnőhet. Ennek csökkentése érdekében célszerű párhuzamosítási módszereket eszközölni [15].

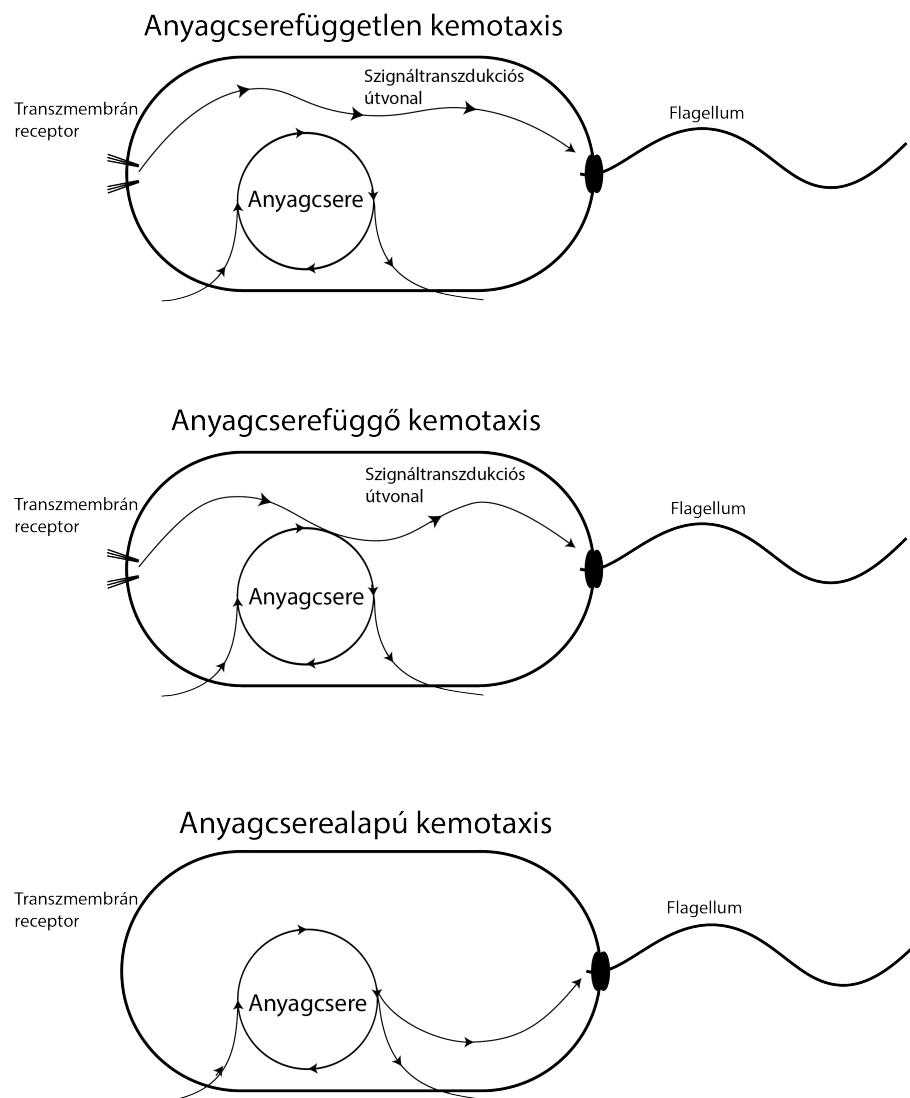
Egy diszkretizálási módszer, aminek a neve FTCS séma, a 2.4 egyenlet formájában szolgáltatja a Laplace egyenletet. [15].

$$\frac{u_{i,j}^{(n+1)} - u_{i,j}^{(n)}}{\Delta t} = D \left[ \frac{u_{i+1,j}^{(n)} - 2u_{i,j}^{(n)} + u_{i-1,j}^{(n)}}{(\Delta x)^2} + \frac{u_{i,j+1}^{(n)} - 2u_{i,j}^{(n)} + u_{i,j-1}^{(n)}}{(\Delta y)^2} \right] \quad (2.4)$$

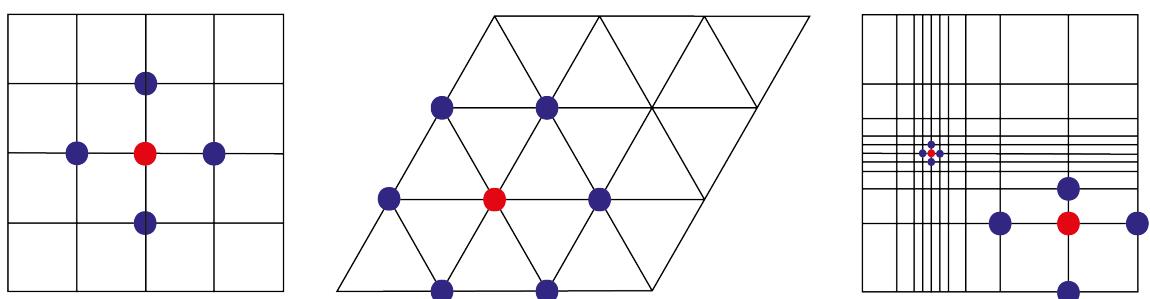
A diffúzió modellezésének nem ez az egyetlen modellezési lehetősége. Létezik például egy olyan elképzelés, amely nem négyzetes, hanem háromszöglapú rácsra képezi le a problémát. Ebben az esetben az adott pontot, amelyre a diffúziót ki akarjuk számolni, hat szomszédja befolyásolja. Belátható, hogy a négyzetes rácshoz képest ez pontosabb megoldást adhat, azonban a háromszöges felbontás miatt komplexebb számításokat kellene végrehajtanunk, továbbá a vizsgált tér leképzése is nehezebb feladat [16].

Egy másik lehetséges megoldás hasonló a négyzetes rácokat használó modellhez, azonban a fő elképzélés az, hogy bizonyos esetekben a tér egy adott pontján nagyobb pontosságra lehet szükségünk más eseménymentes pontokon pedig kevésbé. Ebből az ötletből alakították ki a mozgó rács elképzelését, ami képes változtatni a különböző rácspontjainak a nagyságát annak megfelelően, hogy ott történik valamilyen esemény, aminek szüksége lenne a nagyobb pontosságra vagy sem [17].

Ez a megoldás egyrészt javíthat a pontosságán a szimulációnak, továbbá javíthat a számítási költségeken is, viszont a megvalósítása meglehetősen bonyolult és csak olyan rendszerekben érdemes használni, ahol csak néhány objektum viselkedését figyelnénk meg egyszerre. Mivel a rácst átméretezésével egy új számítási komponens kerül a rendszerbe az ágensek számától függően kell vizsgálni és átméretezni a rácscsot és nem biztos, hogy a modell előnyei kompenzálni tudják a számítási költségek okozta hátrányokat. A 2.5 ábra szemlélteti ezt a három leképzési módszert. [17].



**2.4. ábra.** A kemotaxis modelezési lehetőségei az anyagcsere szempontjából baktériumok esetén. A transzmembrán receptor felel a koncentrációkülönbségek detektálásáért, a flagellum biztosítja a baktérium mozgását, a szignáltranszdukciós útvonal a megfelelő ingert továbbítja a flagellumoknak annak megfelelően, hogy milyen anyagok vannak a környezetben. Mint látjuk az anyagcserefüggetlen és függő esetekben a különbség, hogy az anyagcsere befolyásolhatja a flagellumoknak kiadott parancsot és ezáltal a mozgást. Egberték anyagcserén alapuló megoldása pedig csak az anyagcse-refolyamatokat veszi számításba, a többi komponenst elhagyták a modellből.



**2.5. ábra.** Hárrom lehetséges leképzési módszer a diffúzió számításához. Az ábrákon a piros pont jelenti az adott pillanatban vizsgált koncentrációt, a kék pontok pedig azok a koncentrációk amik alapján kiszámítjuk a piros következő pollanatbeli értékét. Balról jobbra az egyszerű négyzetes leképzést, majd a háromszöges megoldást és végül a mozgó rácsos elképzelést láthatjuk.

## 3. fejezet

# A modell felépítése

A teljes modell több komponens együttműködése révén valósítja meg a feladatot. minden komponens egy-egy különálló modell, ami a feladat egy részproblémáját oldja meg. A sejteket ágens alapú megközelítéssel implementáltam. Képesek a számukra definiált térben mozogni, egymással fizikai interakcióba lépni, anyagcserefolyamatokat lebonyolítani, vagyis tápanyagot felvenni és átalakítani, továbbá osztódni. minden sejt rendelkezik egy sejtciklussal, továbbá tápanyagszükséglettel az osztódáshoz.

A teljes modell réteges felépítését a 3.1 ábra szemlélteti.

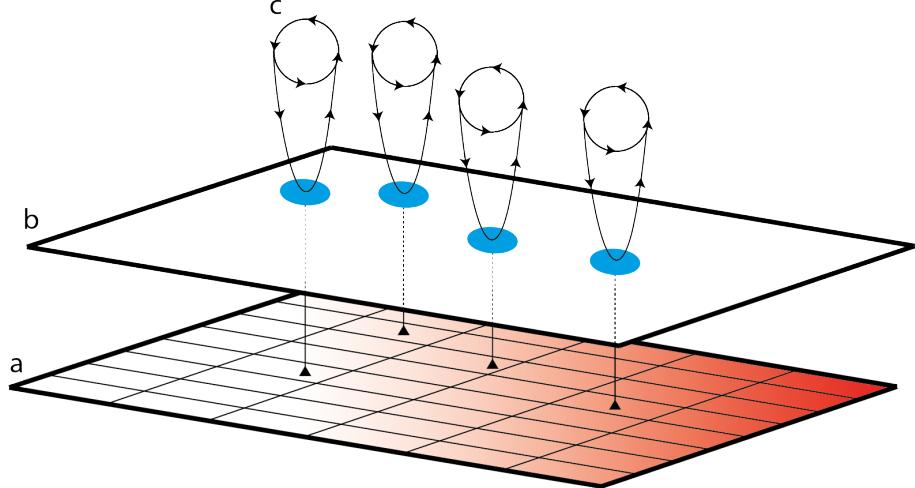
### 3.1. A sejtek mechanikai modellje és mozgása

A legfontosabb eleme a rendszernek a sejtek mozgásáért, egymás közti interakciójáért és az életciklusukért felelős modell. Fontos volt, hogy olyan modellt válasszak, ami ezeket a funkciókat egyszerűen és szemléletesen képes megvalósítani, továbbá legyen viszonylag könnyen beilleszthető az ágensek mindegyikébe egy anyagcsérét reprezentáló modell.

Háromféle lehetőséget vizsgáltam meg amelyekkel a sejtek interakciós tulajdonságait tudom modellezni. Ezek közül véleményem szerint a kemotaxis modellezéséhez a molekuláris dinamikai modell illik a legjobban. A Monte-Carlo eljárásnak alapuló módszer kevésbé determinisztikus továbbá a metabolizmus nehezen beépíthető az alap elgondolásba [10]. A rács alapú modellek esetén pedig a sejt mozgása vagy nem implementált tulajdonság, vagy pedig diszkrét értékeken történhet [11]. Az általam épített modellemben azonban szükség van arra, hogy különálló sejtek tudjanak elmozdulni bizonyos irányokba, és szerettem volna, ha ez az elmozdulás nem egy diszkrét rácson történik.

A molekuláris dinamikai modell képes összetettebb képességekkel rendelkező ágensekkel kezelní a sejteket és könnyebben beépíthető a metabolikus modell a sejtekbe. A mechanikai interakciók révén létrejöhetsz a sejtpopulációk. A Monte-Carlo eljárásokkal szemben itt sokkal egyszerűbb az időt definiálni, továbbá a diffúziót könnyebben figyelembe tudják venni ebben a modellben a sejtek [6].

Az ágensek közötti interakciókért a 3.1 és a 3.3 függvények felelnek. Ezek határozzák meg a sejtek közötti interakciós energiákat és eredményül egy erőt kapunk, aminek hatására elmozdul a sejt. Ezeket a számításokat minden ágens minden környezetében lévő ágenssel elvégzi. Az egyenletekben a  $c_{cca}$ ,  $c_{ccr}$ ,  $m$ ,  $M$ ,  $n$  a vonzásért és a taszításért felelős mérési eredményekből származó konstansok. Ezeket kísérletek alapján határozzák meg, továbbá  $f_{i,j}$  a vonzás leírását segítő együttható.  $\nabla\varphi$  és  $\nabla\psi$



**3.1. ábra.** A komplett szimulációs modellhez szükséges komponensek réteges elrendezése. Az *a* reprezentálja a diffúziós réteget itt áramlik az anyag. A képen a piros szín jelenti a nagyobb koncentrációt. Ebből a rétegből párhuzamosan több is elhelyezkedhet attól függően hányfélé anyagot kezelünk a rendszerben. A *b* az ágensek rétege, itt történik a sejtek mozgása, a köztük lévő interakciók és a sejtciklus által végbement változások (pl.: osztódás vagy halál). A *c* "réteg" lenne az anyagcserefolyamatokért felelős. Mint látható, az ágensek kapcsolatban vannak a diffúziós réteggel, folyamatosan figyelik a koncentrációkat és ha tudnak, anyagot vesznek fel. A metabolikus rétegen pedig az ágensek által felvett anyagok feldolgozása történik és az anyagcserefolyamatok eredménye visszahat az ágens mozgására és sejtciklusára.

láthatók a 3.2 és a 3.4 függvényekben kifejezve.  $\mathbf{x}$  az *i*-edik sejt helyvektorát jelenti,  $R_A$  és  $R_N$  a sejtre jellemző adhéziós és nukleusz távolságot. [6] Mint látjuk ez voltaképp egy szabályrendszer, ami megmondja mi történjen, ha bizonyos távolságon belül kerülnek a sejtek egymáshoz képest.

$$F_{cca}^{ij} = -c_{cca} f_{i,j} \nabla \varphi(\mathbf{x}_j - \mathbf{x}_i; R_A^i + R_A^j, n_{cca}) \quad (3.1)$$

$$\nabla \varphi(\mathbf{r}; R_A, n) = \begin{cases} \left(1 - \frac{|\mathbf{r}|}{R_A}\right)^{n+1} \frac{\mathbf{r}}{|\mathbf{r}|}, & 0 \leq |\mathbf{r}| \leq R_A \\ \mathbf{0} & \text{különben} \end{cases} \quad (3.2)$$

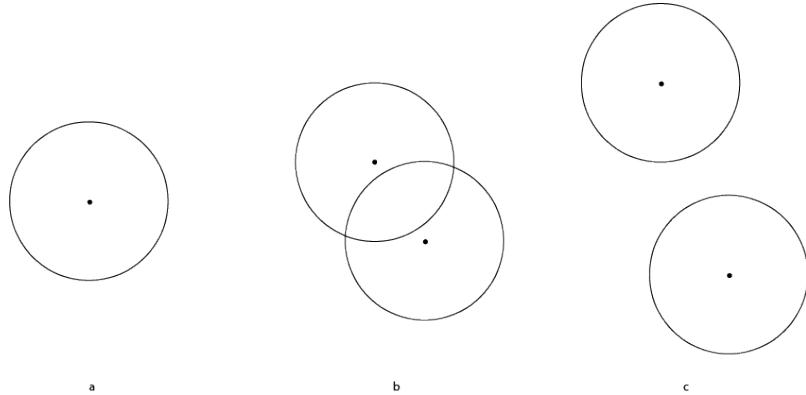
$$F_{ccr}^{ij} = -c_{ccr} \nabla \psi(\mathbf{x}_j - \mathbf{x}_i; R_N^i + R_N^j, R_i + R_j, n_{ccr}) \quad (3.3)$$

$$\nabla \psi(\mathbf{r}; R_N, R, M, m) = \begin{cases} -\left(c \frac{|\mathbf{r}|}{R_N} + M\right) \frac{\mathbf{r}}{|\mathbf{r}|}, & 0 \leq |\mathbf{r}| \leq R_N \\ -\left(1 - \frac{|\mathbf{r}|}{R}\right)^{m+1} \frac{\mathbf{r}}{|\mathbf{r}|}, & R_N \leq |\mathbf{r}| \leq R \\ \mathbf{0} & \text{különben} \end{cases} \quad (3.4)$$

$$c = \left( \left(1 - \frac{R_N}{R}\right)^{m+1} - M \right) \quad (3.5)$$

## 3.2. Sejtosztódás

A sejtek osztódása több tényezőtől függ. Két fontos változót tárolnak el az ágensek. Az egyik az életkorukat határozza meg, míg a másik a felvett tápanyag pillanatnyi mennyiségét. Az osztódás ennek a két változónak az értékétől függ. Amikor egy ágens életkora elérte a kettőződési idő értékét (ez az időtartam sejttípustól függ),



**3.2. ábra.** Az osztódás a modellben két belső változótól függ. Az egyik a felvett tápanyag mennyiségét tárolja, a másik a sejt életkorát. Az *a* ábra mutatja a sejt általános állapotát. Amennyiben elérte egy megfelelő életkort, rendelkezik a szükséges tápanyaggal, továbbá a környezetében fellelhető elég szabad terület az osztódáshoz, akkor a *b* ábrán látható módon ezt véghez is viszi. A *c* ábra reprezentálja, hogy a két sejt külön létezik tovább.

akkor alapvetően megtörténne az osztódás. Ahogyan azt a következő szakaszban pontosítom, az osztódáshoz az életkor elérésén felül szükséges még, hogy megfelelő mennyiségű tápanyag álljon az ágens rendelkezésére. Tápanyag hiányában a populáció növekedése lelassul, vagy leáll. Ez a viselkedés modellezhető például a 3.6 képpel leírt sigmoid függvény használatával.

$$\frac{1}{1 + e^{-\sigma(x-\mu)}} \quad (3.6)$$

Ez a képlet értelmezhető egy osztódási valószínűséggént a rendelkezésre álló tápanyag mennyiségeknek ( $x$ ) függvényében. A  $\mu$  küszöbérték alatti koncentrációhoz egyre alacsonyabb osztódási valószínűségek tartoznak, míg nagy  $\mu$  esetén az osztódás valószínűsége közel 1. A  $\sigma$  paraméter a küszöbérték körüli átmenet meredekségét szabályozza.

Abban az esetben, ha ezek a feltételek adottak a sejtnak meg kell vizsgálnia, hogy megfelelő mennyiségű hely áll-e rendelkezésére az osztódáshoz. A sejtmagok nem fedhetik egymást, továbbá tudjuk, hogy ha legközelebb kerülnek egymáshoz akkor ideális esetben hexagonális elrendezést vehetnek fel, ebből adódóan a közvetlen környezetében maximum 6 sejt tartózkodhat. Az osztódás csak akkor következik be, ha ez a feltétel teljesül.

A sejtosztódást a 3.2 ábra szemlélteti.

### 3.3. Anyagtranszport a környezetben

A szimulációban szükség van egy vagy több forrásra, amiből az ágensek anyagcserejének megfelelő anyagok áramolnak a rendszerben. Ezt egy diffúziós komponenssel valósítom meg. A 2.4-es fejezetben bemutatott modell alapján én is a Laplace egyenlet egy diszkretizált formáját implementálom. Ez a módszer egy rácson helyezi el a térből a különböző koncentrációkat és minden rácspont a négy közvetlen szomszédja alapján határozza meg a saját koncentrációját iterációként [15].

A diffúzió modellezésének egyéb lehetőségei közül ez az egyik legegyszerűbb megközelítés, bár ennek megfelelően, ha nem megfelelő paraméterekkel ruházzuk fel könnyen adhat hibás eredményeket.

További lehetőségek közé tartozik még, hogy a rácsnak a leképzése nem négyzletes, hanem háromszöges alakra vezethető vissza[16] vagy maga a rács dinamikus és bizonyos pontokban besűrűsödik annak érdekében, hogy pontosabb eredményekkel szolgáljon [17].

A választásom azért esik az egyszerű négyzetes rácokra való leképzésre, mert diffúzió szempontjából az a pontoság, amit megad, az én kontextusomban elegendő, kevésbé bonyolult, mint a háromszöges megoldás, továbbá jóval kevesebb kalkulációra van szükség hozzá, mint a mozgó rácsos megoldáshoz.

Egy rácspont koncentrációjának kiszámítását a 3.7 egyenlet mutatja be.

$$u_{i,j}^{(n+1)} = u_{i,j} + D\Delta t \left[ \frac{u_{i+1,j}^{(n)} - 2u_{i,j}^{(n)} + u_{i-1,j}^{(n)}}{(\Delta x)^2} + \frac{u_{i,j+1}^{(n)} - 2u_{i,j}^{(n)} + u_{i,j-1}^{(n)}}{(\Delta y)^2} \right] \quad (3.7)$$

## 3.4. Anyagcsere a sejten belül

### 3.4.1. A tápanyag közvetlen felhasználását feltételező megközelítés

A sejtek anyagcseréje egy bonyolult, jelenleg sem teljesen tisztázott biokémiai reakciósorozat. Lényege nagyon leegyszerűsítve az, hogy a sejt a környezetében elérhető tápanyagot felveszi, hasznosítja, majd a keletkezett melléktermékeket kiválasztja a környezetbe. Ezt a folyamatot legegyszerűbben az alábbi fiktív biokémiai reakcióval írhatjuk le.



Az  $S$  tápanyagot (*substrate*) a sejt átalakítja a  $P$  termékké (*product*). A reakció sebességét az  $r_m$  metabolikus ráta jellemzi. A reakcióegyenlettel leírt koncentráció-változást differenciálegyenletekkel felírva kapjuk a

$$\frac{d[S]}{dt} = -r_m [S] \quad (3.9)$$

$$\frac{d[P]}{dt} = r_m [S] \quad (3.10)$$

összefüggéseket. Mivel ezek a reakciók a sejtekre egyénileg értelmezettek, ezért az egyenletek közelítését gy végzem el, hogy az  $A_i$  ágens által lefedett  $(x, y)$  rácspontokban lévő koncentrációkat állítom be minden szimulációs lépésben az

$$S_{x,y}^{t+\Delta t} = -r_m S_{x,y}^t \Delta t \quad (3.11)$$

$$P_{x,y}^{t+\Delta t} = r_m S_{x,y}^t \Delta t, \quad (3.12)$$

összefüggések szerint, ahol  $\Delta t$  a szimulációs időlépés. Ez megfelel a jól ismert Euler-féle explicit közelítő sémának, és kellően kicsinek választott  $\Delta t$  esetén a közelítés megfelelő pontosságú.

Ez a modell intuitívan úgy fogalmazható, hogy a sejtek a környezetükben lévő tápanyagot egyszerűen átalakítják.

### 3.4.2. A tápanyag felvételét feltételező megközelítés

Az előző modell nem veszi figyelembe azt, hogy a sejt az átalakításhoz az  $S$  tápanyagot először ténylegesen felveszi (vagyis  $S$  koncentrációja a sejt belsejében a felvétel során növekedni fog), valamint, hogy a sejt által felhasználható tápanyag mennyisége, illetve a  $P$  melléktermék előállításának sebessége inkább a belső koncentráció, sem mint a környezeti  $S$  koncentráció függvénye. Ezek alapján az előbbi reakcióegyenletet az alábbiak szerint pontosíthatjuk.



ahol  $S_{env}$  a környezetben jelen lévő tápanyagot,  $S_{int}$  a felvett tápanyagot,  $P$  az előállított mellékterméket jelöli. A reakciók sebességét most az  $r_u$  (*uptake*) és  $r_d$  (*decay*) ráták jellemzik. A reakcióegyenlethez tartozó differenciálegyenlet-rendszer

$$\frac{d [S_{env}]}{dt} = -r_u [S_{env}] \quad (3.14)$$

$$\frac{d [S_{int}]}{dt} = r_u [S_{env}] - r_d [S_{int}] \quad (3.15)$$

$$\frac{d [P]}{dt} = r_d [S_{int}]. \quad (3.16)$$

A differenciálegyenletek közelítése itt is a korábban leírt módon lehetséges.

A modell tehát a sejt belsejét egyfélle „pufferként” értelmezi, amelyet valamekkora sebességgel megtölt a sejt a környezetében elérhető tápanyag mennyiségének függvényében, majd felhasználja ezt a pufferelt tápanyagot egy adott sebességgel.

### 3.4.3. Szabályozott felvétel negatív visszacsatolással

Lehetséges az előző modellen úgy finomítani, hogy a sejt nem kezdi azonnal a környezetében elérhető tápanyagot felvenni, csak akkor, ha a belső tápanyagmennyiség lecsökkent.



ahol a felvétel leíró reakció sebessége függ a sejt belsejében lévő  $S_{int}$  koncentrációtól. Egy lehetséges ilyen függvény az

$$r_u = \hat{r}_u(1 - [S_{int}]),$$

amely alacsony  $S_{int}$  koncentráció esetén gyors, míg magasabb (1-hez közel)  $S_{int}$  koncentráció esetén lassú reakciót jelent majd, így egyfélle negatív visszacsatolást megvalósítva a rendszerben. A reakcióegyenlethez tartozó differenciálegyenlet-rendszer ekkor

$$\frac{d [S_{env}]}{dt} = -r_u (1 - [S_{int}]) [S_{env}] \quad (3.18)$$

$$\frac{d [S_{int}]}{dt} = r_u (1 - [S_{int}]) [S_{env}] - r_d [S_{int}] \quad (3.19)$$

$$\frac{d [P]}{dt} = r_d [S_{int}], \quad (3.20)$$

amelynek megoldása ebben az esetben előállítható is a korábban leírtak szerint.

# 4. fejezet

## Szoftveres megvalósítás

### 4.1. Az elkészített programok

A rendszer három komponensből áll. A legfontosabb maga az elkészített modell, azonban szükség volt további két komponens létrehozására a modell validációja és az eredmények megjelenítése érdekében.

Az egyik egy optimalizáló program, aminek célja, hogy a szimuláció többszörös meghívásával megtalálja a legmegfelelőbb bemeneti paramétereket, amelyek a kísérletekből származó eredményekhez a lehető legközelebbi eredményeket adják. A másik egy egyszerűbb diagramkészítő program, aminek segítségével vizuálisan megjeleníthető az ágensek populációja. A 4.1 ábra mutatja a kapcsolatot ezek között a komponensek között.

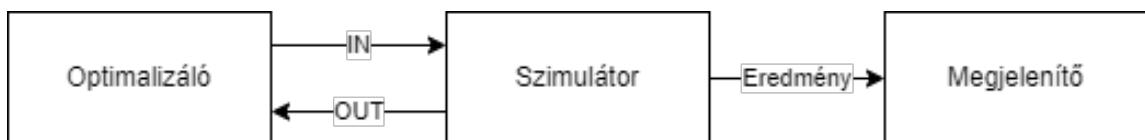
#### 4.1.1. Az optimalizáló és a megjelenítő komponens

##### Optimalizáló

Az ágens alapú modell illesztéséhez szükségem volt egy optimalizáló komponensre, amelyben heurisztikus módszerekkel igyekeztem megtalálni a modell megfelelő bemeneti paramétereit. A két általam használt heurisztika a Nelder–Mead eljárás és a Genetikus algoritmus.

A Genetikus algoritmus egy globális keresési módszer, a Nelder–Mead pedig a lokális optimum megtalálására alkalmas. Az optimalizáló komponens elsősorban egy genetikus algoritmussal igyekszik megtalálni a globális optimumot, majd ha ez megvan, akkor a Nelder–Mead eljárással javít az eredményen. Az 5.2. fejezetben részletesen beszámolok a modellillesztés menetéről.

Az optimalizáló komponens meghívja a szimulátor komponenst és átad neki egy



**4.1. ábra.** A rendszert alkotó komponensek kapcsolata. Balra látható az optimalizáló, középen maga a szimulációs modell, jobbra pedig a program, ami képes grafikonon megjeleníteni az ágenseket. Az optimalizáló és az szimulátor között látható IN fájl tartalmazza az éppen legjobbnak választott bemeneti paramétereket, az OUT fájl pedig a kapott eredményt, ami alapján az optimalizáló tovább finomít.

IN fájt, ami a modell kezdeti paramétereit tartalmazza. Miután lefutott egy szimuláció a szimulátor komponens készít egy OUT fájlt, amit az optimalizáló beolvas, meghatározza az eredmény jóságát és ha szükséges új megoldást keres. A komponens implementációját Python nyelven végeztem.

## Megjelenítő

A megjelenítő komponens feladata, hogy a szimulátor komponens által generált pillanatképeket beolvassa és egy megfelelően felépített pontdiagramon ábrázolja az ágenseket egy Descartes-féle koordináta-rendszerben. A diagramon a pontok reprezentálják a sejteket a háttér színe pedig a tápanyag koncentrációját. A komponenst az optimalizálóhoz hasonlóan Python nyelven implementáltam.

### 4.1.2. A szimulátor komponens

A modell elméleti felépítését a 3. fejezetben tárgyaltam, most pedig bemutatnám a programszintű megvalósítást. A szoftvert C++ nyelven valósítottam meg és CUDA keretrendszerrel gyorsítottam.

Működés tekintetében a szoftver először beolvas egy IN fájlt, ami tartalmazza a szükséges kezdeti paramétereket. Ezek különböző anyagcsereráták, növekedési ráta és kezdeti konfluencia szint (ezeket a kísérletekből származó eredményekből nyertük), továbbá a kezdeti tápanyag mértéke.

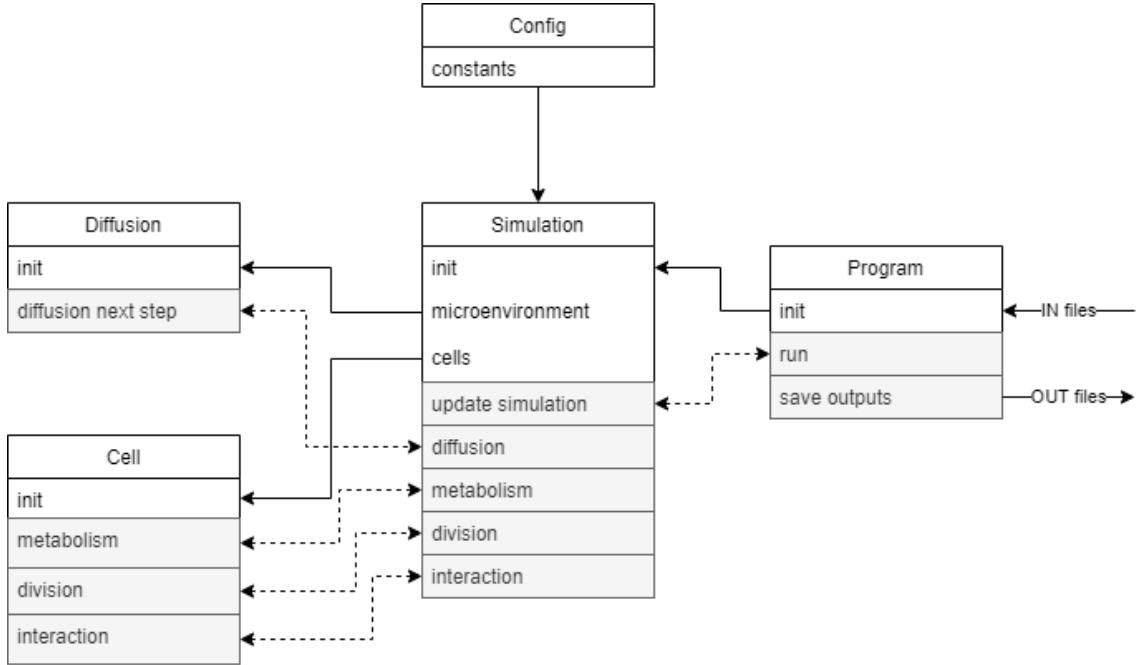
Ezt követően létrejön egy szimulációs objektum, ami létrehozza a meghatározott méretű teret, továbbá a megfelelő mennyiségű kezdeti sejyet. Ezt követően egy cikluson belül megfelelő időközönként frissítésre kerülnek a modell megfelelő komponensei és ennek eredményeként lesz megfigyelhető a növekedési dinamika, az ágensek sejtszerű viselkedése, továbbá a tápanyagszint változása a térben. A szimuláció komponenseit a 4.2 ábra szemlélteti.

A szimuláció során egy ciklus iteráció a valóságban  $1ms$ -nak felel meg, továbbá a tér és az ágensek kiterjedése esetén 1 egység az  $1\mu m$ -nek felel meg. A modellek frissítésének meghívása különböző időközönként történik. Az elkövetkezőkben ismertetett időértékek a szimulációban értelmezettek.

A feldolgozott szakirodalmak alapján az anyagtranszportnak  $15ms$ -onként, a sejtek anyagcseréjének és az osztódás próbájának pedig  $150ms$ -onként célszerű lefutnia.

5 percentként végzem el a sejtek mozgásának próbáját, amihez hozzátarozik az interakciós függvény meghívása is, ugyanis az alapján képesek a sejtek elrendeződni. Szimulációk futtatása során részben tapasztalati úton, részben a Macklin modell alapján határoztam meg, hogy  $15s$ -onként szükséges lefuttatni minden sejtre az interakciós függvényt, ugyanis az esetleges osztódásokkor szükséges a populáció rendezése. Ez annál fontosabb, minél több sejt van a rendszerben és miniél többen hatnak egymásra.

Végül pedig 30 percentként fájlba írom az anyagtranszport mátrix állapotát, és a sejtek pozíóját. Ezek a megjelenítéshez szükségesek. Továbbá elmentem az aktuális konfluencia szintet. Azért 30 percentként írom ezeket ki, mert a valós biológiai kísérletben is 30 percentként végeztek felvételket a sejtekkel. Ezt bővebben a 5. fejezetben ismertetem.



**4.2. ábra.** A modell programszintű felépítése. A Simulation komponens foglalja magába a mikro-környezet inicializálását és a megfelelő részmodellek frissítésének lehetőségét. A Diffusion komponens az anyagtranszport funkciókért felelős, a Cell pedig a sejtek tulajdonságait foglalja magába. A Config komponens általános konstansokat tartalmaz. A Program feladata, hogy inicializálja a szimulációt, a megfelelő időlépésekkel a megfelelő frissítést hajtsa végre, végül pedig az eredmények fájlba írásáért felel.

### 4.1.3. Implementáció

A modell implementációját C++ nyelven valósítottam meg. Két fontos tényező alapján esett a választásom erre a nyelvre. Az egyik, hogy megfelelő implementációval a C++ kódok a leggyorsabbak közé tartoznak és nekem nagy hangsúlyt kellett fektetnem a futásidőre. A másik tényező a CUDA keretrendszer támogatása ami ugyancsak a futásidő gyorsításában játszik fontos szerepet.

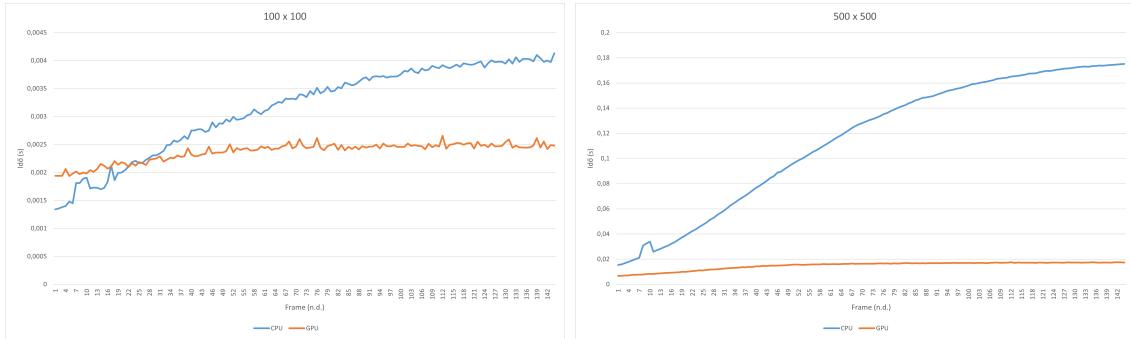
A szimuláció során használt videókártya egy Nvidia GeForce GTX 1650 Super. A GPU Turing architektúrás, 4GB GDDR6-os memóriával rendelkezik, továbbá a memória sebessége 12Gbps.

Az Optimalizáló komponenst és a Megjelenítő komponenst Python nyelven valósítottam meg. Az Optimalizáló komponensnél a fő indok a Python széleskörű matematikai és algoritmikus támogatottsága, a megjelenítésnél pedig az adatvizualizációs előnyeit igyekeztem kihasználni. Mindkét esetben fontos tényező volt a nyelv egyszerű és logikus felépítése.

A kísérletekből származó eredmények feldolgozásához a CellProfiler nevű programot használtam. Ez a program képes a kapott time-lapse felvételeken azonosítani a sejteket és két képkocka között lévő eltérések alapján automatikusan szolgáltatja a kívánt mérési eredményeket. A program használatát a TTK kutatói javasolták.

### 4.1.4. Gyorsítás GPU-val

A szimuláció futtatása során rendkívül sok számítás történik, ezért célszerű optimalizálni a modellt. A kiválasztott részmodellek közül mindegyik elképzélés viszonylag



**4.3. ábra.** Két teljes szimuláció futás idejének összehasonlítása. Balra  $100 \times 100$ -as térben, jobbra pedig  $500 \times 500$ -as esetben. A függőleges tengelyen az időt figyelhetjük meg  $s$ -ban a vízsíntes tengelyen pedig az adott pillanatképet. Ebből 144 készül minden szimulációkor, ez 72 órát tesz ki egy valós biológiai kísérlet esetében. A kék színű grafikon jelenti a szekvenciális a narancssárga pedig a párhuzamosított futásidőt. Jól látható, hogy szimplán az interakciós függvény párhuzamosításával rendkívüli mértékben javult a modell végrehajtási ideje. A bal oldali esetben kb. 100 sejtet tud kezelni a rendszer, a jobb oldali esetben pedig közel  $3000db$ -ot. Megfigyelhetjük, hogy alacsony sejtszám esetén a párhuzamos megvalósítás kezdetben még hátrányos is a szekvenciálishoz képest, azonban ahogy nő a sejtek darabszáma úgy válik egyre inkább elengedhetetlenné az optimalizálás.

jól párhuzamosítható, én ezek közül az interakciós modellt emeltem ki és annak a gyorsítását implementáltam GPU-n a C++-os CUDA könyvtár segítségével.

Az interakciós modell úgy működik, hogy alapvetően minden sejt minden sejtet megvizsgál és módosít a saját elmozdulásvektorán a többi sejt rá ható vonzó vagy taszító ereje alapján. A számítások csökkentése érdekében meghatároztam egy távolságot, amin belül kell legyen egy sejt, hogy interakciós vizsgálat alá helyezzem, viszont még így is hamar rendkívül magasra szökhet a futásidő a populáció növekedésével.

A gyorsítás úgy történik, hogy minden szál egy sejtet kezel a populációból, ezt a sejtet összeméri a többi sejttel, majd egy új tömbbe írja át a frissített eredményt. Ezzel a megoldással csak a szálak beváráására kell ügyelni, konkurens műveletek nem alakulhatnak ki. Végezetül pedig az eredménytömb kerül visszaírásra a szimuláció populációt tartalmazó adatszerkezetébe.

A 4.3 ábrán szemléltetem egy teljes szimuláció futásának eredményét az interakciós függvény szekvenciális és párhuzamosított implementációja esetén.

## 5. fejezet

# Paraméterillesztés és szimulációs eredmények

### 5.1. Kísérleti adatok

#### 5.1.1. A kísérleti elrendezés

A biológiai kísérletben a TTK kutatói az A549-es sejtvonallal (tüdőrák sejt) dolgoztak. Különböző tápanyagok különböző koncentrációjába helyezték ki a sejteket egy 96-well plate-re. Az elsőleges cél mikroszkopikus felvételek készítése volt, hogy megfigyelhető legyen a populáció növekedése a különböző tápanyagkezelések függvényében.

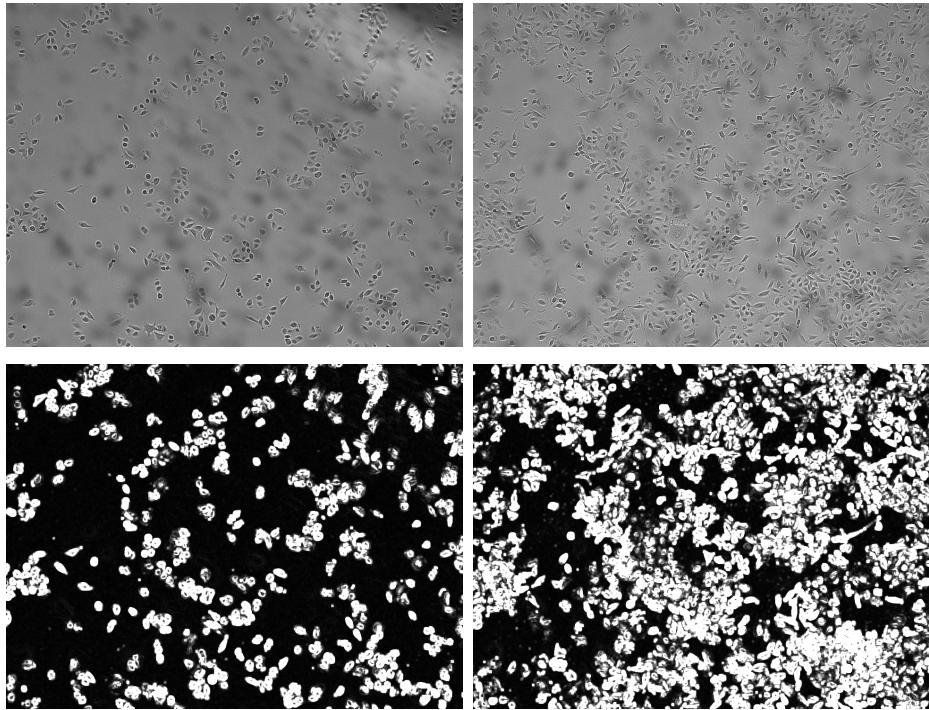
A sejtek tápanyagkezelése egy FBS (Fetal Bovine Serum) nevű anyaggal történt, amely a sejtek fő táplálékának tekinthető. A plate-en két sorban futott kísérlet az esetleges hibák kiküszöbölése érdekében. Mindkét sorban az FBS koncentrációja 100%-ról indulott és a 9. well-ig feleződött ez az érték. A 10. well tartalmazta a kontroll populációt, mely az ideális fejlődési mintázatokat és eredményeket produkált. Itt a sejtek az úgynevezett komplettált RPMI kezelést kapták. Ez 10% FBS-t, 2.054mM L-glutamint és 1% Penicilin/Streptomycin oldatot tartalmazott. Ez utóbbi egy antibiotikum keverék, amely az esetleges fertőzések megakadályozásáért felelős. Általános esetben a sejtek tenyészítése a komplettált médiumban történik.

A kísérletet az FBS-hez hasonlóan elvégezték Glükózzal és Glutaminnal is, azonban sajnos ezek a kezelések a populáció meglehetősen gyors pusztulásával jártak, ezért nem lehetett a szimuláció szempontjából hasznos eredményt kinyerni belőlük.

#### 5.1.2. Live cell imaging felvételek kiértékelése

A kísérletet egy Live cell imagerrel követték nyomon. Ez minden well-ről fél óránként készített egy mikroszkopikus felvételt. Ezeket a képeket összefűzve egy time-lapse videót kapunk a sejtpopulációról. Megfigyelhetjük a videókon a sejtek növekedési dinamikáját, mozgásuk sebességét és irányát, morfológiájuk változását, továbbá a fenotípus váltásokat.

A teljes kísérlet időtartama 72 óra volt. Az eredményként kapott képeket a Cell-Profiler nevű programmal analizáltam [18]. Ezt megelőzően a képeket pixelenként lokális variancia alapján színeztem, ugyanis az imager képein nem minden esetben volt jól elhatárolható a sejt a háttéről. A variancia alapján színezett képeken a



**5.1. ábra.** Live cell imagerrel készült mikroszkópos képek a kísérlet kezdetén és 24 óra múlva (felső sor). Lokális variancia alapján színezett képek: a világosabb pixelek környezetében textúrázottabb a kép, ezek sejtekhez tartoznak (alsó sor).

sejtek (magas intenzitású helyek) már jól elváltak a homogén háttéről (alacsony intenzitású helyek), így egy előre megadott küszöbértékhez viszonyítva meghatároztam az előtér (sejtek) és háttér arányát, amely a plate konfluencia (telítettség) értékével korrelál. A 5.1 ábrán látható két mikroszkópikus felvétel és ezek színezése a variancia alapján.

A konfluencia értékeket felhasználva a CellProfilerben készített feldolgozó eljárás automatikusan meghatározta a populációk növekedését leíró görbüket. Az eredményeket a 5.2 ábrán láthatjuk.

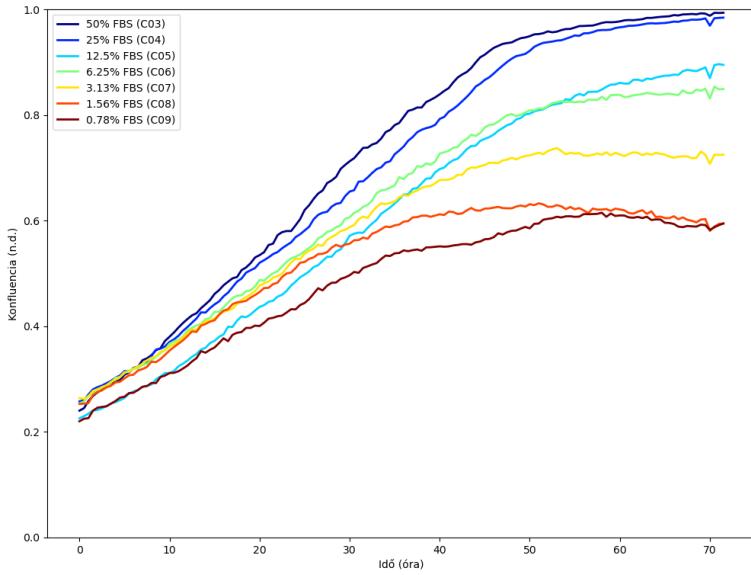
## 5.2. A modell paramétereinek becslése

A kísérletekből kapott növekedési görbük megfelelő alapot szolgáltatnak a modell paramétereinek becslésére, mivel az ágensek időbeli alakulása a szimuláció során könnyen összefüggésbe hozható a fotókon mért konfluenciával.

Előzetes szimulációs futtatások alapján megállapítottam, hogy a modellben szereplő számos paraméter közül néhány hatása a populáció egyedszámára elhanyagolható. Ezek a paraméterek a sejtek különféle átmérői, mozgékonyságuk, az interakciós potenciálfüggvény konkrét paraméterei, valamint a szimulációs térrész alakja voltak. A paraméterbecslés során épp ezért ezeket az értékeket fixen hagytam, részben irodalmi forrásokból beszerzett, részben tapasztalati úton gyűjtött értékekre.

### 5.2.1. A paraméterbecslés menete

Az ágensalapú modell paraméterbecsléséhez a modell szimulációja során szabályos időközönként elmenettetem a rendszerben lévő ágensek darabszámát, amelyet a szi-



**5.2. ábra.** A kísérletekből kapott timelapse felvételek elemzésével előállított reprezentatív növekedési görbék. Egy-egy görbe egy adott kezdeti tápanyagmennyiség melletti populáció sűrűséget (konfluenciát) ábrázol. A biológiai variabilitást a kísérlet ismétlésével csökkentettük (az ábrán nem látható).

mulált térrész méretei alapján összefüggésbe hozhatunk a konfluenciával az alábbi módon.

Egy  $w$  széles és  $h$  magas térrészben akkor maximális a gömbszerű ágensek sűrűsége, ha középpontjaik hatszögrácsban rendeződnek el. Ekkor a  $wh$  területű térrésznek  $\frac{\pi\sqrt{3}}{6}$ -od részét (nagyjából 91%-át) fedik le. A lefedett terület arányából tehát következtethetünk az elméleti maximális számú ágensre az alábbi képlettel.

$$n = \frac{\sqrt{3}wh}{6r^2}, \quad (5.1)$$

ahol  $r$  a sejtek belső részének („magjának”) a sugara.

A szimuláció során három választott tápanyag mennyiség mellett kapott növekedési görbét becsültem. Ehhez olyan paraméterkombinációt kellett találnom, amely választása esetén a különböző tápanyagszinttel induló szimulációk a lehető legjobban közelítik a mérési adatokat. A probléma ebben a formában elsőre talán nem tűnik bonyolultnak, a gyakorlati megvalósítása ugyanakkor egyáltalán nem nyilvánvaló.

### 5.2.2. Helyettesítő modell használata

Egy konkrét szimuláció futtatása nagyságrendileg néhány percre tehető, amennyiben a kísérletek során választott 72 órás időtartamot kívánom megismételni. A modell összes paraméterei közül elhagyva a korábban említetteket, továbbra is öt, egymástól független értéket kell úgy megválasztanom, hogy azok értéktartományáról nem sok előzetes információ áll rendelkezésre. Néhány paraméterkombináció kipróbálása szinte biztosan nem elegendő, a szisztematikus (rács alapú) keresési módszerek pedig a gyakorlatban teljesen kivitelezhetetlen hosszúságú folyamatot jelentenének, hiszen több tízezer szimulációt kellene lefuttatnom.

Az ágens alapú modell bizonyos feltételek esetén (pl. homogén elrendeződésű ágensek) felfogható egy diszkrét állapotokat kezelő populációdinamikai modellként is, vagyis az ágensek tényleges száma közelíthető a klasszikus populációdinamikai módszerekkel. A modell paraméterillesztéséhez így egy helyettesítő (*surrogate*) modellt, más néven *metamodelt* használtam, amely tizedmásodpercek alatt szimulálható, de mégis ad valamilyen támpontot az ágensalapú modell várható viselkedéséről különféle kezdeti feltételek mellett.

### Differenciálegyenlet-alapú helyettesítő modell

Az alkalmazott helyettesítő modell a populáció sűrűségét, a populációra jellemző növekedési rátát, a rendelkezésre álló tápanyag mennyiségét, valamint a populáció által felvett (elfogyasztott) táplálék mennyiségét, illetve a folyamatokhoz kapcsolódó rátákat használja.

$$\frac{dC}{dt} = r_g \cdot C \cdot (1 - C) \cdot \frac{1}{1 + \exp(-\sigma([S_{int}] - \mu))} \quad (5.2)$$

$$\frac{d[S_{env}]}{dt} = -r_u \cdot C \cdot (1 - [S_{int}]) [S_{env}] \quad (5.3)$$

$$\frac{d[S_{int}]}{dt} = r_u \cdot C \cdot (1 - [S_{int}]) [S_{env}] - r_d \cdot C \cdot [S_{int}], \quad (5.4)$$

ahol  $C$  a populáció konfluenciáját (0-1 közötti érték),  $r_g$  pedig a maximális növekedési rátáját adja meg. A további jelölések egyeznek a (3.17)-ben megadottakkal. A fenti modell numerikus közelítését Runge–Kutta-típusú megoldóval,  $\Delta t = 1$  min lépésközzel végeztük, stabilitási problémákat nem tapasztaltunk.

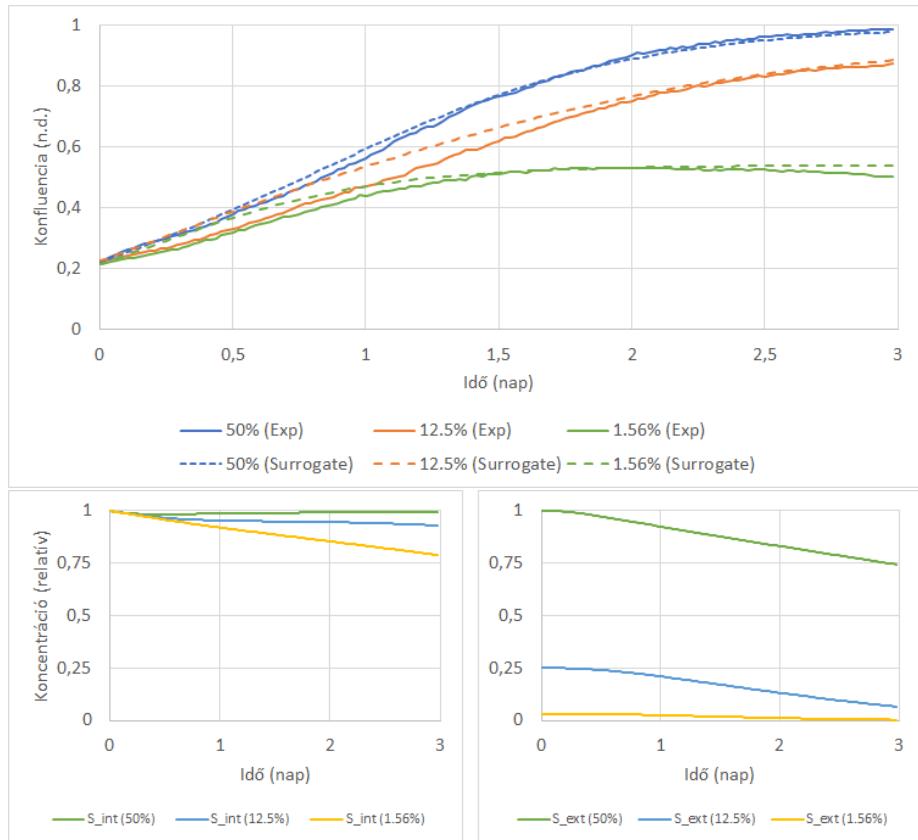
### Modellillesztés

A fenti egyenletekkel definiált helyettesítő modell numerikus megoldásának a kísérletes adatuktól való eltérését a megoldás átlagos négyzetes hibájával (MSE) jellemzük. A modell illesztéséhez az alábbi módszereket alkalmaztam.

**Nelder–Mead heurisztika előzetes jelöltválasztással.** A keresési térben latin hipernégyzetek mintázó eljárással (LHS) adott számú pontot (paraméterkombinációt) szórtunk el. Ezek közül a legjobbnak bizonyulót (minimális hibafüggvény értékűt) használtuk a Nelder–Mead kereső eljárás kezdőértékeként, majd tipikusan 50-100 lépést hajtottunk végre. Az eljárás gyakran gyenge közelítést tudott csak adni, mivel feltehetően lokális minimumban ragadt.

**Genetikus algoritmus alapú minimalizálás.** A keresési térben LHS-sel előállított populációból indultunk ki. A mutációs rátát magasra, a keresztezési rátát alacsonyra állítva hozzávetőlegesen 20-50 generáción keresztül futtattuk a keresést. Az eljárás lényegesen jobb közelítést talált a fenti módszereknél, azonban lényegesen több függvényhívásra volt ehhez szükség.

A genetikus algoritmus által talált paraméterkombinációt részben manuálisan, részben a Nelder–Mead-eljárással tovább finomítva a 5.3. ábrán látható közelítést kaptuk. A tápanyag koncentrációkat elemezve megállapítható, hogy a modell nagyobb kezdeti tápanyag mennyisége mellett lényegében konstans belső mennyiséget prediktál, vagyis a populáció növekedése alig lassul a tápanyaghiány miatt. Ez megmagyarázza az 50% és 12.5% FBS mellett futtatott kísérletek esetén látható magas



**5.3. ábra.** Felső sor: közelítő modellel kapott eredmények a kísérletes növekedési görbékre. Alsó sor: a populáció által felvett (balra), valamint a környezetben elérhető tápanyag koncentrációja (jobbra).

konfluenciát (az ábrán kék és narancsszínű görbék). Az alacsony kezdeti tápanyagszint mellett azonban a belső tartalék hamar felhasználódik, a környezetből alig van csak lehetőség felvenni tápanyagot, így a növekedési ráta gyors csökkenésnek indul (zöld görbe az ábrán).

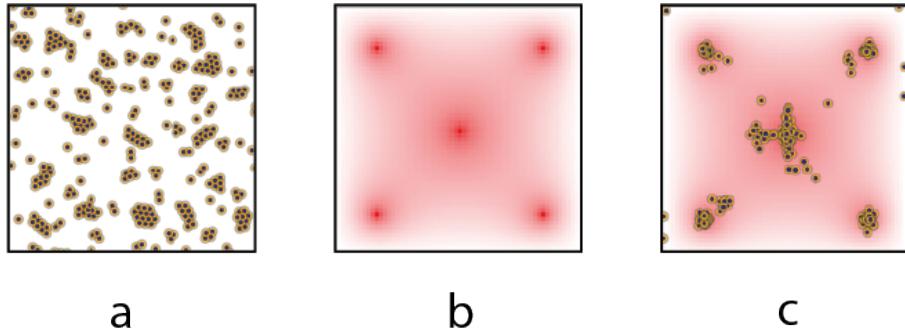
### 5.2.3. Finomítás az ágensalapú modellre

A helyettesítő modell bár sokkal egyszerűbb, épp ezért relatív gyorsan illeszthető, ugyanakkor homogén populációt feltételez, ami viszont nem minden tekintetben igaz az ágensalapú modellre. A helyettesítő modell illesztésével kapott paraméterértékek használatával azonban jó induló értékeket adhatunk egy kereső eljárás számára, ami már relatíve sokkal kevesebb számú függvényhívással tudja illeszteni az ágensalapú modellt a kísérletes adatokra.

Az ágensalapú modell illesztéséhez a fent ismertetett Nelder–Mead-eljárást futtattam a helyettesítő modellből kapott paraméterekeket használva kezdeti keresővektorként.

## 5.3. Néhány szimulációs eredmény

A modell implementációját a különböző részmodellek kódszintű megvalósításával és tesztelésével kezdtem.



**5.4. ábra.** Az implementált részmodellek tesztelésének eredményei. Az *a* jelű képen látható a molekuláris dinamikai részmodell eredménye. A körök reprezentálják a sejteket és látható, hogy a szimuláció végeztével kisebb telepek alakultak ki. A *b* jelű képen a diffúzió látható. Itt a piros szín erőssége jelzi a koncentráció mértékét. Minél sötétebb annál nagyobb a koncentrációja az anyagnak, a fehér gyakorlatilag 0-t jelent. A *c* ábrán pedig ezen két modell összeépítésének eredménye látható, továbbá megfigyelhető a kemotaxis jelensége. Láthatjuk, hogy nem csak véletlenszerűen alakultak ki a telepek, hanem kifejezetten a nagyobb koncentrációk környékére igyekeztek csoportosulni az ágensek.

Első feladatomnak a molekuláris dinamikai modell megvalósítását tűztem ki annak érdekében, hogy a sejtek közötti interakciókat kezelní lehessen a térben, továbbá a sejtek mozgásának mértékét meghatározhassam.

Ezután következett az anyagáramlás implementációja melyhez az FTCS numerikus sémát választottam. Meghatároztam azt a  $\Delta t$  időértéket két diffúziós lépés között és  $\Delta x$  méretet két rácspont között, aminél még stabilan működik a rendszer és a valóság szempontjából is releváns anyagáramlás figyelhető meg.

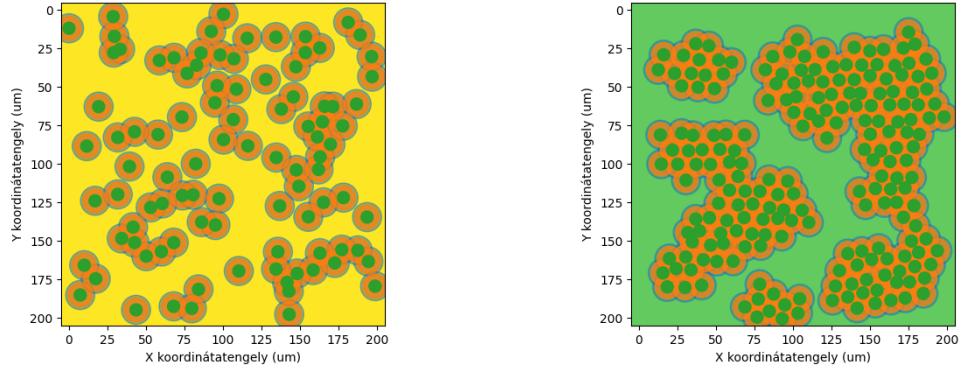
Miután sikerült stabil működést elérni minden az ágensek közötti interakciók és mozgásuk tekintetében, minden pedig az anyagtranszport megfelelő működésében, összeépítettem ezeket és a sejteket felruháztam azzal a képességgel, hogy a térben lévő gradienskülönbségeket érzékelni tudják. Ennek megfelelően képesek reprezentálni a kemotaxis jelenségét, ami a térben lévő kémiai anyag koncentrációtól függő elmozdulást jelent. Jelen esetben a sejt a számára kedvezőbb, nagyobb koncentráció felé halad, de megvalósítható ennek a folyamatnak az ellentéte is. A 5.4 ábrán láthatók ezen implementációk eredményei.

A tesztelések után következett a kísérlet virtuális terének kialakítása. Ebben a térben kezdetben a tápanyag gyakorlatilag homogén módon terült el, így a diffúziós tér annyiban módosult, hogy nem voltak kifejezetten forráspontok, hanem kezdetben minden pontban azonos volt a koncentráció. A szimuláció során a sejtek felhasználják a tápanyagot, a különböző rácspontokon a koncentráció folyamatosan változik. Bár ez a változás nagyon kis mértékű szükség van a diffúzióra a kiegyenlítéshez ezért fontos volt korábban nagyobb koncentrációkülönbségekkel tesztelni a működést.

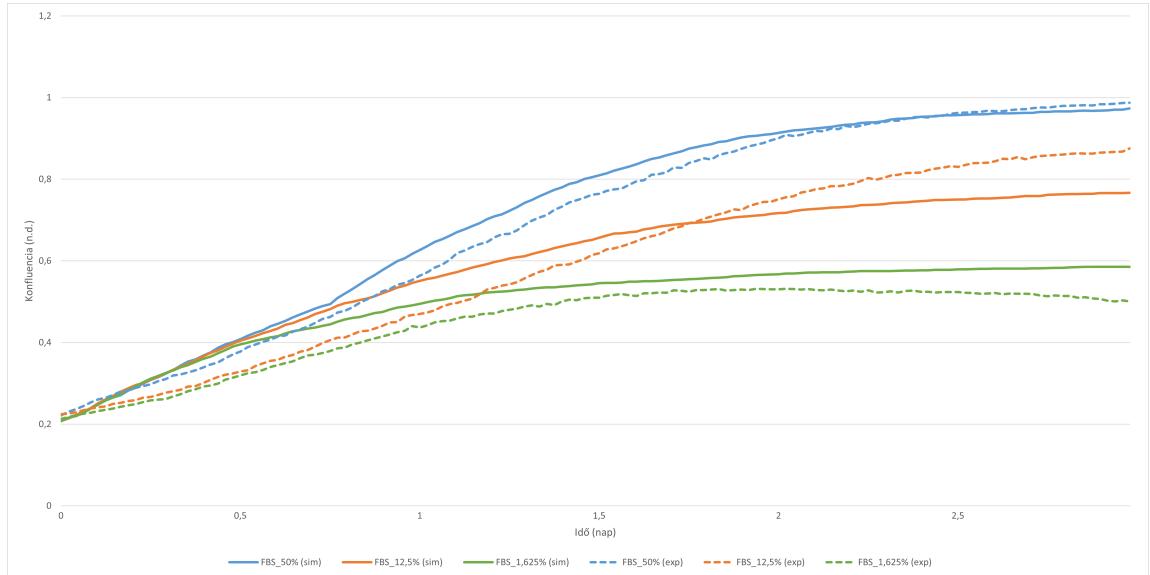
A sejtek kezdeti darabszámát a 5.1 képlet segítségével határoztam meg. A kezdeti szimulációkat 3 különböző méretű térben futtattam. Jelenleg az implementáció csak szekvenciális módon működik. Elkezdtem a párhuzamosítást, azonban egyelőre hibás működést produkál. Bár a szekvenciális szimuláció is viszonylag gyors, a sejtek osztódása által okozott számításigény növekedése miatt és a tér növelhetősége érdekében folytatni fogom a párhuzamosított verzió fejlesztését.

A szimulációs szoftver működését a futtatások alkalmával be lehetett állítani úgy, hogy látszólag megfelelő eredményt szolgáltasson. Ez volt az a pont, ahol szükség

volt a kísérletekből származó eredményekre annak érdekében, hogy felparaméterezzessem úgy a modellt, hogy biológiaileg is releváns működést mutasson. Sajnos az adatok kiértékelése körülményes volt és sokáig tartott, így a releváns működést még nem sikerült elérni. A 5.5 ábrán látható a kísérleti eredményekből felparaméterezett futtatás kimenetele.



**5.5. ábra.** Az elkészült modell grafikus eredménye. Balra a szimuláció kezdeti fázisa látható. Jobbra a szimuláció belül értelmezve egy 12 órával későbbi pillanatkép látható. Megfigyelhetjük, hogy a sejtek az idő előrehaladtával összetapadtak, osztódtak, továbbá a háttér elszíneződéséből látható, hogy a tápanyagot is elkezdték felhasználni.



**5.6. ábra.** Az elkészült modellből származó konfluenciaszintek összevetése a kísérletek eredményeivel. A grafikonon a szaggatott vonallal jelzett eredmények mutatják a kísérletekből származó konfluencia értékeitet adott kezdeti tápanyagkoncentrációt megfelelően. A folytonos vonal pedig a szimuláció eredménye ugyanezen esetekben. A modell eredménye bár nem tökéletes mégis jó közelítést ad a kísérletekből származó konfluenciaszintekhez.

A szimulációk eredményeit a 5.6 ábrán látható grafikon mutatja be. Ezen az ábrán a kísérletekből származó konfluencianövekedés mértékét mutatom be, mellette a szimuláció eredményeivel. Az eredményekből megfigyelhető, hogy a modell képes közelíteni a valós kísérletek eredményeit. Bár nem tökéletes, további finomhangolással tovább javíthatóak az eredmények.

# Összegzés, továbbfejlesztési lehetőségek

Az eddigi eredményekből kiindulva úgy vélem az ágens alapú modellezési módszer egy nagyon hasznos iránymutató eszköz lehet a sejtbiológiai kísérletek támogatására. Mint látható volt az előzőekben, nagyon sok tényezőnek kell helyesnek lennie annak érdekében, hogy megfelelő eredményt szolgáltasson egy ilyen modell.

Habár a jelenlegi állapotában az én modellem még nem közelíti tökéletesen a kísérletekből származó konfluenciaszinteket, az alap elképzélések jó része működőképes és a különböző részegységek a tesztek alapján megállják a helyüket. További finomhangolással elérhető az a megfelelő bemeneti paraméterkombináció, ami révén a modellt illesztettnek tekinthetjük.

A molekuláris dinamikai modellt megfelelőnek találtam a mechanikai interakciók kezelésére, azonban a számításigénye miatt ezt a komponenst párhuzamosítanom kellett. A GPU-s optimalizálás sikeres volt, ezáltal nagy mértékben sikerült gyorsítani a szimulációkat.

Az FTCS sémával végrehajtott diffúzió bár csak nagyon alacsony időléptékkel mutat helyes működést a megvalósítása rendkívül egyszerű és a későbbiekben egy könnyen párhuzamosítható módszer.

Az jelenlegi anyagcserefolyamatot lebonyolító egyszerűbb kinetikai modellt kézőbb szeretném egy a Dr. Matthew Egbert részletesebb modelljéhez hasonló rendszerre cserélni. Ennek oka, hogy pontosabban nyomon lehessen követni a tápanyag felvételét, átalakítását, továbbá a melléktermék termelődését és a különböző *pH* függéseket.

A továbbiakban mindenképp szeretném elérni, hogy minél relevánsabb kimeneti eredményekkel szolgáljon a modell, ezért újra tesztelem a különböző részegységeket, továbbá újra tervezem a paraméterillesztést, szükség esetén új megoldásokat keresek.

Mivel nagyon hamar gyakorlatilag exponenciális mértékben nőhet a futásidő, ezért szükségesnek látom további részegységek párhuzamosítását annak érdekében, hogy nagyobb térben, nagyobb sejtszámmal lehessen szimulációkat végezni rövidebb időn belül. A molekuláris dinamikai modell interakciókat kezelő részének párhuzamosítását már elvégeztem, ezt követően a mozgást, az anyagcsérét és az osztódást szeretném videókártyával gyorsítani. Ezen túlmenően az anyagtranszport frissítését szeretném még párhuzamosítani a későbbiekben.

Miután sikerült a teljes modellt megfelelően felparaméterezní és a kísérleteknek szempontjából is releváns eredményt szolgáltatott, szeretném továbbfejleszteni és egy kifejezetten a kemotaxis jelenségét megvalósító rendszert kialakítani. Remélyeim szerint sikerül ennek megfelelő kísérleteket is lebonyolítani a szimulációk futtatásával párhuzamosan és megalkotható egy olyan három forgatókönyves szimulációsorozat sejtek terén, amelyet Dr. Matthew Egbert vázolt fel baktériumok esetében.

# Kivonat

A számítógépes modellezés és szimuláció egyre fontosabb szerepet tölt be a biológiai kutatások, többek között a sejtbiológia területén alkalmazott kísérletek támogatásában. Mivel a valós biológia (*in vivo* vagy *in vitro*) kísérletek gyakran meglehetősen időigényesek, egy jól felépített *in silico* modell ötletet adhat a valós kísérletek megtervezéséhez, illetve a segítségével választ kaphatunk olyan kérdésekre, amely a valóságban körülönnyen megfigyelhető jellemzőkkal áll kapcsolatban.

A dolgozatom célja egy olyan ágensalapú számítógépes modell létrehozása, amelyben a sejteket reprezentáló ágensek különféle környezeti feltételek (pl. tápanyag mennyisége, tápközeg kémhatása) mellett megfigyelhető növekedési és mozgási mintázatait lehetséges vizsgálni. Az ágensek mozgását és növekedését belső szabályrendszerük irányítja, amelyre közvetetten hatással lehetnek a környezetükben található kémiai anyagok is. Az ágensek képesek ezeknek az anyagoknak a felvételére vagy termelésére, amely folyamatokat egy egyszerű anyagcsere-modell szabályozza. Az anyagtranszport szimulációjához szükséges numerikus eljárásokat masszíván párhuzamos környezetben is implementáltam.

A modell bemeneti paramétereit *in vitro* (tenyésztedényben indított) sejtes kísérletek mikroszkópos felvételeinek elemzésével, az ezekből meghatározott idősoros adatok alapján becsülttem. A kísérletek során a sejtek eltérő tápanyag-szintek mellett a növekedési kinetikáját vizsgáltuk. A kapott szimulációs eredmények alátámasztják, hogy a modell képes leírni az ágensek környezeti feltételektől függő viselkedését, valamint az anyagcsere-folyamatok hatását.

# Abstract

Computer modeling and simulation are acting an increasingly important role in supporting experiments in biological research, including cell biology. Because real-life biology (*in vivo* or *in vitro*) experiments are often quite time-consuming, a well-constructed *in silico* model can provide an idea for designing real-world experiments and provide answers to questions about features that are difficult to observe in reality.

The aim of my dissertation is to create an agent-based computer model in which it is possible to study the growth and movement patterns of cellular agents under different environmental conditions (e.g. nutrient volume, medium pH). The movement and growth of agents are controlled by their internal system of rules, which can also be indirectly affected by the modeled chemicals in their environment. Agents are able to take up or produce these substances, which processes are controlled by a simple metabolic model. I also implemented the numerical procedures required for the simulation of mass transport in a massively parallel environment.

The input parameters of the model were estimated by analyzing microscope images of *in vitro* (kept in a dish) cell experiments based on time series data determined from these. In the experiments, the growth kinetics of the population at different nutrient levels were examined. The simulation results obtained support the ability of the model to describe the behavior of agents depending on environmental conditions as well as the effect of metabolic processes.

# Köszönetnyilvánítás

Köszönöm konzulensemnek, Kiss Dánielnek a mérhetetlen türelmet, segítséget és iránymutatást a felkészüléssel, modellépítéssel és adatelemzéssel kapcsolatos új ismereteim megszerzésében.

Köszönöm továbbá a Természettudományi Kutatóközpont Gyógyszerrezisztencia Kutatócsoportjának, elsősorban Dr. Füredi Andrásnak és Gombos Balázsnak a biológiai kísérletek megtervezését és kivitelezését, valamint a lehetőséget, hogy részt vehettem a kutatásban.

# Irodalomjegyzék

- [1] Charles M. Macal és Michael J. North. “Tutorial on agent-based modeling and simulation”. *Proceedings - Winter Simulation Conference* (2005), 2–15. old. ISSN: 08917736. DOI: [10.1109/WSC.2005.1574234](https://doi.org/10.1109/WSC.2005.1574234).
- [2] D. C. Walker és tsai. “The epitheliome: Agent-based modelling of the social behaviour of cells”. *BioSystems* 76 (2004), 89–100. old. ISSN: 03032647. DOI: [10.1016/j.biosystems.2004.05.025](https://doi.org/10.1016/j.biosystems.2004.05.025).
- [3] Marian Gheorghe, Mike Holcombe és Petros Kefalas. “Computational models of collective foraging”. *BioSystems* 61 (2001), 133–141. old. ISSN: 03032647. DOI: [10.1016/S0303-2647\(01\)00164-2](https://doi.org/10.1016/S0303-2647(01)00164-2).
- [4] Nuno Fachada, V Lopes és Agostinho Rosa. “Agent-based modelling and simulation of the immune system: a review”. *EpiA 2007 Lncs (Lnai)* 4874 (2007), 300–315. old.
- [5] Shu Heng Chen és Chung Chih Liao. “Agent-based computational modeling of the stock price-volume relation”. *Information Sciences* 170 (2005), 75–100. old. ISSN: 00200255. DOI: [10.1016/j.ins.2003.03.026](https://doi.org/10.1016/j.ins.2003.03.026).
- [6] Paul Macklin, Mary E. Edgerton, Alastair M. Thompson és Vittorio Cristini. “Patient-calibrated agent-based modelling of ductal carcinoma in situ (DCIS): From microscopic measurements to macroscopic predictions of clinical progression”. *Journal of Theoretical Biology* 301 (2012), 122–140. old. ISSN: 00225193. DOI: [10.1016/j.jtbi.2012.02.002](https://doi.org/10.1016/j.jtbi.2012.02.002).
- [7] Grazziela P. Figueiredo, Tanvi V. Joshi, James M. Osborne, Helen M. Byrne és Markus R. Owen. “On-lattice agent-based simulation of populations of cells within the open-source Chaste framework”. *Interface Focus* 3 (2013). ISSN: 20428901. DOI: [10.1098/rsfs.2012.0081](https://doi.org/10.1098/rsfs.2012.0081).
- [8] D. C. Rapaport. *The Art of Molecular Dynamics Simulation*. USA: Cambridge University Press, 2004. ISBN: 0521825687.
- [9] David Landau és Kurt Binder. *A Guide to Monte Carlo Simulations in Statistical Physics*. USA: Cambridge University Press, 2005. ISBN: 0521842387.
- [10] D Drasdo, R Kree és J S Mccaskill. “Monte Carlo approach to tissue-cell populations”. 52.6 (1995).
- [11] Bartłomiej Waclaw, Ivana Bozic, Meredith E. Pittman, Ralph H. Hruban, Bert Vogelstein és Martin A. Nowak. “A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity”. *Nature* 525 (2015), 261–264. old. ISSN: 14764687. DOI: [10.1038/nature14971](https://doi.org/10.1038/nature14971).
- [12] Heather Kelly. “Parallel Monte Carlo Simulations of Vapour-Liquid Systems”. (2006).

- [13] Svetlana Volkova, Marta R. A. Matos, Matthias Mattanovich és Igor Marín de Mas. “Metabolic Modelling as a Framework for Metabolomics Data Integration and Analysis”. *Metabolites* 10.8 (2020). ISSN: 2218-1989. DOI: 10.3390/metabo10080303. URL: <https://www.mdpi.com/2218-1989/10/8/303>.
- [14] Matthew D. Egbert, Xabier E. Barandiaran és Ezequiel A. di Paolo. “A minimal model of metabolism-based chemotaxis”. *PLoS Computational Biology* 6 (2010). ISSN: 1553734X. DOI: 10.1371/journal.pcbi.1001004.
- [15] Wojciech Artichowicz és Dariusz Gasiorowski. “Computationally efficient solution of a 2D diffusive wave equation used for flood inundation problems”. *Water (Switzerland)* 11 (2019). ISSN: 20734441. DOI: 10.3390/w11102195.
- [16] Michael J. Saxton. “Modeling 2D and 3D Diffusion”. 400 (2007), 295–321. old. ISSN: 1064-3745. DOI: 10.1007/978-1-59745-519-0\_20.
- [17] Zhicheng Hu és Zhihui Liu. “Heat Conduction Simulation of 2D Moving Heat Source Problems Using a Moving Mesh Method”. *Advances in Mathematical Physics* 2020 (2020). ISSN: 16879139. DOI: 10.1155/2020/6067854.
- [18] Anne E. Carpenter és tsai. “CellProfiler: image analysis software for identifying and quantifying cell phenotypes”. *Genome Biology* 7.10 (2006. okt.), R100. ISSN: 1474-760X. DOI: 10.1186/gb-2006-7-10-r100. URL: <https://doi.org/10.1186/gb-2006-7-10-r100>.

# Ábrák jegyzéke

- 2.1. A képen a Macklin-féle modellben implementált sejtek sematikus ábrája látható. A baloldali ábra a sejt gömbökre történő felosztását mutatja be.  $R_N$  jelenti a sejtmagot reprezentáló gömb sugarát,  $R$  a sejt tényleges mérete,  $R_A$  pedig az adhéziós távolság, amin belül képes összetapadni két sejt. A piros szaggatott vonal szemlélteti egy valós sejt szabálytalan morfológiáját. Látható, hogy a kék kör ezt a méretet próbálja közelíteni. Jobbra látható, hogyan is képezték le az adhéziós tulajdonságot. Azzal, hogy az  $R_N$  távolságon belül találkoznak azt modellezik ahogy a való életben nyúlnának egymás felé bizonyos sejtek. . . . . 7
- 2.2. A Macklin modellben szereplő interakciós függvény. Látható, hogyha  $R_N$  távolságon belül kerülnek egymáshoz a sejtek, akkor az energia megugrik és taszítani fogják egymást. Azonban ahogy távolodnak az energiaszint egy idő után 0 lesz.  $R_N$  és  $R_A$  között történik meg a vonzás. Az  $s$  egy vonzás és taszítás közti egyensúlyi állapotot jelöl, ezen a ponton tekinthetünk két sejtet összetapadtnak. . . . . 8
- 2.3. Rácsalapú modell, amelyben apró kockák szimbolizálják a sejteket. Balra egyetlen sejtet látunk, középen egy pár száz elemet tartalmazó populációt, jobbra pedig milliós nagyságrenddel szerepelnek az ágensek. Alacsony számú ágens esetén nem túl szemléletes a modell, azonban a milliós nagyságrendnél már egy sokkal szébb végeredményt kapunk. Forrás: [11] . . . . . 9
- 2.4. A kemotaxis modelezési lehetőségei az anyagcsere szempontjából baktériumok esetén. A transzmembrán receptor felel a koncentrációküllönbségek detektálásáért, a flagellum biztosítja a baktérium mozgását, a szignáltranszdukciós útvonal a megfelelő ingert továbbítja a flagellumknak annak megfelelően, hogy milyen anyagok vannak a környezetben. Mint látjuk az anyagcserefüggetlen és függő esetekben a küllönbség, hogy az anyagcsere befolyásolhatja a flagellumoknak kiadott parancsot és ezáltal a mozgást. Egberték anyagcserén alapuló megoldása pedig csak az anyagcserefolyamatokat veszi számításba, a többi komponenst elhagyta a modellből. . . . . 14
- 2.5. Hárrom lehetséges leképzési módszer a diffúzió számításához. Az ábrákon a piros pont jelenti az adott pillanatban vizsgált koncentrációt, a kék pontok pedig azok a koncentrációk amik alapján kiszámítjuk a piros következő pollanatbeli értékét. Balról jobbra az egyszerű négyzetes leképzést, majd a háromszöges megoldást és végül a mozgó rácsos elképzelést láthatjuk. . . . . 14

- 3.1. A komplett szimulációs modellhez szükséges komponensek réteges elrendezése. Az  $a$  reprezentálja a diffúziós réteget itt áramlik az anyag. A képen a piros szín jelenti a nagyobb koncentrációt. Ebből a rétegből párhuzamosan több is elhelyezkedhet attól függően hányfélé anyagot kezelünk a rendszerben. A  $b$  az ágensek rétege, itt történik a sejtek mozgása, a köztük lévő interakciók és a sejtciklus által végbement változások (pl.: osztódás vagy halál). A  $c$  "réteg" lenne az anyagcsere-folyamatokért felelős. Mint látható, az ágensek kapcsolatban vannak a diffúziós réteggel, folyamatosan figyelik a koncentrációkat és ha tudnak, anyagot vesznek fel. A metabolikus rétegen pedig az ágensek által felvett anyagok feldolgozása történik és az anyagcserefolyamatok eredménye visszahat az ágens mozgására és sejtciklusára. . . . . 16
- 3.2. Az osztódás a modellben két belső változótól függ. Az egyik a felvett tápanyag mennyiségét tárolja, a másik a sejt életkorát. Az  $a$  ábra mutatja a sejt általános állapotát. Amennyiben elérte egy megfelelő életkort, rendelkezik a szükséges tápanyaggal, továbbá a környezetében fellelhető elég szabad terület az osztódáshoz, akkor a  $b$  ábrán látható módon ezt véghez is viszi. A  $c$  ábra reprezentálja, hogy a két sejt külön létezik tovább. . . . . 17
- 4.1. A rendszert alkotó komponensek kapcsolata. Balra látható az optimalizáló, középen maga a szimulációs modell, jobbra pedig a program, ami képes grafikonon megjeleníteni az ágenseket. Az optimalizáló és az szimulátor között látható IN fájl tartalmazza az éppen legjobbnak választott bemeneti paramétereit, az OUT fájl pedig a kapott eredményt, ami alapján az optimalizáló tovább finomít. . . . . 20
- 4.2. A modell programszintű felépítése. A Simulation komponens foglalja magába a mikrokörnyezet inicializálását és a megfelelő részmodellek frissítésének lehetőségét. A Diffusion komponens az anyagtranszport funkciókért felelős, a Cell pedig a sejtek tulajdonságait foglalja magába. A Config komponens általános konstansokat tartalmaz. A Program feladata, hogy inicializálja a szimulációt, a megfelelő időlépések körében a megfelelő frissítést hajtsa végre, végül pedig az eredmények fájlba írásáért felel. . . . . 22
- 4.3. Két teljes szimuláció futás idejének összehasonlítása. Balra 100 x 100-as térszínen, jobbra pedig 500 x 500-as esetben. A függőleges tengelyen az időt figyelhetjük meg  $s$ -ban a vízsíntől a levegőig. Ebből 144 készül minden szimulációkor, ez 72 órát tesz ki egy valós biológiai kísérlet esetében. A kék színű grafikon jelenti a szekvenciális a narancssárga pedig a párhuzamosított futásidőt. Jól látható, hogy szimplán az interakciós függvény párhuzamosításával rendkívüli mértékben javult a modell véghajtási ideje. A bal oldali esetben kb. 100 sejtet tud kezelni a rendszer, a jobb oldali esetben pedig közel 3000db-ot. Megfigyelhetjük, hogy alacsony sejtszám esetén a párhuzamos megvalósítás kezdetben még hátrányos is a szekvenciálishoz képest, azonban ahogy nő a sejtek darabszáma úgy válik egyre inkább elengedhetetlené az optimalizálás. . . . . 23

5.1. Live cell imagerrel készült mikroszkópos képek a kísérlet kezdetén és 24 óra múlva (felső sor). Lokális variancia alapján színezett képek: a világosabb pixelek környezetében textúrázottabb a kép, ezek sejtekhez tartoznak (alsó sor). . . . .	25
5.2. A kísérletekből kapott timelapse felvételek elemzésével előállított reprezentatív növekedési görbék. Egy-egy görbe egy adott kezdeti tápanyagmennyiség melletti populáció sűrűséget (konfluenciát) ábrázol. A biológiai variabilitást a kísérlet ismétlésével csökkentettük (az ábrán nem látható). . . . .	26
5.3. Felső sor: közelítő modellel kapott eredmények a kísérletes növekedési görbékre. Alsó sor: a populáció által felvett (balra), valamint a környezetben elérhető tápanyag koncentrációja (jobbra). . . . .	28
5.4. Az implementált részmodellek tesztelésének eredményei. Az <i>a</i> jelű képen látható a molekuláris dinamikai részmodell eredménye. A köröök reprezentálják a sejteket és látható, hogy a szimuláció végeztével kisebb telepek alakultak ki. A <i>b</i> jelű képen a diffúzió látható. Itt a piros szín erőssége jelzi a koncentráció mértékét. Minél sötétebb annál nagyobb a koncentrációja az anyagnak, a fehér gyakorlatilag 0-t jelent. A <i>c</i> ábrán pedig ezen két modell összeépítésének eredménye látható, továbbá megfigyelhető a kemotaxis jelensége. Láthatjuk, hogy nem csak véletlenszerűen alakultak ki a telepek, hanem kifejezetten a nagyobb koncentrációk környékére igyekeztek csoportosulni az ágensek.	29
5.5. Az elkészült modell grafikus eredménye. Balra a szimuláció kezdeti fázisa látható. Jobbra a szimuláció belül értelmezve egy 12 órával későbbi pillanatkép látható. Megfigyelhetjük, hogy a sejtek az idő előrehaladtával összetapadtak, osztódtak, továbbá a háttér elszíneződéséből látható, hogy a tápanyagot is elkezdték felhasználni. . . . .	30
5.6. Az elkészült modellből származó konfluenciaszintek összevetése a kísérletek eredményeivel. A grafikonon a szaggatott vonallal jelzett eredmények mutatják a kísérletekből származó konfluencia értékeit adott kezdeti tápanyagkoncentrációt megfelelően. A folytonos vonal pedig a szimuláció eredménye ugyanezen esetekben. A modell eredménye bár nem tökéletes mégis jó közelítést ad a kísérletekből származó konfluenciaszintekhez. . . . .	30