



Dokumentácia k projektu

# Aplikácia celulárnych automatov v biológii

IMS - Modelování a simulace

**Autori:** Peter Hostačný (xhosta03)  
Dávid Mikuš (xmikus15)

**Dátum:** 8. decembra 2015

# Obsah

<b>1</b>	<b>Úvod</b>	<b>1</b>
1.1	Hlavné zdroje faktov . . . . .	1
1.2	Overovanie validity modelu . . . . .	1
<b>2</b>	<b>Rozbor témy, použitých metód a technológií</b>	<b>1</b>
2.1	Popis použitých postupov . . . . .	2
<b>3</b>	<b>Koncepcia</b>	<b>2</b>
3.1	Stavový automat . . . . .	3
<b>4</b>	<b>Architektúra simulačného modelu</b>	<b>4</b>
4.1	Mapovanie konceptuálneho modelu do simulačného . . . . .	4
4.2	Použitie simulačného programu . . . . .	5
<b>5</b>	<b>Podstata simulačných experimentov a ich priebeh</b>	<b>5</b>
5.1	Postup experimentovania . . . . .	5
5.2	Dokumentácia jednotlivých experimentov . . . . .	6
5.2.1	Experiment č.1 - nezávislosť na veľkosti matice . . . . .	6
5.2.2	Experiment č.2 - Zväčšovanie pomeru R:S zväčšuje infekciu . . . . .	7
5.2.3	Experiment č.3 - Zhlukovanie infikovaných buniek . . . . .	7
5.3	Závery experimentov . . . . .	8
<b>6</b>	<b>Zhrnutie simulačných experimentov a záver</b>	<b>8</b>

# 1 Úvod

Táto práca experimentálne overuje niektoré závery odborného článku [1] o celulárnom automate pre simuláciu šírenia infekčnej choroby - hepatitídy typu B.

Hepatitída typu B je infekčná choroba spôsobená vírusom hepatitídy (HBV). Ide o zápalové ochorenie pečene, ktorým trpia asi 2 miliardy ľudí po celom svete a zomiera na ňu približne 2 milióny ľudí ročne.

Existujúce modely pre simuláciu šírenia tohto vírusu sú založené na predpoklade, že populácia vírusov a buniek je rovnomerne rozložená. Reálne ale prostredie, v ktorom sa šíri nákaza, nie je homogénne a teda ani vírus sa nešíri rovnomerne. Príkladom nehomogenity sú napr. mŕtve bunky, ktoré "zabraňujú" šíreniu nákazy v priestore (keďže vírus sa môže reprodukovať iba vo vnútri živých buniek). K simulácii šírenia vírusu v takomto prostredí je teda potrebné vziať do úvahy tieto fakty a v modeli uvažovať rôzne typy buniek.

Autori článku [1] toto implementovali pomocou jednoduchého dvojdimenzionálneho celulárneho automatu. Ich experimenty sa zaoberali vplyvom rôznych parametrov systému na proces šírenia infekcie a ich nastavovaní na také hodnoty, aby model čo najlepšie vystihoval realitu. Naša implementácia tohto modelu bude opísaná v kapitolách 3 a 4. Počas experimentovania sa budeme snažiť potvrdiť/vyvrátiť ich tvrdenia a vysvetliť naše výsledky experimentovania.

## 1.1 Hlavné zdroje faktov

Na tejto práci sa podieľali obaja autori uvedený na úvodnej strane dokumentácie. Hlavný zdroj faktov bol článok [1], ktorého závery simulácie sme overovali. Jednotlivé hodnoty boli prevzaté aj z iných zdrojov, na ktoré sa odkazuje aj spomínaný článok. Tie dôležitejšie sú taktiež odkazované na konci tejto práce, v sekcii *Literatúra*.

## 1.2 Overovanie validity modelu

Validita nášho modelu bola overovaná experimentálne, porovnávaním výstupov zo simulačného modelu [8, str. 44] s informáciami vo vedeckom článku [1]. Parametre boli nastavované na hodnoty, ktoré sú použité v článku a boli skúmané výstupy našej simulácie. Celé overovanie validity by sa dalo zhrnúť do dvoch častí - overovanie, či sa nám podarilo správne namodelovať pravidlá [8, str. 317] zo spomínaného článku a experimentovanie za účelom kontroly výsledkov článku. Niektoré tvrdenia autora sme dokázali potvrdiť, avšak vyskytlo sa aj viacero tvrdení, ktoré sa nám potvrdiť nepodarilo (viac v kapitole 5).

Najprv sme teda overovali každé pravidlo celulárneho automatu.

príklad: pri pravidle pre zachovanie pomeru medzi rezistentnými a náchylnými bunkami, je treba viesť štatistiku počtu buniek oboch druhov a za behu sa udržiavať ich pomer na zadanej hodnote - toto, ako aj ostatné fakty/pravidlá, podľa ktorých sme modelovali náš systém sa nám podarili úspešne implementovať aj overiť (ostatné sú popísané v kapitole 4.1).

Následne sme pomocou experimentov porovnávali naše výsledky a výsledky článku (kapitola 5)

# 2 Rozbor témy, použitých metód a technológií

Hepatocyty sú bunky, ktoré tvoria základ tkaniva pečene. Majú na starosti väčšinu metabolických pochodov, ktoré v pečeni prebiehajú. Pre tieto bunky je charakteristická tesná väzba s okolitými bun-

kami. Častice vírusu (inak nazývané aj viriony) HBV napádajú bunky pečene a replikujú sa v nich (pre svoju replikáciu používajú výhradne zdravé bunky).

Regeneračná aktivita buniek v pečeni je veľmi vysoká. Zabezpečujú to žľčové epitelové bunky - tie sú schopné rýchlej reprodukcie a môžu sa ďalej deliť [2]. Po delení vznikajú buďto nové žľčové epitelové bunky alebo hepatocyty, pričom hepatocyty nie sú schopné ďalšieho delenia. Pomer počtu buniek týchto dvoch skupín ovplyvňuje hlavne vek jedinca a v časovom rámci niekoľkých (desiatok) týždňov ostáva tento pomer približne rovnaký.

London a Blumberg vo svojom diele [3] uvádzajú, že vírus sa rýchlejšie replikuje v epitelových bunkách ako v hepatocytoch. Takisto uvádza, že hepatocyty odumierajú po infikovaní rýchlejšie.

Niektoré bunky môžu byť nakazené chybným vírusom (napr. mutant HBV, neschopný ďalšej replikácie), čo sa ale stáva vo veľmi malom množstve prípadov (asi 0.1%) [5]. Bežná životnosť zdravých buniek v pečeni je asi 6 mesiacov [6].

Odumieranie infikovaných buniek nemá na svedomí samotný vírus, ale cytotoxické T-lymfocyty [4]. Sú to bunky, ktoré sprostredkovávajú imunitnú odpoveď tak, že rozpoznávajú bunky infikované parazitom-vírusom a postupne ich likvidujú. Nie sú zabíjané okamžite po nákeze, pretože vírus mutuje a imunitná odpoveď potrebuje nejaký čas na reakciu.

Vďaka vynikajúcej schopnosti prirodzenej obnovy, sú odumreté bunky veľmi rýchlo nahrádzané bunkami novými. Z tohto vyplýva, že imunodeficitní pacienti utrpia len mierne poškodenie pečene, ale sú viac náchylní k vyvinutiu stavu nosičstva (chronické ochorenie).

## 2.1 Popis použitých postupov

Pre čo najreálnejšie namodelovanie trojrozmerného objektu je najlepšie využiť priestorovú reprezentáciu modelu, k čomu sa vynikajúco hodí celulárny automat [7]. Dokážeme ním zjednodušiť reálny systém buniek v tele natoľko, aby sme v rozumnom čase dokázali získavať výsledky z experimentov. Jeho implementácia vďaka viacrozmerným poliam v počítačoch je veľmi jednoduchá, efektívna a hlavne veľmi dobre paralelizovateľná.

Vďaka priestorovej reprezentácii sa dá predpokladať jednoduché modelovanie rôznych orgánov s využitím vedomostí z biológie (o veľkosti, hustote a rozmiestnení buniek a podobne).

Bunky sú v celulárnom automate modelované ako prvky uchovávajúce svoj stav v priestore a čase, pričom všetky tieto veličiny sa menia diskkrétne.

V našom modeli je celulárny automat modelovaný ako dvojrozmerné pole, pretože budeme skúmať šírenie vírusu lokálne. Chceme overiť vplyv vybraných parametrov výsledného modelu (ako ukazuje článok) a to, že nehomogenita priestoru má nezanedbateľný vplyv na šírenie vírusu.

## 3 Koncepcia

Ako už bolo spomenuté v predchádzajúcich kapitolách, v našom modeli budeme reprezentovať jednotlivé bunky pomocou prvkov dvojrozmerného poľa. Kvôli zjednodušeniu a efektívnosti v simulácii nebudeme uvažovať vírusové častice. Miesto toho budeme modelovať šírenie vírusu ako prechod nákazy z infikovanej bunky na bunky okolité.

Takisto nebudeme uvažovať imúnne bunky a bunky zabezpečujúce imunitnú reakciu. Imunitná reakcia, ktorej výsledkom je odumieranie infikovaných buniek, bude modelovaná pomocou skrátenia životnosti týchto buniek.

Potrebu delenia buniek pri ich úbytku (regenerácia) budeme modelovať ako nahrádzanie starých, odumretých buniek novými zdravými bunkami. Keďže rýchlosť tejto regenerácie tkaniva je v pečeni

veľmi vysoká, bude tomu odpovedať aj pravdepodobnosť nahradenia mŕtvej bunky zdravou.

### 3.1 Stavový automat

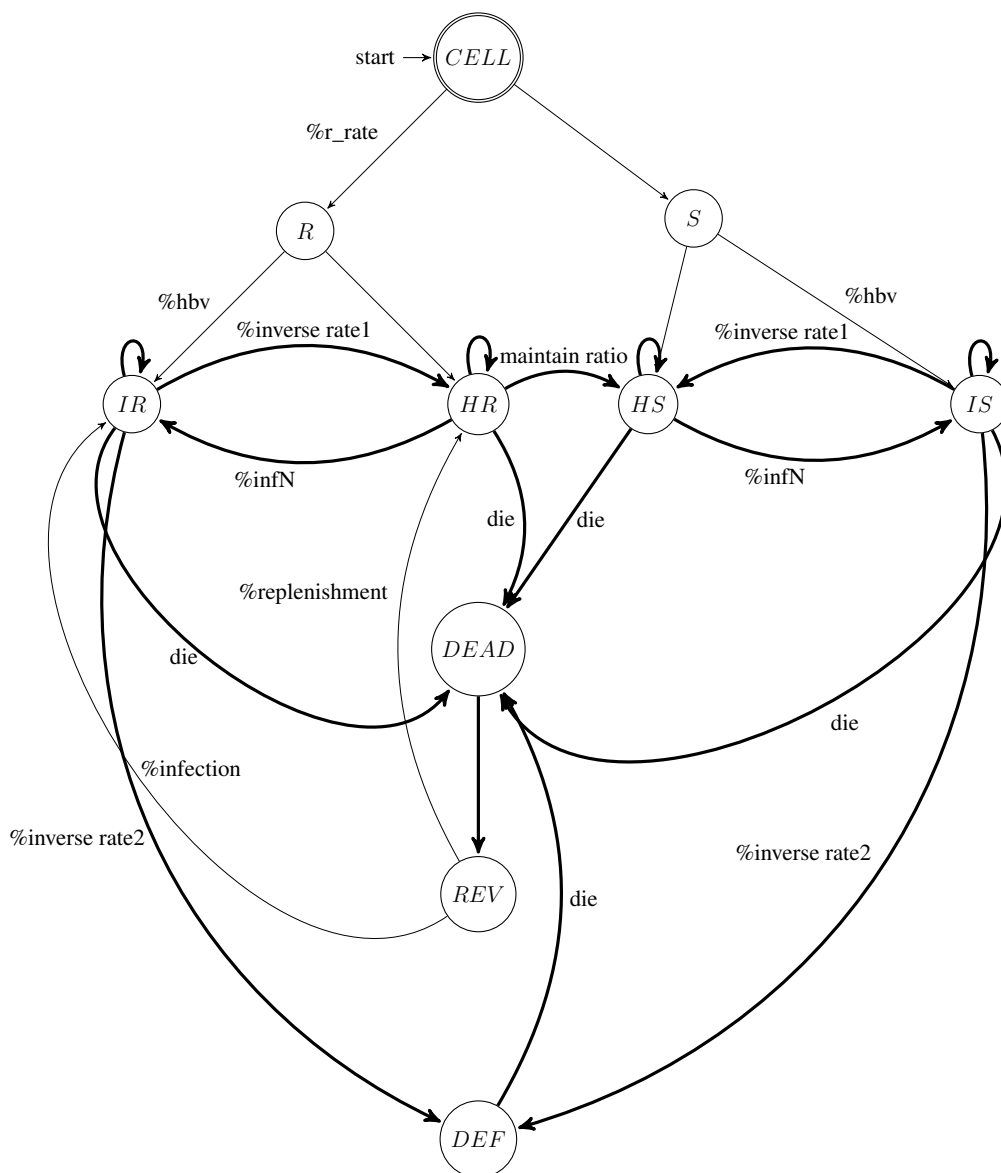
Stavový automat (obr. 1) zobrazuje prechody medzi stavmi buniek tak, ako sú implementované neskôr v simulačnom modeli. Znak % pred názvom prechodu označuje, že prechod sa vykoná s určitou pravdepodobnosťou.

Na začiatku sú všetky bunky (prechod %r\_rate na obr. 1) rozdelené na R (resistant) a S (susceptible). Podľa počiatkovej infekcie h<sub>bv</sub> sa z nich stávajú buďto infikované alebo zdravé bunky. Na tomto mieste sa spúšťa simulácia a prechody, ku ktorým dochádza synchronne, tj. tie, kde sa mení modelový čas [8, str. 21], sú vyznačené hrubšie. Ostatné prechody sú popísané v sekcii 4.1.

Vysvetlenie názvov stavov:

REV = revived; DEF = defectively infected

IX = infected X; HX = healthy X ... pričom  $X \in \{R, S\}$



Obr. 1: Stavový automat

## 4 Architektúra simulačného modelu

Simulačný model bol napísaný v jazyku C++ s využitím knižnice 'Bitmap Image Reader Writer' pre vykreslenie stavu celulárneho automatu do BMP formátu. Program obsahuje dve triedy nazvané *CA* a *Cell*.

Trieda *Cell* obsahuje informácie o bunke (jej aktuálny stav a životnosť). *CA* nám reprezentuje stav celulárneho automatu (tj. obsahuje polia buniek a parametre ovplyvňujúce systém).

Na začiatku simulácie je pole buniek inicializované. Podľa nastaveného parametru `CA::hbv`, ktorý určuje počiatočnú infekciu buniek v systéme, sa vytvoria zdravé aj infikované bunky v potrebnom pomere (náhodne-rovnomerne rozložené v poli).

Následne sa menia stavy buniek pre ďalší krok, pomocou metódy `CA::nextStep()` a pomocou definovaných pravidiel - metóda `CA::gen()`. Pri každom kroku celulárneho automatu sa zaznamenáva počet zdravých/infikovaných/mŕtvych buniek kvôli štatistike a grafom.

### 4.1 Mapovanie konceptuálneho modelu do simulačného

Stavy bunky:

- zdravá bunka: stav `HEALTHY_R` alebo `HEALTHY_S`
  - R (resistant) = žľčové epitelové bunky
  - S (susceptible) = hepatocyty
- infikovaná bunka: stav `INFECTED_R` alebo `INFECTED_S`  
(R/S rovnako ako v predchádzom príklade)
- bunka nakazená chybným vírusom: stav `DEF_INFECTED`
- mŕtva bunka: stav `DEAD`

Regeneračná aktivita je modelovaná ako pravidlo 4a) z článku [1], kedy sú nahrádzané mŕtve bunky živými s pravdepodobnosťou `CA::replenishment` (simulácia delenia buniek). Pomer medzi R a S bunkami sa udržiava pomocou stáleho sledovania ich počtu a preklápania potrebného počtu buniek (podľa `CA::r_rate`) so stavom `HEALTHY_R` do stavu `HEALTHY_S` (prechod *maintain ratio* na obr. 1).

Rovnica pre určovanie, či bude zdravá bunka v danom kroku nakazená od svojho okolia (prechod *%infN* na obr. 1) sa odvíja z rovnice:

$$INFECT\_RATE\_R * n/8 + INFECT\_RATE\_S * m/8$$

Pričom *n* je počet infikovaných R buniek a *m* je počet infikovaných S buniek v okolí najbližších 8 susedných políčok.

Pre modelovanie nákazy, ktorá sa v S bunkách šíri rýchlejšie ako v R, je zvolený parameter `INFECT_RATE_S > INFECT_RATE_R` ( $0.6 > 0.5$ ).

Nakazenie chybným vírusom je modelované ako šanca každej infikovanej bunky v ľubovoľnom kroku na prechod do stavu `DEF_INFECTED` (prechod *%inverse rate2* na obr. 1) Pravdepodobnosť tohto javu je daná parametrom `CA::inverse_rate2`.

Životnosť bunky sa kontroluje v každom kroku. Bunky majú počítadlo (`Cell::lifeTime`), ktoré reprezentuje, aký dlhý čas už strávili v systéme ako zdravé/nakazené. Keďže poznáme životnosť

R aj S buniek, ktorá je pevne stanovená, môžeme toto kontrolovať pri prechode maticou a bunky, ktoré v danom kroku prekročili svoju životnosť, prepíname do stavu DEAD (prechod *die* na obr. 1).

Odumieranie infikovaných buniek je modelované ako skrátenie ich životnosti oproti zdravým bunkám:

```
CA::cell_lifespan >> CA::infect_r_lifespan > CA::infect_s_lifespan
```

Oneskorená imunitná reakcia, ktorú má za následok mutácia vírusu, je do modelu prenesená ako nenulová doba života infikovaných buniek.

Posledné pravidlo modeluje zavedenie nových infikovaných buniek do systému. Je implementované ako šanca pri "oživovaní" bunky, že bude infikovaná (prechod *%infection* na obr. 1) Šanca je veľmi malá: `CA::infection = 0.1%`.

## 4.2 Použitie simulačného programu

Program sa kompiluje pomocou príkazu `make`. Na spustenie názornej ukážky simulácie slúži príkaz `make run`.

Niektoré parametre simulácie sa dajú nastavovať z príkazovej riadky:

Parameter	Popis
-s N	Vykoná sa N krokov simulácie ( $N \geq 1$ ), východzia hodnota je 1
-l N	Plocha simulácie $N \times N$ ( $N \geq 3$ ), východzia hodnota je 1000
-r N	Pri simulácii bude použitý pomer $N$ ( $0.0 \leq N \leq 1.0$ ) medzi počtom R a S buniek.
-f FILE	Generovanie súboru/ov vo formáte BMP (program dopĺňa číselný prefix a príponu .bmp)
-d FILE	Po skončení simulácie sa do FILE uložia štatistiky v každom kroku simulácie
-g N	Každých N krokov sa vygeneruje BMP súbor

## 5 Podstata simulačných experimentov a ich priebeh

Experimentovaním sa budeme overovať závery a niektoré tvrdenia skúmaného článku. Tie sa budeme snažiť dokázať alebo vyvrátiť pomocou modelu, ktorý sme implementovali podľa dostupných informácií z článku.

### 5.1 Postup experimentovania

Experimentovaním v tomto projekte rozumieme spúšťanie simulácie s rôznymi parametrami, sledovanie výstupov a ich následné zhodnotenie a porovnávanie s výsledkami článku (porovnávanie tvarov grafu, vlastností rozloženia rôznych typov buniek a podobne).

Výstupy simulácie sú:

- Grafické - obrázky vo formáte BMP, ktoré zachytávajú stav CA v určených krokoch simulácie. Ide vlastne o zobrazenie matice, kde každá bunka je reprezentovaná jedným pixelom, ktorého farba odpovedá stavu bunky.

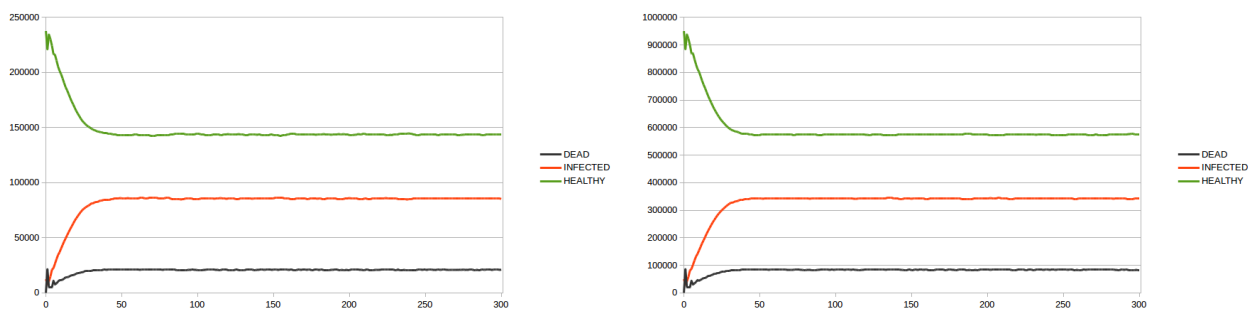
- Textové - výstupom simulácie je textový dokument, v ktorom je zaznamenaný počet buniek v rôznych stavoch. Tieto čísla sa ďalej exportujú do tabuľkového dokumentu, z ktorého sa vykreslí graf. Na tomto grafe je vidieť, ako sa počas simulácie menil napr. počet zdravých buniek, infikovaných a mŕtvych.

## 5.2 Dokumentácia jednotlivých experimentov

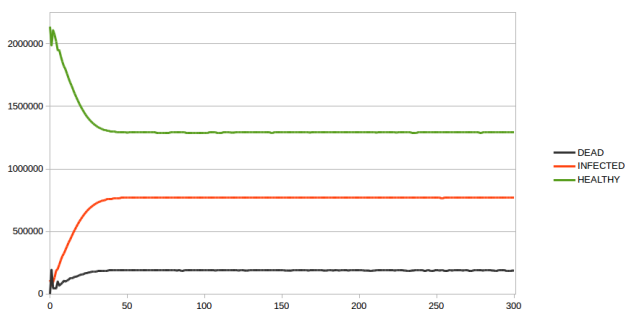
### 5.2.1 Experiment č.1 - nezávislosť na veľkosti matice

V odkazovanom článku [1] je tvrdenie, že simuláciu neovplyvňuje veľkosť matice, ktorá tvorí CA. Toto tvrdenie je dôležité kvôli tomu, že je časovo menej náročné spraviť viacero experimentov na menších maticiach a spriemerovať ich, ako jeden experiment na veľkej matici. Ak je toto tvrdenie pravdivé tak pri rozsiahlejšej simulácii by takéto priemerovanie menších simulácií citel'ne urýchlilo experimentovanie s modelom.

V tomto experimente sme sa teda rozhodli overiť toto tvrdenie. Spúšťali sme simuláciu s rôznym parametrom  $-1$  a zistili sme, že pri zmene veľkosti matice naozaj nedochádza k zmenám štatistických údajov. Na obrázkoch 2 a 3 je znázornený priebeh simulácie s rôznymi veľkosťami matice a z grafov je vidieť, že priebeh sa naozaj nemení. Menia sa iba čísla na osi Y - čo je logické, keďže meníme veľkosť poľa - ale pomer {infikovaných, zdravých, mŕtvych} k celkovému počtu buniek ostáva rovnaký.



Obr. 2: Celulárny automat 500x500 (vľavo) a 1000x1000 (vpravo)



Obr. 3: Celulárny automat 1500x1500

Konvencia:

Zelená farba predstavuje počet zdravých buniek v systéme, infikované sú označené červenou farbou a mŕtve čiernou.



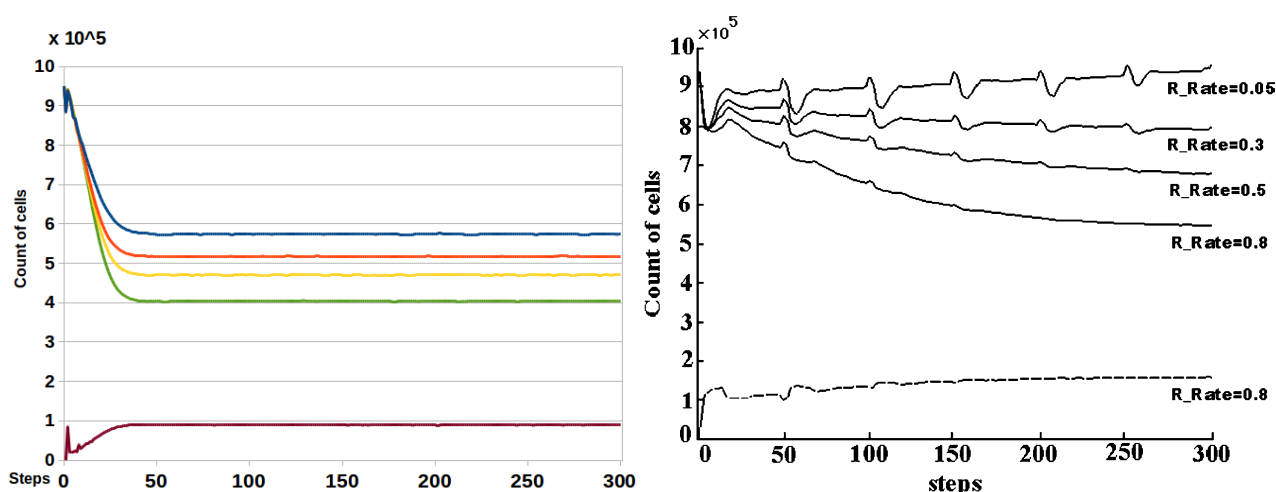
### 5.2.2 Experiment č.2 - Zväčšovanie pomeru R:S zväčšuje infekciu

Z grafu, ktorý prezentuje článok [1] a je zobrazený aj na obr. 4 (vpravo), je vidieť, ako sa mení počet zdravých/mŕtvych buniek v závislosti na parametri  $r\_rate$ . Čím je tento parameter menší, tým je počet zdravých buniek vyšší. V grafe, ktorý sme dostali z nášho simulátora (obr. 4, vľavo), môžeme pozorovať rovnakú vlastnosť.

Všetky parametre simulátora boli nastavené na rovnaké hodnoty ako v článku, grafy sa ale viditeľne líšia vo svojom tvare a v hodnote, na ktorej sa ustálujú krivky (v našom modeli sa infekcia rozmáha rýchlejšie a plynulejšie ako v modeli z článku).

Jediný zákmit, ktorý sa dá pozorovať v našom grafe je spôsobený pridelením náhodnej životnosti buniek na začiatku (0 - CELL\_LIFESPAN) všetkým bunkám. To znamená, že ak sa takýto čas prideli infikovanej bunke, pri prvom kroku simulácie sa často zistí, že čas ktorý jej bol pridelený je vyšší ako jej životnosť (životnosť infikovanej bunky je oveľa menšia). To znamená, že v prvom simulačnom kroku zahynie skokovo naozaj veľa infikovaných buniek - čo sa odrazí aj v grafe. Táto implementácia je prebratá z článku, kde sa nespomínala žiadna výnimka pre generovanie počiatočného času infikovaným bunkám.

Na obrázku vpravo je vidieť pravidelné "zákmity", ktoré sa v našom modeli nevyskytovali. Napriek štúdiu implementácie ich modelu a dôkladnej kontrole toho nášho, sme neprišli na to, čo mohlo tieto zákmity spôsobiť. Vo svojom modeli sme implementovali všetky fakty zo spomínaného článku, preto sa domnievame, že "referenčný" model obsahoval niektoré parametre, ktoré v článku spomenuté neboli a tým pádom sa ich model choval inak.



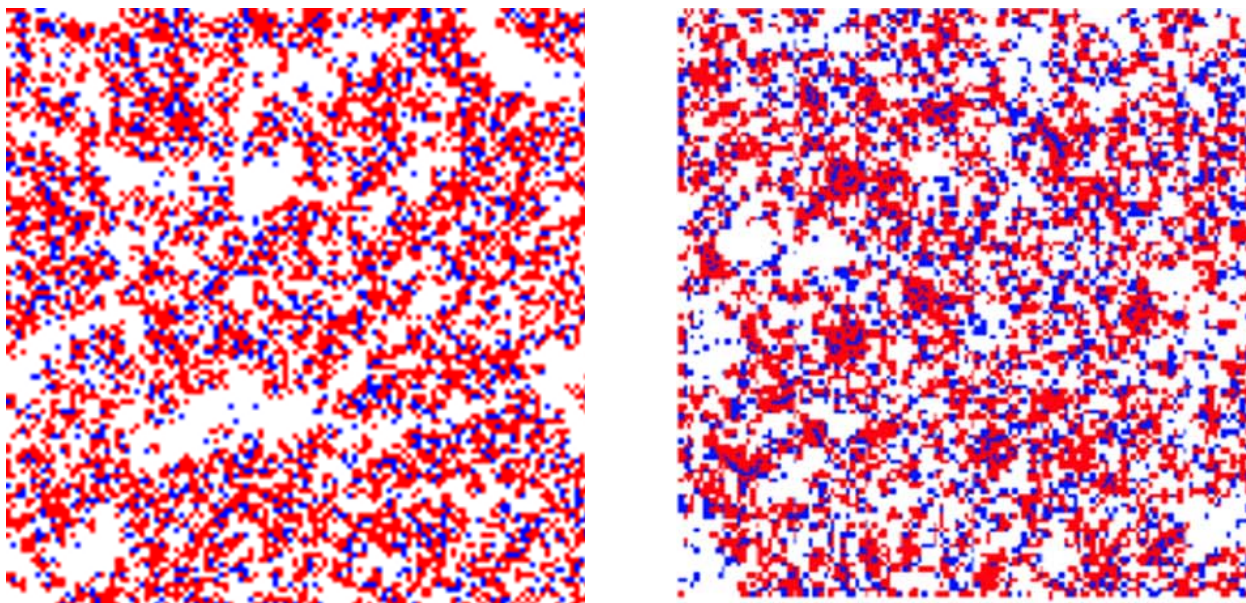
Obr. 4: Porovnanie výstupu nášho simulátora (vľavo) a grafu z článku (vpravo)

### 5.2.3 Experiment č.3 - Zhlukovanie infikovaných buniek

Tretím experimentom sme sa rozhodli overiť zhľukovanie infikovaných buniek. V článku [1] je uvedené, že by infikované bunky mali tvoriť zhľuky (miesta, kde je podiel infikovaných buniek viditeľne vyšší ako v okolí). Ako príklad článok uvádza obrázok, kde je zobrazený stav CA po 30 týždňoch (obr. 5 vpravo).

Pre overenie tohoto faktu sme použili rovnaké parametre a simuláciu sme nechali bežať takisto 30 krokov (týždňov). Náš výsledok (obr. 5 vľavo) sa značne líši od ich výsledku v koncentrácii infikovaných buniek, čo už bolo zistené v predchádzajúcom experimente, avšak zhľuky sa tu naozaj

objavujú. Infikované bunky sú zobrazené červenou farbou - na obrázku je tie miesta pekne vidieť. Ak by koncentrácia infikovaných buniek nebola taká vysoká, obrázky by boli podobnejšie, každopádne tvrdenie, že sa bunky zhukujú je naozaj pravdivé.



Obr. 5: Porovnanie výstupu nášho simulátora (vľavo) a obrázku z článku (vpravo)

### 5.3 Závery experimentov

Na simulátore sme previedli celkovo 3 experimenty. Postupne sme overovali viacero tvrdení a výsledkov článku, pričom nie všetky sa nám podarilo potvrdiť (viď. experiment 2 -> 5.2.2).

Z výsledkov experimentov sme zistili, že na simuláciu nemá vplyv rozmer poľa buniek - pomer počtu jednotlivých buniek za behu simulácie zostáva rovnaký. Potvrdilo sa, že väčší pomer počiatočných buniek R ku S spôsobuje zvýšenú infekciu. Tak isto sme overili, že infikované bunky sa medzi sebou zhukujú, tým pádom vznikajú miesta so zvýšenou koncentráciou infikovaných buniek. Výnimku tvorí rýchlosť infekcie, ktorú sa nám nepodarilo potvrdiť aj pri dôkladnom preskúmaní nášho simulačného modelu.

## 6 Zhrnutie simulačných experimentov a záver

Z výsledkov experimentu vyplýva, že článok obsahuje veľa pravdivých tvrdení, ktoré sa nám podarilo overiť. Narazili sme ale na viacero nezrovnalostí, ktoré nie sú veľmi dobre v článku vysvetlené - usúdili sme, že článok obsahuje neúplné informácie alebo sa v našom simulátore nachádza chyba.

Po dôkladnom skúmaní nášho modelu a rôznych overovaniach funkcionality rôznych jeho častí sa skôr prikláňame k prvej variante, teda že článok nie je úplný/presný, a že v ňom bola zamlčaná relevantná informácia. Neznamená to, že by sa v ňom nachádzali nepravdivé informácie - iba autori možno neuviedli všetky detaily potrebné k vytvoreniu autentického simulačného systému.

## Literatúra

- [1] X. Xiao, S.H. Shao, K.C. Chou: *Biochemical and Biophysical Research Communications*, Volume 342, Issue 2, 7 April 2006, Pages 605-610.
- [2] V.M. Factor, S.A. Radaeva, S.S. Thorgeirsson, *Origin and fate of oval cells in dipin-induced hepatocarcinogenesis in the mouse*, Am. J. Pathol. 145 (1994) 409–422.
- [3] W.T. London, B.S. Blumberg, *A cellular model of the role of hepatitis B virus in the pathogenesis of primary hepatocellular carcinoma*, Hepatology 2 (1982) 10S–14S.
- [4] G.J. Alexander, *Immunology of hepatitis B virus infection*, Br. Med. Bull. 46 (1990) 354–367.
- [5] N.V. Naoumov, R. Schneider, T. Grotzinger, M.C. Jung, S. Miska, G.R. Pape, H. Will, *Precore mutant hepatitis B virus infection and liver disease*, Gastroenterology 102 (1992) 538–543.
- [6] C. Seeger, W.S. Mason, *Hepatitis B virus biology*, Microbiol. Mol. Biol. Rev. 64 (2000) 51–68.
- [7] S. Wolfram, *Cellular automation as models of complexity*, Nature 311 (1984) 419–424.
- [8] PERINGER, P.: Modelování a simulace. [online], september 2015. (VUT FIT v Brně), IMS - slajdy z přednášek.