UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Božidara Cvetković

ANALIZA MOŽNOSTI REALIZACIJE PRIMITIVNIH RAČUNALNIŠKIH STRUKTUR NA OSNOVI DNK GRADNIKOV

Diplomska naloga
na univerzitetnem študiju

Mentor: izr.prof.dr. Miha Mraz
Somentor: prof.dr. Roman Jerala

Ljubljana, 2008

Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

Besedilo je oblikovano z urejevalnikom besedil LaTeX.

Namesto te strani **vstavite** original izdane teme diplomskega dela s podpisom mentorja in dekana ter žigom fakultete, ki ga diplomant dvigne v študentskem referatu, preden odda izdelek v vezavo!

ZAHVALA

Zahvaljujem se vsem, ki so mi kakorkoli pomagali pri izdelavi diplomskega dela. Posebej bi se rada zahvalila mentorju dr. Mihu Mrazu za vložen trud, ter somentorju dr. Romanu Jerali za strokovno pomoč na področju biotehnologije.

— Božidara Cvetković, Ljubljana, September 2008.

KAZALO

Zahvala **i**

Povzetek **ix**

Abstract **xi**

1 Uvod **1**

2 Osnove koncepta biokock **5**

2.1 Zgodovina področja 5

2.2 Delitev biokock 6

2.2.1 Osnovni elementi 7

2.2.2 Naprave 8

2.2.3 Sistemi 9

2.2.4 Povzetek poglavja 9

3 Gradniki in okolja realizacije **11**

3.1 Izbrani gradniki 12

3.1.1 Regulator 12

3.1.2 Mesto vezave ribosoma 13

3.1.3 Zaporedje kodirajočega proteina 13

3.1.4 Terminatorji 14

3.1.5 Konjugator 14

3.1.6 Reporterji 14

3.1.7 Inverterji 15

3.1.8 Generatorji proteinov 15

3.1.9 Signalizatorji 15

3.1.10 Meritve 15
3.2 Okolje 15
3.2.1 Plazmidi 16
3.2.2 Organizmi 20
3.2.3 Mediji 21
3.2.4 Povzetek poglavja 22

4 Modeliranje biološkega vezja 23
4.1 Algoritem 24
4.2 Matematični model 25
4.2.1 Booleanov model 26
4.2.2 Modeli z diferencialnimi enačbami 26
4.2.3 Stohastični modeli 27
4.3 Testiranje in monitoring vezja 29
4.4 Povzetek poglavja 29

5 Logične strukture 31
5.1 Vhodne naprave 32
5.1.1 Vhodni signali 32
5.2 Izhodne naprave 34
5.2.1 Izhodni signali 34
5.3 Biološka logična vrata 35
5.3.1 Modeliranje logičnih vrat 37
5.3.2 Bio NOT vrata 39
5.3.3 Bio AND vrata 40
5.3.4 Bio OR vrata 41
5.3.5 Bio NOR vrata 42
5.3.6 Bio NAND vrata 43
5.4 Ostale logične naprave 44
5.5 Povzetek poglavja 46

6 Sistemi z večjim številom celic 49
6.1 Medcelična komunikacija 50
6.2 Uporabnost večceličnih sistemov 52

6.3 Primer 53
6.3.1 Ideja projekta 54
6.3.2 Eksperiment 55
6.3.3 Rezultati 58
6.4 Povzetek poglavja 62

7 Zmogljivost in varnost **63**
7.1 Zmogljivost 63
7.2 Varnost 64

8 Zaključek **67**

Dodatek A **69**

Seznam slik **71**

Literatura **74**

Izjava **77**

Slovar okrajšav

3OC6HSL .... 3-oxohexanoil-homoserin lakton, najpogosteje uporabljena vrsta AHL

E.coli ..... *Escherichia coli*, prokariontski organizem

S.cerevisiae ... *Saccharomyces cerevisiae*, eukariontski organizem

AHL ..... Acil Homoserin Lakton, kemična snov uporabna kot vhodni signal

aTc ..... Anhidrotetraciklin

CDS ..... Kodirno zaporedje

CMOS ..... Complementary metal–oxide–semiconductor

DNK ..... Deoksiribonukleinska kislina

EcoRI ..... *E.coli* restrikcijski encim tipa II

FaPS ..... Število vezanih transkripcijskih faktorjev na sekundo

GFP ..... Green Fluorescent Protein, zelen fluorescentni protein

iGEM ..... International Genetically Engineered Machine

IPTG ..... Isopropil-beta-D-thiogalactopiranozid

Lac ..... Oznaka označuje izkoriščanje in/ali uporabo laktoze

LacZ ..... Encim β-galaktosidaza

LB ..... L-Broth, standarden medij

M9 ..... Definiran minimalen medij za gojenje bakterij

M9-CA ..... Definiran minimalen medij za gojenje bakterij z dodano casamino kislino

M9-CAglu .... Definiran minimalen medij za gojenje bakterij z dodano casamino kislino in glukozo

M9-glu ..... Definiran minimalen medij za gojenje bakterij z dodano glukozo

mRNK ..... Informacijska RNK

OriT ..... Izvor prenosa, (angl. *Origin od transfer*)

PoPS ..... Število prehodov polimeraz na sekundo

PoPSdc ..... Število prehodov polimeraz na sekundo ob eni kopiji DNK

PstI ..... Restrikcijski encim, ki izvira iz organizma *Providencia stuartii*

RBS ..... Mesto vezave ribosoma, (angl. *Ribosom Binding Site*)

RFP ..... Red Fluorescent Protein, rdeč fluorescentni protein

RiPS ..... Število vezanih ribosomov na sekundo

RiPSmc ..... Število vezanih ribosomov na sekundo ob eni mRNK kopiji

RNK ..... Ribonukleinska kislina

SC ..... Definiran minimalen medij za gojenje kvasovk po sintezi

SD ..... Definiran minimalen medij za gojenje kvasovk

SiPS ..... Število pretoka signalov na sekundo

SOB ..... Super Optimal Broth, standarden medij

SOC ..... Super Optimal Catabolite repression, standarden medij

SpeI ..... Restrikcijski encim, ki izvira iz organizma *Sphaerotilus*

TraJ ..... Geni za prenos, (angl. *Transfer genes*)

tRNK ..... Transportna RNK

VLSI ..... Very Large Scale Integration

X-Gal ..... 5-bromo-4-kloro-3-indolil-β-D-galaktopiranozid, substrat za LacZ, ki ob hidrolizi tvori moder produkt

XbaI ..... Restrikcijski encim, ki izvira iz organizma *Xanthomonas badrii*

YAC ..... Umetni kromosom kvasovke, (angl. *Yeast artificial chromosome*)

YPD ..... Standarden medij za gojenje kvasovk

Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

Božidara Cvetković

Analiza možnosti realizacije primitivnih računalniških struktur na osnovi DNK gradnikov

POVZETEK

Namen dela je predstaviti bioračunalništvo kot enega od možnih načinov procesiranja v prihodnosti. Temeljna opora je odprto skladišče "Biobricks", kjer so DNK gradniki po funkciji jasno porazdeljeni po posameznih skupinah.

Z iskanjem analogij med funkcijami elektronskih gradnikov, ki sestavljajo današnje računalnike in strukturami, ki jih ponujajo biološka DNK zaporedja, smo predstavili način sestavljanja osnovnih logičnih struktur, ki nam bodo v pomoč pri sestavljanju struktur s poljubnimi funkcijami. Temeljne logične strukture, ki sestavljajo digitalne naprave, so logična vrata. Z izdelavo bioloških logičnih vrat, ki jih uvrstimo v polni nabor, lahko sestavimo biološko strukturo s poljubno funkcijo. Tako kot logična vrata v eletronskih vezjih, potrebujejo tudi biološka vhodne signale in izhodni signal. Težava s katero se srečamo pri povezovanju različnih gradnikov v biološko vezje je neuniverzalen tok med njimi. Da bi lahko povezali izhod ene strukture na vhod druge strukture, smo predstavili dodaten funkcional, ki spremeni intenziteto izražanja gena v signal.

Zanesljivo biološko vezje je rezultat zaporedja korakov v katerih načrtujemo, izbiramo ustrezne gradnike, modeliramo, analiziramo in testiramo. V ta namen je podrobno opisan algoritem z vsemi koraki, ki nas pripelje do delujoče biološke strukture.

Medcelična komunikacija je pomembna za delovanje večjih in inteligentnejših sistemov, kjer je sinhronizirano delovanje pomembno. Kakšni so pogoji za komunikacijo med sosednjimi celicami in delujoč primer sistema sestavljenega iz dveh celic, ki med seboj komunicirata, smo pokazali na koncu.

Naloga je podlaga za nadaljnje raziskovanje in realizacijo biološkega sistema, ki bo zanesljivo izvajal poljuben proces.

Ključne besede: biokocke, bioračunalništvo, nanoračunalništvo, bio logična vrata, večcelični sistemi.

University of Ljubljana
Faculty of Computer and Information Science

Božidara Cvetković

Feasibility analysis of realization of primitive computer structures based on DNA assembly elements

ABSTRACT

The purpose of this work is to introduce biocomputing as a possible way of processing in the future. The registry of standard biological parts named Biobricks is our fundamental source of assembly parts, which are already distributed among individual groups by function.

By searching for similarities between functions of electronic components and functions we can produce from DNA sequence, we presented a method to compose primitive logic structures. Digital devices are mainly made of logic gates. By creating biological gates and classifying them as a functionally complete set we can by definition assemble any logic function we can think of. Connecting output to input signal in a biological circuit is a challenge comparing to connection in electronic circuits, because here we don't have a universal current between components. In order to create it we introduced an additional device that converts intensity of protein generation into signal.

A reliable biological circuit is the result of planning, choosing proper assembly elements, modeling, analyzing and testing. We introduced an algorithm that will lead us step-by-step towards a working and reliable biological structure.

Multicellular and intelligent systems mostly have to be synchronized; for that purpose we presented the necessary communication structures, challenges and terms for reliable and efficient communication function.

This paper is the groundwork for future research and realization of biological system that will reliably perform any chosen process.

Key words: biobricks, biocomputing, nanocomputing, bio logic gates, multicellular system.

1 Uvod

Mejnik v preteklosti pred katerim se ni zgodilo nič pomembnega za razvoj računalništva bomo težko postavili. Rečemo pa lahko, da je vsaka realizacija stroja za računanje na nek način prispevala k razvoju računalnika oziroma naprave za procesiranje, kot jo poznamo danes.

Današnje računalnike večinoma sestavljajo vezja izdelana v siliciju, zgrajena v *VLSI* tehniki iz *CMOS* tranzistorjev. Oznaka *VLSI* označuje vezja visoke gostote, kjer danes že na milijone individualnih tranzistorjev sestavlja en čip. Hiter razvoj in potreba po zmanjševanju naprav je botrovala hitremu večanju števila tranzistorjev na čip in naglemu padcu velikosti tranzistorjev. Po Moorovem[1] zakonu, naj bi se vsakih 18 mesecev oziroma približno vsakih 24 mesecev podvojilo število tranzistorjev na elektronskem vezju. S tem zakonom so povezane skoraj vse metrike zmogljivosti digitalne elektronike, kot sta naprimer hitrost procesiranja in kapaciteta spomina.

Ta zakon oziroma napoved velja že več kot 40 let, vendar se z zmanjševanjem ve-

[1] Gordon Earle Moore, rojen 3. januarja, 1929 v San Franciscu, California, je soustanovitelj podjetja Intel Corporation in avtor Moorovega zakona, ki je bil objavljen 19. aprila 1965 v časopisu Electronics Magazine.

likosti tranzistorjev na velikost atomov naglo približujemo temeljnim fizikalnim omejitvam, ki onemogočajo trenutni tehnologiji zanesljivo delovanje tako majhnih gradnikov. Sam Moore je leta 2005 za internetni časopis Techworld [1] komentiral svoj zakon s stališča naglega zmanjševanja velikosti tranzisorjev in potrdil fizikalne ovire.

Iz tega razloga se že nekaj časa razmišlja o alternativnih tehnologijah, ki bi se jih lahko uporabilo za vezja nove generacije. Možne smeri v katere bi se intenziven razvoj lahko začel so kvantno računalništvo, bioračunalništvo, računalniki na osnovi nanocevi (angl. *nanotubes*) in kvantni celični avtomati [2–4]. Ta področja so še vedno v eksperimentalni fazi, vendar so že v samem začetku raziskovanja dala izjemne rezultate v kreiranju elementarnih tranzistorjev, spominskih celic in reševanju matematičnih problemov.

To delo je v celoti namenjeno obravnavi sestavljanja novih DNK struktur za uporabo v bioračunalništvu.

Bioračunalništvo je oblika računalništva, ki uporablja DNK, biokemijo in molekularno biologijo, namesto tradicionalne silicijeve tehnologije. Prepleta se z bionanotehnologijo, katere cilj je kreiranje novih DNK struktur z uporabo molekularnih značilnosti, nukleotidnih kislin in encimov. Novo nastale DNK strukture se lahko nato uporabijo kot gradniki biološkega vezja v bioračunalništvu. Okolje v katerem biološko vezje izvaja svojo funkcijo je celica v živem organizmu. V naravi celice je samoreprodukcija in cepitev kot oblika razmnoževanja, kar prinaša prednost majhne porabe energije in hitrega kreiranja več celic z enako DNK strukturo in s tem tudi funkcijo.

Raziskovanja in kreiranja novih DNK struktur, so se intenzivno lotili na univerzi MIT, kjer so si v ta namen zastavili cilj sistematizirati DNK zaporedja v posamezne funkcionalne skupine. S tem so želeli olajšati razumevanje funkcije, ki jo nosi posamezno DNK zaporedje. Z natančno analizo delovanja vsake novo kreirane strukture so omogočili, da lahko s principom "bottom-up" iz enostavnih gradnikov sestavimo zahtevnejše strukture s kompleksnejšimi funkcijami. DNK strukture so vnesli v vsem dostopno skladišče in projekt poimenovali "Biobricks".

Omenjeno skladišče bomo vzeli za temeljno zalogo DNK gradnikov, katere bomo po posameznih skupinah podrobno opisali, na podlagi funkcij gradnikov naredili modele logičnih struktur in jih definirali v polni nabor. Polni nabor logičnih struktur nam bo omogočil sestavljanje poljubnih logičnih funkcij iz DNK-ja. Spoznali bomo postopek in metode v posameznih korakih postopka, ki nam bodo v pomoč pri gradnji zanesljivega biološkega računalniškega sistema. Da bomo v prihodnje lahko sestavili tudi večji in

inteligentnejši sistem, bomo premotrili zgradbo struktur, ki so bistvene za komunikacijo med celicami.

Zaključili bomo z analizo primera odprte in zaprte zanke z vgrajenimi strukturami za komunikacijo in s tem pokazali potencial bioračunalništva, kot tehnologije procesiranja v prihodnosti.

2 Osnove koncepta biokock

Današnje tehnologije omogočajo, da lahko iz nedavno še mističnega DNK s tehniko rekombiniranja, rezanja in lepljenja, umetno ustvarimo novo zaporedje DNK, ki nosi svojo funkcijo. "Biobricks" je standard, kjer so takšna zaporedja sistematizirana in nam omogočajo izbiro ustreznih gradnikov za sestavo kompleksnejše strukture.

Glede na standard združevanja gradnikov, ki spominja na sestavljanje otroških kock v večje konstrukcije, bomo v nadaljnjem tekstu, namesto angleškega poimenovanja "biobricks", privzeli besedo *biokocke*.

2.1 Zgodovina področja

Biokocke so plod ideje Toma Knighta, raziskovalca v laboratoriju za računalništvo in umetno inteligenco na MIT-ju (Massachusetts Institute of Technology), da bi predstavil principe inženirstva v sintezni biologiji. Sestavljati je začel svojo zbirko enostavnih zaporedij DNK, s ciljem graditi biološke računalniške strukture. Pridružila sta se mu Drew Endy, profesor biološkega inženiringa, prav tako z MIT-ja in Christopher Voigt, sintezni biolog z UCSF-ja (University of California - San Francisco). Rezultat naveze treh ljudi

se je izrazil kot javno skladišče ali repozitorij vseh nastalih DNK zaporedij, pod skupnim imenom *Biobricks*.

Ugotovitev, da bi iz finančnega stališča skladišče težko hitro polnili sami, je vodila k organiziciji letnega tekmovanja, imenovanega *iGEM* (International Genetically Engineered Machine). To je tekmovanje za dodiplomske in podiplomske študente s celega sveta, kjer tekmujejo v sestavljanju bioloških sistemov iz biokock. Študentje polnijo skladišče s sestavnimi deli ter končnimi projekti in hkrati relativno poceni raste količina gradnikov v javnem skladišču.

Raznolikost in sistematiziranost biokock v skladišču bo pripomogla h končnemu cilju ideje Toma Knighta, da bi se ustavaril jasen pristop k nanoračunalništvu (angl. *nanocomputing*) z uporabo bioloških organizmov.

2.2 Delitev biokock

Gradnike biokock hierarhično delimo na tri skupine [5]:

- Osnovni elementi (angl. *Parts*),
- Naprave (angl. *Devices*),
- Sistemi (angl. *Systems*).

O hierarhični delitvi biokock govorimo, ker za izdelavo osnovnih elementov potrebujemo določeno zaporedje DNK-ja, za izdelavo naprav potrebujemo osnovne elemente in za izdelavo sistemov potrebujemo naprave. S pomočjo hierarhične delitve se lahko osredotočimo na sestavljanje kompleksnejših bio sistemov, ne da bi morali pri tem do potankosti vedeti kaj se dogaja na nižjih nivojih. Simboli gradnikov pozameznih vrst so na sliki 2.1.

Definicija povezovanja med vrstami znotraj skupin je še vedno nejasna uporabnikom in izdelovalcem biokock. Privzeli bomo eno od veljavnih definicij (Reshma Shetty in Barry Canton v [6]), ki pravi da s stališča povezovanja biokock velja, da za izdelavo osnovnih elementov uporabimo zaporedja DNK, nastalo biokocko pa lahko povežemo z drugimi osnovni elementi, naprave lahko povezujemo preko standardnih nosilcev signalov (PoPS, RiPS) z drugimi napravami in da je sistem že zaokrožena celota, katere ne moremo povezati z ostalimi vrstami preko standardnih nosilcev signalov.

V nadaljevanju je pregled skupin glede na pomen in delovanje.

Slika 2.1 Simboli biokock.

2.2.1 Osnovni elementi

Osnovni elementi so sestavljeni iz zaporedja DNK, ki nosi svojo biološko funkcijo. Gradnike iz skupine osnovnih elementov je mogoče združiti z drugimi gradniki iz iste skupine po standardu združevanja biokock (angl. *standard assembly*).

V skupino osnovnih elementov spadajo gradniki, katere lahko pomensko opredelimo kot *DNK* (angl. *DNA*), *RNK* (angl. *RNA*), *regulator* (angl. *Regulatory*), *mesto vezave ribosoma* (angl. *ribosom binding sites*), *zaporedje kodirajočega proteina* (angl. *protein coding*), *terminator* (angl. *terminators*) in *konjugator* (angl. *conjugation*). Za hiter pregled lastnosti gradnikov te skupine, bomo razložili kako z uporabo teh nastane veriga proteinov.

Ob vezavi RNK polimeraze (glej razlago v dodatku A) na DNK, se dvojna veriga DNK razpre in se s transkripcijo kodogene verige ustvari mRNK. Transkripcija se lahko ustavi z uporabo *terminatorjev*. Terminatorji so sekundarne zančne strukture DNK, ki zagotovijo, da se polimeraza odstrani iz DNK. Transkripcija se ustavi. Priporočljivo je uporabiti dvojni terminator, ker niso dovolj zanesljivi, da bi en sam terminator ustavil prepis. Ustvarjen mRNK je pripravljen na vezavo z ribosomom.

Del mRNK na katerega se veže ribosom je *mesto vezave ribosoma*. Mesto vezave

ribosoma je pomembno, ker s pravilno izbiro tega gradnika lahko določimo natančnost in učinkovitost vezave ribosoma in s tem začetka translacije mRNK v protein.

Z uporabo gradnika *zaporedje kodirajočega proteina*, izberemo zaporedje s katerim določimo funkcijo proteina in proteinske verige, ki bo nastala kot produkt translacije. Proteinska veriga se nato zvije in se lahko pripne na regulatorno regijo drugega DNK, katero izberemo iz *regulatorjev*. Regulatorji so uravnalni elementi, promotorji, ki definirajo regulatorno regijo in poskrbijo za inicializacijo prepisa DNK v RNK. S promotorji lahko uravnavamo hitrost in smer prepisa.

Element iz gradnikov *konjugatorjev* izberemo takrat, kadar želimo prenos genskih informacij z oddajne celice na sprejemno celico.

Rezultat uspešne združitve množice osnovnih elementov je bolj kompleksna enota, ki jo imenujemo naprava.

2.2.2 Naprave

Naprava je funkcionalna celota sestavljena iz množice osnovnih elementov. Naprava je lahko reporter (angl. *reporters*), inverter (angl. *inverter*), generator proteina (angl. *protein generator*), signalizator (angl. *signalling*), sestavljena naprava (angl. *composite device*) ali *merilna naprava* (angl. *measurement*).

Reporterji vizualno vrednotijo rezultat oziroma stanje naprave, kar je podobno svetilnim napravam v elektronskih vezjih.

Inverter je logična naprava, ki enako kot v elektronskih napravah invertira vrednost, le da je v tem primeru ta vrednost koncentracija določenega proteina.

Generatorji proteinov, glede na zaporedje kodirnega proteina, ki ga vsebujejo, tvorijo protein ob sprejemu iniciatorja, ta je ponavadi signal PoPS.

Naprave, namenjene komunikaciji med celicami so v skupini *signalizatorjev*. Te naprave so bistvenega pomena pri gradnji sistemov, saj omogočajo komunikacijo med več celicami naenkrat in tudi med različnimi vrstami celic. Signalizatorji se delijo glede na oddajne in sprejemne celice.

Med naprave spadajo tudi naprave, ki same po sebi nimajo funkcije in jih je potrebno združiti z drugo napravo, da bi se ustvarila funkcionalnost. Takšne naprave spadajo v skupino *sestavnih naprav*.

Da bi lahko predvideli delovanje potrebujemo parametre s katerimi lahko ovrednotimo moč promotorjev. Ti parametri so minimalni in maksimalni nivo signala, prenosna

funkcija, hitrost prepisovanja, presluh itd.. Te prametre potrebujemo tudi zato, da bomo lahko povezali naprave med seboj. Napravo povežemo z eno ali več napravami le s pomočjo skupnih standardnih signalov.

Najpogostejša signala, ki se uporabljata sta PoPS in RiPS. Signal PoPS označuje število prehodov polimeraz na sekundo, hitrost transkripcije, izpeljanka iz tega signala je signal PoPSdc, ki poda število PoPS na eno kopijo DNK. Signal RiPS, označuje število iniciranih ribosomov na sekundo, hitrost translacije, izpeljanka za število RiPS na kopijo mRNK, pa je RiPSmc. Ti signali so elementi gradnikov *merilnih naprav*. V literaturi [7], ki ni povezana s standardom "Biobrick" se omenjajo tudi drugi signali, kot sta FaPS (število transkripcijskih faktorjev na sekundo) in SiPS (število signalov na sekundo).

2.2.3 Sistemi

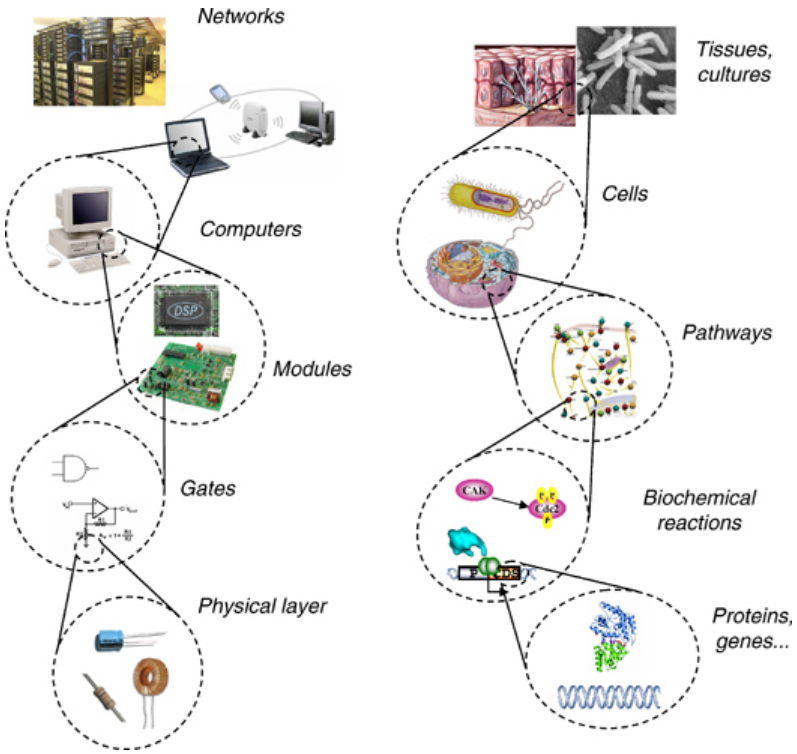
S povezovanjem naprav med seboj lahko sestavimo funkcionalno zaključene sisteme. Sistem ne more komunicirati z ostalimi sistemi s pomočjo standardih signalov. V to skupino spadajo vsi zaključeni projekti, kot naprimer oscilator.

2.2.4 Povzetek poglavja

Za lažjo predstavo hierarhije v sintezni biologiji obstaja več različnih analogij, ki temeljijo na pojmih iz računalniške znanosti. Ena zanimivejših primerjav povzeta po [8], je prikazana na sliki 2.2.

V tem poglavju smo se seznanili z gradniki, ki jih lahko uporabimo pri gradnji primitivnega ali kompleksnega sistema. Uspešno delovanje sestavljene biokocke je odvisno tudi od plazmida v katerega vstavimo biokocko, vrste organizma v katerega vstavimo celico in medija v katerem lahko organizem vrši prevajalno funkcijo definirano z zaporedjem biokock.

V naslednjem poglavju si bomo podrobneje pogledali gradnike, ki so pomembni pri realizaciji logičnih naprav in organizme ter medije v katerih so realizirani.



Slika 2.2 Analogija sintezne biologije.

3 Gradniki in okolja realizacije

Računalniške strukture so sestavljenje iz več povezanih elektronskih komponent, ki dajo strukturi funkcijske in logične značilnosti. Kakšne so te značilnosti, je odvisno od uporabljenih elektronskih komponent, interakcije med njimi in od vpliva okolja. Izbira elektronskih komponent je popolnoma prepuščena snovalcu vezja, pomembno je, da končna funkcionalnost računalniške strukture ustreza predvideni. Snovalec je omejen z velikostjo vezja, s tem je omejeno število komponent, ki jih lahko uporabi, in s kompatibilnostjo izbranih elektronskih komponent med seboj.

Vse predhodno navedbe, ki jih upoštevamo pri snovanju računalniških struktur bi lahko prenesli na snovanje bioloških struktur, le da namesto elektronskih komponent zdaj izbiramo med biokockami. Tudi tokrat ima snovalec svobodo pri izbiri biokock, ki jih bo uporabil za gradnjo svojega sistema. Pomembna omejitev, ki jo mora upoštevati je velikost. Velikost biokocke, dolžina DNK zaporedja, ne sme presegati dolžine zaporedja katerega lahko plazmid sprejme. Število plazmidov, ki jih lahko vstavi v celico je odvisno od velikosti celice in število celic je odvisno od izbranega organizma.

Bistvena razlika med biološkim vezjem in elektronskim vezjem je izolacija.

V elektronskih vezjih so komponente med katerimi ne želimo interakcije izolirane med seboj in od zunanjega medija.

V bioloških vezjih so okoliščine popolnoma drugačne. Komponente niso izolirane niti od medija, niti med seboj. V celici je veliko različnih molekul, ki lahko pridejo v stik z genskim zaporedjem katerega vstavimo. Protein, ki ga biokocka sintetizira lahko uide v celični medij in pride v stik z ostalimi biokockami, kar se lahko izrazi kot nepravilno delovanje ali mutacija. Iz teh razlogov moramo pri izbiri biokock paziti, da je interakcija med izbranimi zelo nizka.

Glavne funkcije današnjih računalniških struktur so *pomnjenje*, *odločanje* in *prenos informacij*.

Ker želimo z biokockami sestaviti biološke računalniške strukture, si poglejmo kateri so tisti gradniki, ki lahko v večji meri pripomorejo k realizaciji željenih funkcij in okolja v katerih so te strukture realizirane.

3.1 Izbrani gradniki

Analogije, ki se pogosto pojavljajo v kontekstu skupine osnovnih elementov biokock so upornik, žica, kondenzator, tranzistor in diode. To so osnovne elektronske komponente, ki jih uporabljamo pri gradnji elektronskih vezij.

Med vsemi vrstami biokock, ki jih ponujajo posamezne skupine, so naslednje tiste, s katerimi lahko vplivamo na ciljno delovanje sistema, da bo primerljiv računalniškim strukturam. Kompatibilnost med posameznimi biokockami, lahko preverimo v skladišču [5] vseh biokock, čeprav bi načeloma naj bili vsi osnovni elementi med seboj povezljivi.

3.1.1 Regulator

Regulatorji so uravnalni elementi, ki definirajo regulatorno regijo, uravnavajo hitrost in smer prepisovanja. Cilj regulatornega elementa je inicializacija prepisa DNK v RNK ob vezavi RNK polimeraze na promotor.

Glede na uravnavanje hitrosti v regulatorni regiji, poznamo dve vrsti regulatornih elementov:

- represivni elementi, opisujejo represivno stanje prepisovanja v regiji,
- inducibilni elementi, opisujejo inducibilno stanje prepisovanja v regiji.

Z *represivnimi elementi* opišemo regulatorno regijo, kjer lahko z naknadim induciranjem posebnega liganda na promotor hitrost prepisovanja znižamo oziroma zaustavimo.

Z *inducibilnimi elementi* opišemo regulatorno regijo v *konstitutivnem stanju*, kjer z naknadnim induciranjem dodatnega liganda na promotor pospešimo prepisovanje.

Za učinkovitejšo uporabo regulatornih elementov je priporočljiva uporaba sestavljenih promotorjev, ker so manj občutljivi na presluhe in tako lahko sprejmejo več PoPS. Prav tako je pomembno za katere kontrolne ligande se odločimo, saj večini učinkovitost na prostem pade. Iz prakse se priporočajo sladkor arabinoza, anhidrotetraciklin (aTC) in Izopropil-beta-D-tiogalaktopiranozid (IPTG).

Smer prepisovanja, je odvisna od izbrane verige dvoverižne DNK, na katero se RNK polimeraza veže. Prepisuje lahko naprej, DNK v mRNK ali nazaj, mRNK v komplement DNK-ja.

3.1.2 Mesto vezave ribosoma

Prisotnost ribosoma je ključnega pomena pri translaciji mRNK v protein. Mesto vezave ribosoma je zaporedje mRNK blizu start kodona (glej razlago v dodatku A), na katerega se ribosom veže pred translacijo. Po uspešni vezavi se začne ribosom premikati vzdolž mRNK in izvaja translacijo dokler ne doseže stop kodona.

S pravilno izbiro mesta vezave lahko določimo natančnost in učinkovitost vezave ribosoma in s tem začetek translacije mRNK v protein.

3.1.3 Zaporedje kodirajočega proteina

Biokocka tega gradnika je zaporedje, ki nosi informacijo potrebno za kreiranje proteinskih verig z ustrezno funkcijo.

Kodirno zaporedje (angl. *coding sequence (CDS)*) se začne s start kodonom (ATG v DNK in AUG v mRNK) in konča pred stop kodonom. Za zanesljiv zaključek kodirnega zaporedja se pogosto uporabita dva različna stop kodona (TAA;TAA). Proteini, ki so rezultat kodiranja lahko nosijo funkcijo encima, represorja, aktivatorja, kodirnega zaporedja reporterja ali ostale proteine, ki ne spadajo v prej navedene kategorije.

Enako kot pri regulatorjih, je smer prepisovanja določena z vezanjem RNK polimeraze na eno od verig dvoverižne DNK. Prepisovanje se lahko izvaja naprej, DNK v mRNK, ali pa nazaj, mRNK v komplement DNK-ja.

3.1.4 Terminatorji

Terminator je DNK zaporedje, katerega funkcija je prekinitev prepisovanja DNK v RNK.

3.1.5 Konjugator

Konjugator je plazmid, ki pripomore k prenosu genskih informacij z oddajne celice na sprejemno celico.

Konjugacijski plazmid, za razliko od ostalih plazmidov, vsebuje informacijo o izvoru prenosa genov, gen ($OriT$), in o genih ki se bodo prenesli, gen ($TraJ$).

Potek prenosa genskih informacij se prične tako, da se del ene verige cirkularnega DNK konjugacijskega plazmida, ki definira izvor prenosa, odreže in se prenese v sprejemno celico. Polimeraze v sprejemni celici poskrbijo da se kreira komplement odrezanega dela verige in s tem celica ugotovi, od kod bodo genske informacije prišle. Ob aktivaciji gena TraJ, se ločijo geni v oddajni celici in se kaskadno prenesejo v sprejemno celico, kjer se združijo z ostalimi geni.

Najpogosteje uporabljeni konjugacijski plazmidi so plazmidi tipa F in R (glej poglavje 3.2.1).

3.1.6 Reporterji

Funkcija reporterjev je vizualno vrednotenje izhoda oziroma stanja naprave ali sistema. Biološko gledano so to fluorescentni proteini. Reporterske biokocke omogočajo uporabo štirih barv za testiranje rezulatatov. Te so zelena, rdeča, rumena in modro-zelena (angl. *Cyan*).

Reporterje delimo na tri podvrste in sicer glede na regijo kodiranja reporterja, na osnovne reporterje in na inducirne reporterje.

Reporterji glede na regijo kodiranja reporterja, sami ne morejo izraziti potrebnega fluorescentnega proteina, zato jim je potrebno dodati še regulatorno regijo in primerno mesto vezave ribosoma. To so zaporedja kodirajočega proteina za regulatorje.

Podvrsta osnovnih reporterjev so reporterji prejšnje podvrste, ki že imajo dodano mesto vezave ribosoma in regulatorno regijo. Ti regulatorji so zmožni sami izraziti protein in jih lahko deaktiviramo z dodatkom tetraciklina v sistem.

Tretja podvrsta vsebuje reporterje, ki aktivirajo regulatorno regijo z naknadnim induciranjem določenega proteina oziroma kemijske snovi.

3.1.7 Inverterji

Inverter ima enak pomen kot v računalništvu, le da v tem primeru dobi na vhod določeno koncentracijo represorja A in na izhodu izrazi določeno koncentracijo represorja B.

3.1.8 Generatorji proteinov

Tipično so generatorji proteinov sestavljeni iz promotorja, mesta vezave ribosoma, regije kodirajočega proteina in enega ali več terminatorjev. Te naprave ob aktivaciji iz mRNK in kodirnega zaporedja generirajo ustrezen protein. Kodiranje aktivira vhodni dregljaj, običajno signal PoPS.

3.1.9 Signalizatorji

Signalizatorji so naprave, ki omogočajo komunikacijo med posamezno celico in bližnjimi sosednjimi celicami. Takšen način komunikacije nam omogoča sinhronizirano delovanje večje populacije celic in komunikacijo med celicami v katerih so naseljeni različni sistemi.

Celica lahko pošlje signal ali pa sprejme povprečen signal od vseh sosednjih celic, ki nosijo enako signalizatorsko napravo. Glede na funkcijo se delijo na sprejemne in oddajne naprave.

3.1.10 Meritve

Ker želimo, da bo delovanje sestavljenega sistema iz biokock zanesljiv, je potrebno ovrednotiti vhode, izhode in notranje delovanje.

Meritvene biokocke nam omogočajo meritev relativne moči promotorja, delovanje pa lahko ovrednotimo z že prej omenjenimi signali PoPS in RiPS (glej poglavje 2.2.2).

3.2 Okolje

Da bi biokocke producirale željeno funkcijo, jih je potrebno vstaviti v primerno okolje. Tehnologija, ki nam omogoča, da lahko prenesemo genske informacije iz enega organizma v drugega se imenuje *rekombinantna DNK*. Postopek rekombiniranja, katerega bomo v nadaljevanju podrobneje spoznali, lahko na kratko opišemo po stopnjah procesa [9]:

- izbrano biokocko združimo z drugo DNK entiteto (klonirni vektor - plazmid), da bi ustvarili novo rekombinantno DNK molekulo,

- klonirni vektor z novo rekombinantno DNK molekulo se vstavi in ohrani v gostiteljski celici,
- celice, ki sprejmejo nov DNK konstrukt, se identificirajo in izberejo, ostale se uničijo,
- odvisno od značilnosti novega DNK konstrukta se kreirajo željeni proteini znotraj gostiteljske celice.

Za uspešno izvedbo navedenega postopka, je potrebno upoštevati značilnosti uporabljenih elementov in medsebojno kompatibilnost teh. Po funkciji moramo skrbno izbrati plazmid, da bo ustrezal gostujoči celici organizma v katerem želimo, da bo naš sistem deloval. Da bi organizem preživel, mu moramo zagotoviti vse potrebne življenske resurse, kar storimo z ustrezno izbiro medija.

V nadaljevanju si podrobneje poglejmo, kakšna je vloga posamezne okoljske entitete, značilnosti teh in stopnje postopka v katerem so pomembne.

3.2.1 Plazmidi

Plazmid je zunaj kromosomska (glej razlago kromosoma v dodatku A), dvoverižna cirkularna DNK molekula v bakterijskih celicah, ki se lahko replicira neodvisno od bakterije. V plazmidu je lahko pd manj kot tisoč do petsto tisoč paznih parov [9]. Število kopij plazmida v gostiteljski celici je lahko od enega do sto kopij na celico, v posebnih okoliščinah tudi tisoč.

Na plazmide lahko gledamo, kot na samostojne življenjske oblike podobne virusom. Plazmid se enako kot virus učinkovito avtonomno podvojuje v ustreznem gostiteljskem okolju. Razlikujeta se v načinu delovanja, virusi so večinoma paraziti medtem ko plazmidi delujejo simbiotično. Plazmid vsebuje uporabne dele DNK, ki pomagajo preživeti celici in plazmidu v stresnih okoliščinah.

Glede na namen plazmida lahko definiramo štiri vrste plazmidov [9]:

- *F plazmidi* (angl. *F plasmids*) nosijo informacijo o prenosu svojih genov iz ene celice v drugo,
- *R plazmidi* (angl. *R plasmids*)vsebujejo gen, ki jih naredi odporne na določen antibiotik ali strup,

- *Razgrajevalni plazmidi* (angl. *degradative plasmids*) vsebujejo množico genov, ki omogočijo presnovo nenavadnih substanc, kot je e.g.toluen ali sialična kislina,
- *Nejasni plazmidi* (angl. *cryptic plasmids*), ki ne nosijo funkcionalnih genov.

Posamezen plazmid lahko spada v več kot eno vrsto prej naštetih plazmidov. Plazmide uporabljamo kot DNK entiteto s katero se združi izbrana biokocka, da bi dobili novo rekombinirano DNK molekulo.

Poglejmo si kako poteka združevanje biokocke s plazmidom in kloniranje plazmida.

Kloniranje plazmida

Termin kloniranje se v klasični biologiji nanaša na novo celico ali organizem, ki izvira iz starševske celice ali organizma. V moderni biologiji se termin nanaša na množico identičnih celic, ki jih z metodo podvojevanja dobimo iz starševske celice. Rekombinantna DNK tehnika nam omogoča, da plazmid v starševski celici sami oblikujemo in s tem dobimo klone lastne celice.

Za kloniranje uporabimo plazmid, ki ga imenujemo *konstrukcijski vektor*, v katerega bomo vstavili biokocko. Ta vektor je najpogosteje plazmid bakterije *E.coli*.

Plazmid v katerega vstavimo eno ali več biokock ima lahko tri regije in sicer, gen za odpornost (angl. *antibiotic resistance marker*), začetno mesto podvojevanja (angl. *replication origin*) in regijo kamor lahko vstavimo naš DNK (angl. *cloning site*).

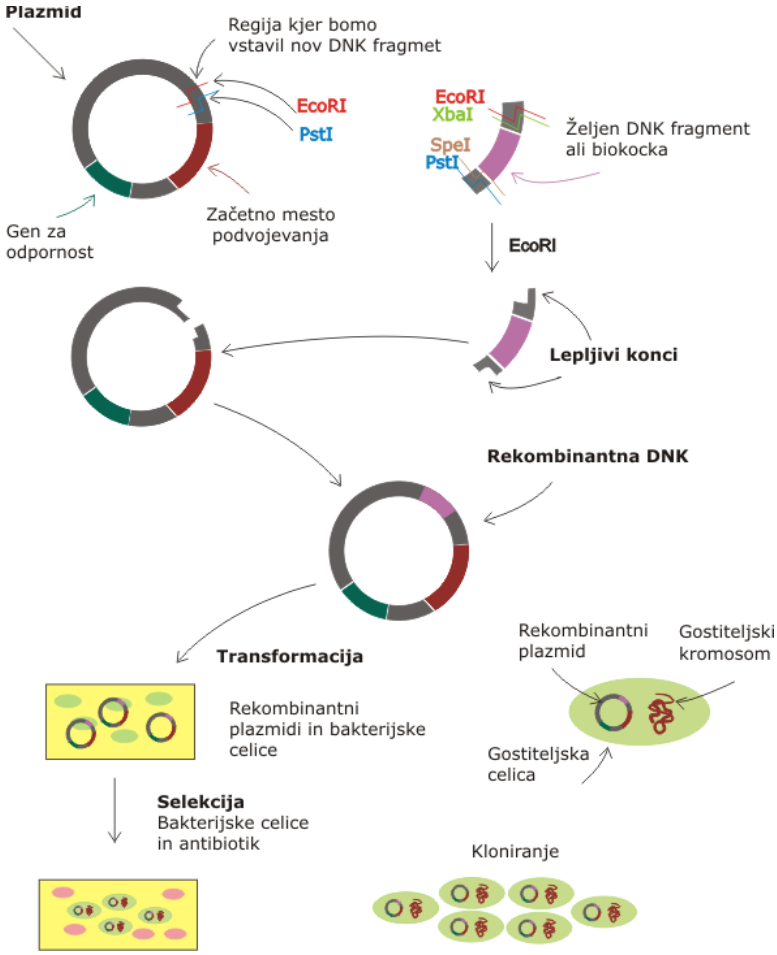
Gen za odpornost vsebuje enega ali več antibiotikov, ki delujejo kot selektivni agenti. Uporaba tega gena omogoča, da ob pojavitvi koncentracije specifičnega antibiotika celica preživi le, če vsebuje ta antibiotik v plazmidu. Antibiotiki, ki se običajno uporabijo, so *ampicillin*, *tetracycline*, *chloramphenicol kanamycin*, *hygromycin B*, *neomycin* in *streptoycin*.

Vsebovanost začetnega mesta podvojevanja omogoča podvojevanje plazmida v gostiteljski celici. Če plazmidu ta regija manjka, se ne bo podvojeval. Izbrano biokocko ali več biokock vstavimo v regijo, ki je za to namenjena.

Da bi vstavili naše DNK zaporedje, moramo najprej razcepiti plazmid. Cepitev omogočajo restrikcijski encimi (angl. *restriction enzymes*). V standardu biokock se za cepitev plazmida in kreiranje vektorja uporabljajo encim restrikcijska endonukleaza ali restrikcijski encim *Eco*RI in restrikcijski encim *PstI*. Encima cepita plazmid na določenem predelu, ki ga spoznata zaradi palindromskega DNK zaporedja na posamezni verigi. Z

razcepom dobimo dva konca, ki ju imenujemo *lepljivi konci* (angl. *sticky ends*).

Poleg prej omenjenih restrikcijskih encimov se uporabljajo tudi drugi restrikcijski encim, kot naprimer *XbaI* in restrikcijski encim *SpeI*. Uporabljata se za cepitev plazmidov, ki vsebujejo biokocke, katere želimo vstaviti v konstrukcijski plazmid. Obstaja več protokolov sestavljanja različnih fragmentov DNK v celoto. Glede na uporabljen protokol so določeni tudi uporabljeni restrikcijski encimi. Na sliki 3.2, je predstavljen koncept sestavljanja dveh biokock v enotno biokocko.



Slika 3.1 Cepitev plazmida in potek kloniranja plazmida, ki vsebuje rekombinantno DNK.

V nastalo praznino plazmida lahko vstavimo sestavljeno biokocko, katere konci ustrezajo zaporedju na lepljivih koncih. Encim DNK ligaza poveže plazmid in biokocko v končno celoto. Takšen plazmid imenujemo *rekombinantni plazmid* z novim DNK fragmentom, katerega želimo podvojevati.

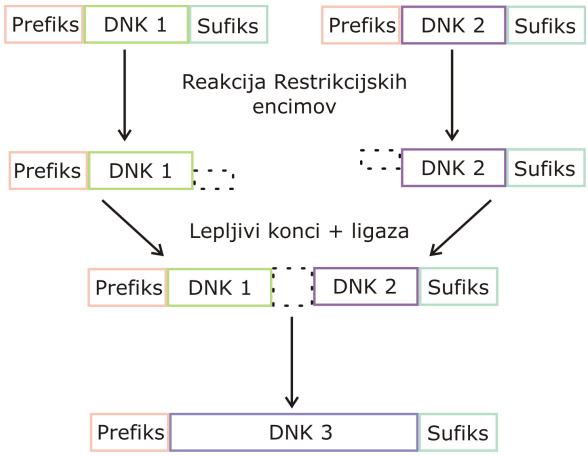
Tako lahko ustvarimo zalogo (angl. *pool*) večih plazmidov z različnimi DNK fragmenti in enakim genom za odpornost.

V naslednji stopnji postopka dodamo ustvarjenim plazmidom bakterijske celice, to so običajno *E.coli*. V procesu, ki se imenuje transformacija (angl. *transformation*), manjše število celic sprejme rekombinirane plazmide, medtem ko večina celic ne.

Da bi ločili celice z našimi plazmidi od tistih ki jih ne potrebujemo, dodamo zmes celic in plazmidov v hranljiv medij, ki ima dodan specifičen antibiotik. Celice, katerih plazmidi vsebujejo antibiotik bodo preživele, tiste brez antibiotika bodo umrle. Temu procesu ločevanja rečemo *selekcija* (angl. *selection*).

Celice, ki so preživele, se bodo zaradi prisotnosti mesta podvojevanja podvojevale in rasle z uporabo encimov v gostujoči celici. Tako nastane kolonija identičnih celic, klonov.

Za vnašanje daljšega DNK zaporedja, kot ga je zmožen sprejeti plazmid, so razvili *E.coli* virus. To je *bacteriophag* λ, ki služi kot vektor in lahko sprejme od 15 tisoč do 20 tisoč baznih parov. Za še daljša zaporedja, pa se za vektor uporablja *umetni kromosom kvasovke* (YAC).



Slika 3.2 Koncept rezanja dveh različnih biokock, DNK 1 in DNK 2, z restrikcijskimi encimi in sestavljanje teh v enotno bikocko, DNK 3.

Iz stališča delovanja računalniškega sistema v dinamičnem okolju, je izraba energije okolja za reprodukcijo in s tem obnavljanje energije prednostna karakteristika. Narediti moramo le en delujoč prototip in ga vstaviti v primerno gojišče. Z reprodukcijo imamo lahko čez noč več tisoč identičnih kopij.

Plazmid je naraven sestavni del bakterije, medtem ko v celicah eukariontskih organizmov temu ni tako. Zgradba eukariontske celice je kompleksnejša od prokariontske celice.

Značilnosti posameznih vrst celic si bomo podrobneje ogledali v naslednjem poglavju.

3.2.2 Organizmi

Pod organizmi razumemo vrsto organizma, katere celico uporabimo za vnos plazmida.

Živeče organizme delimo na dve večji skupini in sicer na *prokariontske organizme* (bakterije) in *eukariontske organizme* (živali, ljudje, rastline, glive). Razlika med vrstama celic je strukturne narave. Glavne razlike so vsebovanje jedra (angl. *nucleus*) v katerem je kromosomska DNK, kemična sestava celične stene in prisotnost organelov.

V nadaljevanju si poglejmo primerka najpogosteje uporabljenih organizmov obeh vrst.

Escherichia coli

Escherichia coli ali *E.coli* je bakterija, ki spada v prokariontske organizme in je ena najbolj preučevanih bakterij na svetu. V zadnjih petdesetih letih so znanstveniki zbrali ogromno količino informacij o njeni genetiki, sestavi in življenju, zato je iz vrst prokariontskih organizmov najbolj primerna za manipuliranje.

E.coli je v obliki palice, gram negativna, nepatogena, kratka in premična. Največkrat jo najdemo v človeškem črevesju, malokdaj pa v zemlji in vodi. Razmnožuje se in živi v medijih, ki vsebujejo amino kisline, vitamine, soli, sledilne elemente (angl. *trace elements*) (Fe^{2+}, Co^{2+}, Zn^{2+}, Cu^{2+} in Mo^{2+}) in izvor ogljika. Čas v katerem se generira nova generacija bakterij v bogatem mediju pri 37°C je 22 minut.

Preživi lahko v okolju brez kisika (anaerobno okolje), vendar za bolj učinkovito produciranje rekombinantnih proteinov je že rutina, da je v okolju s kisikom (aerobno okolje).

Namesto bakterije *E.coli* se uporabljajo tudi drugi bakterijski mikroorganizmi. Delijo se v dve skupini glede na funkcijo: mikroorganizmi, ki proizvajajo gene za posebne funkcije in na mikroorganizme, ki so genetsko spremenjeni za izvajanje določene naloge bolj učinkovito.

Bakterija *E.coli* je najpogosteje uporabljen prokarionstki organizem v kombinaciji z biokockami.

Saccharomyces cerevisiae

Saccharomyces cerevisiae je enocelični eukariontski organizem iz družine kvasovk.

Uporaben je za preučevanje celičnega cikla, ker je enostaven za reprodukcijo v labaratoriju ima pa kompleksno notranjo strukturo celice kot rastline in živali. *S. cerevisiae* je bil prvi eukariontski organizem, katrega gensko zaporedje je v celoti znano. Metabolizem in molekularna biologija tega enoceličarja je tudi dovolj znana, da ga lahko uporabimo kot *E.coli* v prokariontih.

Kot laboratorijski organizem, se *S. carevisiae* razmnožuje v enakem mediju, kot ga uporabljamo za razmnoževanje *E.coli*, čeprav razmnoževanje poteka popolnoma drugače kot pri prokariontih.

S.carevisiae lahko preživi in goji v dveh oblikah, kot haploidna ali diploidna celica. Haploidne celice živijo po principu mitoze (glej razlago v dodatku A) in za njihovo preživetje je lahko usodno stresno okolje. Diploidne celice živijo življenski cikel mitoze, ampak stresno okolje pri njih lahko sproži, da diploidna celica podleže mejozi (glej razlago v dodatku A) in se reformira v haploidno.

Razmnoževanje lahko poteka samo med haploidnimi celicami, kar je podobno kot pri človeških celicah in s tem je zanimivo za uporabo z rekombinantno DNK. Organizmi kvasovk so tudi edina vrsta eukariontskih organizmov, ki se uporabljajo z biokockami.

3.2.3 Mediji

Mediji so hranljiva okolja, gojišča, kjer lahko mikroorganizme vzgajamo in kloniramo. Poznamo *standardne medije* (angl. *standard media*), kjer imajo organizmi na voljo bogato okolje polno primernih organizmu hranljivih snovi in *definirane medije* (angl. *defined media*), kjer so vsebovane le izbrane snovi.

Standarden medij za več vrst bakterij v različnih stanjih je medij imenovan *LB* (L-Broth) in ustreza kot medij za končno izvedbo sestavljene naprave iz biokock.

Standarden medij namenjen bakterijski transformaciji se imenuje *SOC* (Super Optimal Catabolite repression), sestavljen iz glukoze in medija *SOB* (Super Optimal Broth), ki je namenjen gojenju bakterij za kemično kompetentne celice.

YPD je standarden medij za gojenje kvasovk.

Medij *M9* je definiran minimalen medij za gojenje bakterij. Prednosti tega medija sta, da je poceni in da ima zelo nizko autofluorescenco. Obstaja tudi v obliki z različnimi dodatki. Oblika *M9-glu* vsebuje glukozo, *M9-CA* vsebuje kazaminske kisline (angl. *casamino*

acids) (glej razlago v dodatku A) in *M9-CAglu* vsebuje glukozo in casamino kislino. Dodatki omogočajo, da se hitrost rasti bakterij poveča (z glukozo) ali pa da omogoči gojenje bakterij, ki potrebujejo dodane aditive.

Minimalen definiran medij za organizme kvasovk je *SD*, medij *SC* je definiran medij za kvasovke po končani sintezi.

3.2.4 Povzetek poglavja

V tem poglavju smo spoznali biokocke, ki so pomembne pri sestavljanju in testiranju bioloških sistemov, tudi računalniških. Spoznali smo se s postopkom s katerim vstavimo biokocko v živ organizem in pogoje, v katerih lahko izbran organizem preživi.

Ker so biokocke v skladišču večinoma kreirane za uporabo v bakterijskih celicah, bodo vse v nadaljevanju omenjene naprave sestavljene v celici *E.coli*. Od navedenih medijev je najbolj optimalen medij LB, ker ustreza vsem stadijem v katerem se bakterija lahko nahaja.

V nadaljevanju bomo spoznali algoritem za načrtovanje sistema, sestavili osnovne biološke računalniške strukture za odločanje in poskusili definirati poln nabor biokock v primernem mediju z namenom sestaviti kompleksnejšo strukturo.

4 Modeliranje biološkega vezja

Delujoče končno vezje, računalniško ali biološko, je rezultat večkratnega ponavljanja ciklov izbiranja primernih gradnikov, optimizacije in testiranja. Ponavljanje določenih ciklov je še posebno pomembno pri sestavljanju DNK struktur, saj s kompleksnostjo raste tudi število interakcij med gradniki, s ponavljanjem pa lahko opredelimo izvor deviacij in karakteristike posameznih parametrov. Poznavanje delovanja posameznega sestavnega dela celice ne pomeni razumevanja delovanja celice kot celote.

S povečevanjem števila gradnikov se zaplete napovedovanje delovanja. Rešitev problema napovedovanja je uporaba matematičnih orodij. Vezje analiziramo in ga opišemo kot matematični model ali kombinacijo več matematičnih modelov, kar nam omogoči teoretično definirati dinamiko strukture ter izvor in posledice stohastičnosti pri izražanju genov.

Za uspešno realizacijo biološkega vezja v praksi moramo teoretično vezje testirati *in silico* (glej razlago v dodatku A) in že v času oblikovanja izbrati primerne gradnike, ki bodo pripomogli pri nadaljnjem monitoringu vezja v realnem okolju.

Poglejmo si enostaven algoritem, opisan v članku Marguet et al.[10], ki nas v treh

fazah pripelje do željenega biološkega vezja *de novo* (glej razlago v dodatku A).

4.1 Algoritem

Priprava na modeliranje

1. **Določitev cilja biološkega vezja**

2. **Izbira ustreznega gostiteljskega organizma glede na:**
   - enostavnost genskega manipuliranja,
   - hitrost podvajanja,
   - preživetje v željenem okolju,
   - notranje lastnosti organizma.

3. **Izbira ustreznih biokock (osnovni deli, naprave)**
   - ni potrebno, da je DNK zaporedje iz gostiteljskega organizma,
   - če primerna biokocka ne obstaja, jo lahko kreiramo sami,
   - več informacij kot imamo o biokocki, lažje jo bomo uporabili,
   - priporočljiva je uporaba reporterjev, saj s tem lahko ovržemo ali sprejmemo pravilnost delovanja posamezne konstrukcije.

Modeliranje

4. **Priprava matematičnega modela**
   - začnemo z najmanjšim modelom, ki zajame dinamiko vezja.

5. **Raziskava dinamike vezja *in silico* (glej razlago v dodatku A)**
   - Ali dobim željen rezultat?
   - Kateri parametri so najbolj kritični?
   - Kako se dinamika spreminja ob spremembi parametrov?

Impementacija, testiranje in razhroščevanje

6. **Določitev ustrezne DNK implementacije vezja**

7. **Sestava večih verzij, v kolikor obstaja kritičen parameter**

8. **Testiranje vezja, na podlagi tega se odloči ali je potrebno preoblikovanje, ponovno testiranje ali popravek.**

Algoritem tvorijo tri faze. Prva faza, v algoritmu označena kot *priprava na modeliranje*, nas vodi skozi proces izbire ustreznih resursov, s katerimi bomo uresničili idejo o biološkem vezju. V prejšnjem poglavju smo opredelili funkcionalne značilnosti in omejitve posameznih biokock in organizmov, ki so nam na razpolago in s tem tudi zadostne informacije za uspešno izvedbo vseh stopenj prve faze.

Cilj druge faze algoritma, označene kot *modeliranje*, je analitična potrditev pravilnosti delovanja konstrukcije iz izbranih gradnikov, ki smo jo teoretično sestavili v prejšnji fazi. Za analizo *in silico* uporabimo programsko opremo, ki je namenjena modeliranju biokock. Pred tem z ustreznimi matematičnimi orodji opišemo dinamiko vezja. Matematična orodja bomo spoznali v poglavju 4.2. Če se izkaže, da matematični model ne izraža cilja, katerega smo si zastavili, moramo ugotoviti v kateri fazi je napaka. Napaka v drugi fazi pomeni vnovično analizo matematičnega modela. V kolikor je napaka že v prvi fazi moramo ugotoviti ustreznost izbranih resursov in jih po potrebi spremeniti ali zamenjati.

Vloga matematične analize na vseh področjih je analitično definirati delovanje in s tem zmanjšati število neuspešnih poskusov. V našem primeru z matematično analizo delovanja privarčujemo veliko denarja in časa, ki bi ga porabili pri takojšnji implementaciji rezultatov prve faze. Ko se izkaže matematični model za pravilen, naredimo še testiranje *in silico*.

Ko potrdimo pravilno delovanje v teoriji ga moramo potrditi še v praksi. V tretji fazi algoritma, *implementacija, testiranje in razhroščevanje*, implementiramo v drugi fazi potrjen rezultat izbire prve faze. Po uspešni implementaciji sledi postopek ugotavljanja pravilnosti delovanja konstrukcije v željenem okolju. O načinih testiranja bomo več povedali v poglavju 4.3.

V nadaljevanju si poglejmo načine na katere lahko izvedemo optimizacijo [11] in kako testiramo implementirano konstrukcijo.

4.2 Matematični model

Za pravilno delovanje strukture je potrebno analizirati in upoštevati vse vplive, ki lahko spremenijo željeno delovanje sistema. Optimizirati je potrebno število gradnikov in zmanjšati negativen medsebojni vpliv. Cilj je obdržati konsistenco med genomom (glej

razlago besede v dodatku A) in izraženim fenotipom (glej razlago besede v dodatku A).

Vloga matematičnega modela je integrirati vplive posameznega fragmenta v celoto, saj vsak posamezni gradnik doprinese k večjemu številu interakcij v celici. Izvajanje procesov v celici se izraža tudi navzven in vpliva na pravilnost medcelične komunikacije.

V uporabi so *deterministični* in *stohastični* matematični modeli. Izbira modela je odvisna od števila celic v populaciji. Za analizo vpliva šuma v eni sami celici se uporablja stohastični model, za populacijo več celic pa determinističen.

Da bi lahko sistem opisali kot matematični model, je potrebno predhodno definirati produkt vsakega posameznega gena z biokemičnimi, biofizičnimi ali genetskimi tehnikami.

Produkt na ravni mRNK ugotovimo z DNK mikro-mrežo (angl. *DNA microarray*), na ravni proteinov pa z masno spektrometrijo (angl. *mass spectrometry*).

Posamezni modeli, ki so uporabljeni za analizo dinamike v časovni domeni [11], so podrobneje opisani v nadaljevanju. Analizo dinamike v prostorski domeni bomo spoznali v poglavju 6.1, kjer bomo podrobneje spoznali principe in uporabnost medcelične komunikacije.

4.2.1 Booleanov model

Booleanov model je predstavnik *kvalitativnih modelov*. Uporablja se za analizo genskih omrežij in je najenostavnejši model za karakterizacijo dinamike v večjih populacijah.

Binarna logika je preslikana na vsak posamezen gen in omogoča da je lahko v dveh stanjih. Sistem je predstavljen kot logično omrežje, v katerem interakcija med posameznimi geni določa njihovo naslednje stanje v času. Ne glede na enostavnost modela, ga ne smemo prekomerno poenostaviti, saj lahko pride do dvoumnih napovedi obnašanja sistema, kar povzroča nepravilnosti.

Uporaben je pri analizi samo-organizacijskih mrež (angl. *Self-Organizing Maps*), ki so eden od načinov nenadzorovanega učenja in evolucije v kontekstu nevronskih mrež.

4.2.2 Modeli z diferencialnimi enačbami

Uporaba diferencialnih enačb je primer *kvantitativnega modela*, ki se lahko uporabi za modeliranje enoceličnega sistema in sistema večje populacije celic.

Ko oblikujemo model z diferencialnimi enačbami, moramo gledati na biološki sistem kot na mrežo kemičnih reakcij. Vsako kemično reakcijo ovrednotimo z njeno stoihiometrično vrednostjo (glej razlago pojma stoihiometrija v dodatku A) in izrazimo s

stoihiometrično matriko. Z analizo stoihiometrične matrike dobimo obsežen vpogled v delovanje in pregled karakteristike topologije sistema. V formulaciji stoihiometričnih vrednosti manjka povezava s časovnim prostorom, kar nam onemogoči napovedovanja časovne evolucije biološkega sistema. Da bi to omogočili, vpeljemo dodatne informacije o hitrosti posameznih reakcij.

Model, ki zajema kemijsko kinetiko imenujemo *kinetični model* in je izražen kot množica standardnih diferencialnih enačb, ki opisujejo spremembo parametrov v interakciji.

Kinetični model je determinističen, saj ob enakih začetnih pogojih numeričnih konfiguracij, pridemo vedno do istega končnega stanja. Uporaba diferencialnih enačb v celoti zanemari naključne procese, zato je primeren za modeliranje vezja v katerem šum ne naredi velikih sprememb.

Združitev stoihiometričnih vrednosti in kinetičnega modela je kompleksen model, vendar sorazmerno s kompleksnostjo modela raste tudi globina vpogleda v delovanje v času.

Kinetični model lahko združimo tudi z enačbami, ki opisujejo procese pri prenosu. S tem spet zapletemo matematično reševanje, vendar s tem dobimo več informacij o obnašanju sistema. Omenjena združitev je uporabna, ko moramo poleg kemijske kinetike upoštevati tudi promet med komponentama v interakciji. Takšen model je običajno opisan z dodatnimi parcialnimi diferencialnimi enačbami.

Uporabnost kinetičnih modelov je zelo široka, vendar ima to slabost da implicitno privzamemo kontinuiteto interakcij med celičnimi strukturami. Slabost se izrazi v analizi znotraj celičnih procesov. Pri tem si lahko pomagamo s stohastičnim modelom.

4.2.3 Stohastični modeli

Dinamika v posamezni celici vključuje interakcijo manjšega števila molekul, kar se lahko izrazi v visoki ravni šuma tudi pri sintezi enega samega proteina. Model, ki ta problem upošteva pri analizi, imenujemo *stohastičen model*.

Vpliv šuma se v stohastičnem modelu izraža skozi rezultat ponovitve izračuna modela. Posamezna ponovitev izračuna nam vrača različne rezultate zaradi upoštevanja določene stopnje naključnosti. V članku [12] je predstavljen način filtriranja in kontrole šuma ter njegova uporabnost.

Cilj uporabe stohastičnega modela ni predvidevanje šuma v določeni časovni točki, ampak definiranje obsega funkcije in koliko lahko minimiziramo funkcijo ob upoštevanju

naključnosti sistema.

Skrit markovski model

Znan stohastičen proces je Markovski proces iz katerega je izpeljan skrit markovski model (angl. *Hidden Markov model*), ki je široko uporaben. Lahko ga uporabimo za napovedovanje sekundarne strukture proteina [13] in za napovedovanje genov [14].

Skrit markovski model ima tri bistvene elemente. Ti so verjetnostna tranzicijska matrika, verjetnostna matrika izražanja in opazovano zaporedje dogodkov. Z dodatkom učnega zaporedja in uporabo Viterbi algoritma [15] lahko naučimo markovski model napovedovanja sekundarne strukture in napovedovanja genov.

Vrednosti, ki jih vzamemo pri predikciji genov s skritim markovskim modelom so:

- eksoni, introni, itd., so skrita stanja,
- množica nukleotidnih trojčkov v zaporedju so simboli.

Pogostost različnih trojčkov oddanih v stanju eksona (glej razlago besede v dodatku A) je drugačno od tistih oddanih v stanju introna (glej razlago besede v dodatku A). Dejstvo je da stanju introna vedno sledi stanje eksona, kar poveča odvisnost sosednjih stanj. Različne oddajne verjetnosti in tranzicijske verjetnosti med sosednjimi stanji nam omogočajo rekonstrukcijo skritih stanj eksonov in intronov in s tem predikcijo sintetiziranega gena.

Vrednosti, ki jih vzamemo pri napovedovanju sekundarne strukture proteina so:

- α-vijačnica, β-ploskev in zavoj, kot skrita stanja,
- amino kisline.

Strukturalni elementi so skrita stanja, ki oddajajo različne aminokisline z drugačno frekvenco. In ker so nekatere amino kisline večkrat najdene v vijačnici in manj pogosto v ploskvi in zavoju dobimo verjetnosti s katerimi lahko predvidevamo naslednje stanje.

Za zanesljivo predikcijo moramo skrito markovsko verigo naučiti pridobiti verjetnostno tranzicijsko matriko in oddajne verjetnosti. Z znanimi tranzicijskimi in oddajnimi verjetnostmi lahko uporabimo Viterbi algoritem za napovedovanje skritega zaporedja.

Viterbi algoritem je algoritem dinamičnega programiranja za iskanje najbolj verjetnega zaporedja skritih stanj v modelu. To zaporedje se imenuje tudi Viterbi pot in v našem primeru bi to bilo zaporedje genov ali sekundarnih struktur.

4.3 Testiranje in monitoring vezja

Pravilnost definiranega matematičnega modela najprej testiramo *in silico*, torej v programskem okolju, ki omogoča modeliranje in simulacijo dinamičnih sistemov. Primer takšnega okolja je Simulink, paket programskega orodja Matlab.

Da bi ugotovili, ali vezje deluje tako kot v simulaciji, ga moramo testirati tudi v okolju, v katerem lahko realno opazujemo njegovo dinamiko. Eden on načinov je uporaba že omenjenih fluorescentnih proteinov oziroma reporterskih biokock (glej poglavje 3.1.6).

Reporterske biokocke lahko ovrednotijo vsebovanost stopnje proteina v translacijskih procesih ali prisotnost proteina v času transkripcije, kar je izredno elegantna rešitev monitoringa *in vivo*.

Če želimo monitoring v času translacije, vstavimo ustrezno regulatorsko biokocko v bralni okvir (angl. *reading frame*) (glej razlago besede v dodatku A) željenega ciljnega proteina. Translacija sintetizira željeni protein in fluorescentni protein v isto molekulo, kar nam signalizira uspeh sinteze.

Monitoring v času transkripcije lahko izvedemo tako, da se zaporedje kodirajočega proteina in reporter vežeta za istim promotorjem. Paralelno se izrazita fluorescentni in željen protein. Cilj monitoringa je opazovanje aktivnosti promotorja v času transkripcije.

Reporterske biokocke, fluorescentni proteini, so primerni za monitoring dinamike v posamezni celici. Lokaliziramo lahko deviacijo in karakteriziramo celularni šum.

Z vezavo več reporterjev z mRNK vezalnim proteinom, lahko ugotovimo prisotnost ene same molekule, individualne mRNK *in vivo*.

4.4 Povzetek poglavja

Spoznali smo algoritem po katerem lahko uspešno zgradimo lastno biološko vezje *de novo*. Postopek je dolgotrajen in velikokrat je lahko rezultat neuspešen. Z raziskovanjem, uporabo ustrezne kombinacije matematičnih modelov in uporabo reporterskih biokock, se lahko čas iskanja napak precej skrajša.

5 Logične strukture

Ljudje mislimo in sklepamo iz izjav v pogovoru, ki so bodisi resnične ali neresnične. Z upoštevanjem teorije izjavnega računa, ki nam omogoča, da na podlagi izjavnih veznikov sklepamo o resničnosti ali neresničnosti celotne izjave, lahko izjavo razdelimo na posamezne dele, ki so sestavljene s pomočjo izjavnih veznikov oziroma logičnih veznikov (IN, ALI, NE itd.). Resničnost tako formulirane izjave je odvisna samo od resničnosti sestavnih delov, zato so izjavni vezniki definirani s pomočjo pravilnostnih tabel.

Sestavljen sistem lahko prav tako razdelimo na sestavne dele, povezane z napravami ki izvajajo funkcijo logičnih veznikov. Matematični zapis in sklepanje človeškega izročila se prav nič ne razlikuje od zapisa in sklepanja funkcije logične naprave. Tako lahko po pravilih sklepanja izjavnih izrazov potrdimo ali ovržemo pravilnost delovanja sestavljene logične naprave.

Za nekoga, ki bo želel sestaviti kompleksno strukturo, biološko ali računalniško, je matematično izjavno sklepanje zelo uporabno in potrebno. V procesu sestavljanja bo moral uporabiti vezne gradnike, ki bodo kot sestavljena celota oziroma naprava vrnili željeno vrednost. Da bo res tako, pa si bo pomagal z izjavnim sklepanjem.

Karl Raimund Popper [1] je pravilno trdil, ko je rekel, da se ideje ne da nikoli potrditi za pravilno, vendar je zadostno da najdemo en primer, ki bo idejo zanikal, da jo lahko za vedno ovrednotimo kot nepravilno. Matematični izjavni račun in izjavno sklepanje nam omogočata in olajšata ravno to.

V tem poglavju si bomo pogledali, katere so tiste naprave, ki aktivirajo delovanje sistema in katere so tiste, ki nam predstavijo rezultat funkcije sistema. Za naprave, ki izvajajo željeno funkcijo in so med začetnim vhodom in zaključkom, bomo uporabili vezne gradnike ali logična vrata, ki delujejo kot izjavni vezniki.

5.1 Vhodne naprave

V računalništvu je vhodna naprava katerakoli periferna naprava, ki lahko sprejme informacijo oziroma signal in ga posreduje procesni enoti na interpretiranje.

Biološke naprave sestavljene iz biokock obravnavajo PoPS signal kot vhodni signal. Lahko si predstavljamo da je signal PoPS "tok" izražanja gena. Za zagon naprave potrebujemo tok (PoPS), vendar omenjenega signala ne moremo aplicirati direktno, ampak potrebujemo ustrezen konvertor. Vgradimo ga na začetek naprave, tako da bo preoblikoval kemične vhodne signale v PoPS signale.

5.1.1 Vhodni signali

Kandidati za vhodne kemične signale so:

- svetloba,
- IPTG,
- aTc,
- AHL.

Poglejmo si kakšne so lastnosti uporabe posameznih kandidatov.

Uporaba *svetlobe* za vhodni signal, je v realnih napravah nezaželjena zaradi težavnosti dinamičnega nadziranja njene intenzitete. V primerjavi z dodajanjem reagenta v raztopino, je nadzor intenzitete svetlobe na vhodu veliko bolj zapleten. Težavo predstavlja tudi zahteva, da mora biti svetlobni senzor vedno izpostavljen svetlobi. V različnih virih

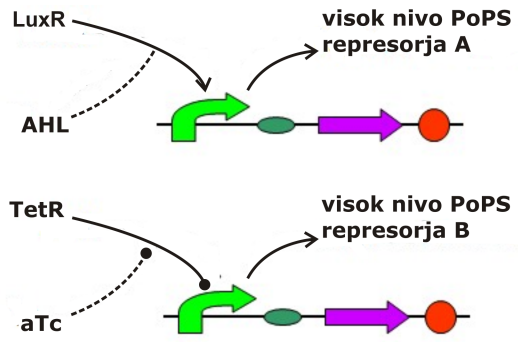
[1] Sir Karl Raimund Popper (Julij 28, 1902 – September 17, 1994) avstrijski in britanski filozof. Velja za enega najbolj vplivnih filozofov in kritikov znanosti.

sicer lahko zasledimo uporabo svetlobe kot vhodnega signala, kot naprimer v [16], kjer so raziskovalci naredili fotografije z uporabo bakterije *E.coli*.

IPTG sproži transkripcijo s promotorjev, katere regulira *LacI* represor. Pravzaprav je IPTG *reset* signal za celoten sistem, ker ob vsebovanosti te kemične snovi vsi pari represor-operator v trenutku izgubijo funkcijo. Vsi izhodi se postavijo na ena, ker se vsi represivni signali na vhodu ignorirajo.

ATc je uporaben pri celularni komunikaciji, v koncentrirani obliki pa se uporablja kot nasprotni učinek izražanja represivnega proteina TetR.

AHL so majhne signalne molekule, ki so uporabne v sistemih, kjer se naenkrat odloča več decentraliziranih enot (angl. *quorum sensing*). Najpogosteje uporabljen je *3OC6HSL*, vrsta iz AHL družine, v kombinaciji z LuxR receptorjem kot senzorjem. LuxR je aktivator, AHL pa njegov ojačevalec.



Slika 5.1 Primer zgradbe dveh vhodnih naprav s prefiksno dodanim konvertorjem. Zgornji primer dobi kot vhodni signal AHL na receptor LuxR in se izrazi kot visoko stanje PoPS. Na spodnjem primeru sta signal in receptor povezana z inhibrno povezavo, kar pomeni da je signal TetR neaktiven v primeru indiciranega aTc. Ker sta signal TetR in konvertor tudi povezana z inhibirno povezavo, je nivo PoPS-a visok, kadar je signal TetR neaktiven. Zgradba je povzeta po ideji iz [5].

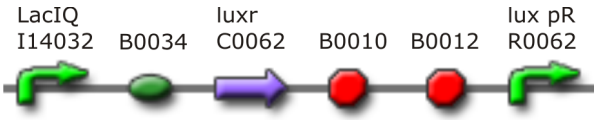
PoPS konvertor

Naštete kemične vhodne signale moramo preoblikovati v kvantitativno vrednost PoPS z ustreznim konvertorjem. Zgradba enostavne vhodne naprave s prefiksno dodanim konvertorjem je na sliki 5.1. Rezultat zgornjega in spodnjega dela je sprememba kemičnih signalov v PoPS signal *lac* aktivatorja LuxR in represorja TetR.

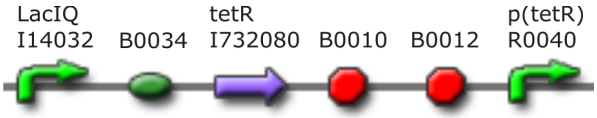
V skladišču biokock lahko najdemo PoPS konvertorje v skupini sestavnih biokock. Oba v nadaljevanju navedena konvertorja sta vzeta iz skladišča [5], biokocka na sliki 5.2,

katero lahko najdemo v skladišču pod oznako *BBa_J09855*, pa je primer konvertorja AHL v PoPS.

Biokocka na sliki 5.3, katero lahko najdemo v skladišču pod oznako *BBa_I732083*, je primer konvertorja aTc v PoPS.



Slika 5.2 Biokocka z oznako BBa_J09855 je konvertor AHL v PoPS.



Slika 5.3 Biokocka z oznako BBa_I732083 je konvertor aTc v PoPS.

To sta le dva primera konvertorja. V skladišču lahko najdemo tudi konvertorje za ostale kemične signale v skupini sestavne bikocke.

Konvertor potrebujemo zgolj na začetku DNK zaporedja sistema. Med moduli ki sestavljajo sistem poteka komunikacija preko generičnih PoPS signalov, ki se generirajo znotraj modulov z uporabo PoPS regulatorjev (glej poglavje 5.3.1).

5.2 Izhodne naprave

Namen izhodnih naprav je prikazati rezultat interpretacije vhodnih signalov v procesni enoti. V sistemih, ki so sestavljeni iz biokock je pravtako potrebno potrditi pravilnost delovanja (glej poglavje 4.3).

5.2.1 Izhodni signali

Za kvantitativen in kvalitativen prikaz rezultata v konstruktu sestavljenem iz biokock lahko za izhodne signale uporabimo:

- LacZ in
- fluorescentne proteine.

Lahko bi uporabili tudi drugačne reporterje, vendar sta omenjena dva načina najbolj raziskana in s tem najbolj prikladna.

Gen *LacZ* je bakterijski encim β-galaktozidaza (angl. *β-galactosidase*). Z dodanim umetnim X-gal substratom kot barvilom, lahko ob cepitvi z encimom β-galaktozidazo enostavno izsledimo dinamiko s prostim očesom. Primeren je za ocenjevanje lastnosti kandidata za logična vrata.

Kandidata, ki prestane LacZ test, lahko z *fluorescentnimi proteini* dodatno ovrednotimo in izberemo najboljšega. Kot smo že omenili v algoritmu (glej poglavje 4.1), ni potrebno, da željena biokocka že obstaja, tvorimo lahko svojo. V primeru, da se zgodi ravno to, je najbolje izbrati fluorescentni protein, o katerem imamo največ informacij. Trenutno je to zeleni fluorescentni protein (GFP), pridobljen iz meduze *Aequeora victoria*.

V skladišču lahko najdemo tudi reporterje, ki lahko predstavijo vrednost PoPS in RiPS. Reporter take vrste najdemo pod oznako *BBa_J13003*.

Slabost trenutnega skladišča je, da imamo premalo informacij o delujočih biokockah. Kratek opis funkcije posamezne biokocke je obetaven, vendar v dodatnem opisu, ki naj bi bil poglobljen in v katerem bi morale biti vse karakteristike in lastnosti biokocke, velikokrat ne najdemo zadostnih informacij.

5.3 Biološka logična vrata

Logična vrata izvajajo logične operacije glede na enega ali več logičnih vhodov in izrazijo rezultat operacije v obliki enega logičnega izhoda. V elektronskih vezjih so logična vrata implementirana z uporabo diod in tranzistorjev, v sintezni biologiji pa jih sestavimo iz ustreznih DNK zaporedij in kemičnih snovi.

Vhodi in izhodi logičnih vrat so v elektroniki predstavljeni kot napetostni nivoji, kar doprinese k enostavnosti pri povezovanju izhoda na enega ali več vhodov drugih logičnih vrat. Posamezen izhod lahko povežemo s končnim številom vhodov, to število določa pojem *fanout*.

Povezovanja bioloških logičnih vrat ne moremo tako poenostaviti, saj potrebujemo dodatne funkcionale, ki spremenijo tok izražanja gena v signal, katerega lahko povežemo na vhod drugih bioloških logičnih vrat.

Osnovne logične operacije logičnih vrat so:

- negacija *NOT*,
- konjunkcija *AND*,
- disjunkcija *OR*.

Kombinacija logičnih operacij osnovnih logičnih vrat so sestavljena logična vrata:

- negirana konjunkcija *NAND*,
- negirana disjunkcija *NOR*,
- ekskluzivna disjunkcija *XOR*,
- negirana ekskluzivna disjunkcija *XNOR*.

Logično operacijo, ki jo izvajajo logična vrata v matematiki zapišemo z izjavnimi vezniki ali z drugim imenom logičnimi operatorji.

Charles Sanders Peirce [2] je dokazal, da se lahko s kombinacijo izjavnih veznikov NOT in AND ali NOT in OR zapiše vse logične operacije, čemur danes pravimo *poln nabor operatorjev.*

Definicija 5.1: *Množica N izjavnih veznikov je poln nabor, če za vsak izjavni izraz A obstaja enakovreden izjavni izraz B, ki vsebuje samo veznike iz nabora N.*

$$\{\neg, \wedge\}, \{\neg, \vee\}, \{\uparrow\}, \{\downarrow\} \tag{5.1}$$

Zapis (5.1), je matematični zapis množice izjavnih veznikov oziroma logičnih operatorjev v polnem naboru. V prvi množici sta logična operatorja NOT in AND, v drugi NOT in OR in sta množici polnega nabora standardnih logičnih operatorjev. V tretji množici je NAND, ki je sam po sebi poln nabor, ker je kombinacija polnega nabora množice NOT in AND. V četrti množici je poln nabor operator NOR, ki je prav tako kombinacija standardnih operatorjev iz množice polnega nabora NOT in OR.

Računalniške strukture so sestavljene iz množice logičnih vrat, ki so ustrezno vezane v zaporedje, ki določa željeno funkcijo. Zaporedje vezave logičnih vrat lahko določimo z različnimi metodami, s katerimi opišemo željeno funkcijo. Ob upoštevanju definicije 5.1 in množice polnega nabora (5.1), lahko sklepamo, da z uspešno realizacijo bioloških logičnih vrat iz množice polnega nabora, lahko sestavimo katerokoli biološko računalniško strukturo.

[2] Charles Sanders Peirce (September 10, 1839 – April 19, 1914) ameriški logik, matematik, filozof in znanstvenik.

5.3.1 Modeliranje logičnih vrat

Vsaka naprava je fizični objekt, ki vsiljuje svojo funkcijo. Pomembni karakteristiki vsake naprave sta njena *prenosna funkcija* in *latenca*. Prenosna funkcija predstavlja stabilno razmerje med vhodom in izhodom, latenca pa je čas, ki ga naprava potrebuje, da se izhod ustrezno odzove na vhodni signal.

Da bomo lažje razumeli analizo in delovanje logične naprave, bomo rekli, da je sestavljena iz dveh delov, in sicer iz *generatorja proteina* in iz *PoPS regulatorja*. V nadaljevanju bomo ločeno analizirali in izrazili prenosno funkcijo in latenco ter ju na koncu združili v prenosno funkcijo ter latenco celotne naprave. Na sliki 5.4 je označeno, kateri del je generator proteina in kateri je PoPS regulator.

Generator proteina

Število generatorjev proteina v logičnih vratih je odvisno od števila vhodov. Pripraviti moramo model, ki nam bo omogočil izraziti prenosno funkcijo in latenco posameznega generatorja proteina.

Vhodni signal v generator je PoPS vhodni signal P_{in}, izhod pa bo koncentracija represorskega proteina R. Prenosna funkcija mora opisati R kot funkcijo od P_{in}.

Upoštevati moramo tako fizični proces med vstopom RNK polimeraze in sintezo ustreznega proteina, kot tudi izgubo mRNK in proteina zaradi degradacije vsakega od teh. Če analiziramo napravo v celici, ki se razvija, potem moramo tudi upoštevati oslabitev mRNK in nivoja proteina zaradi celične delitve.

Vse naštete vplive lahko združimo v diferencialno enačbo, ki opisuje spremembo R-ja po času [6].

$$\frac{d[R]}{dt} = f(P_{in}) - k_d * [R] \tag{5.2}$$

Uvedli smo dve novi vrednosti $f(P_{in})$ in k_d. Funkcija $f(P_{in})$ je zveza med P_{in}, vhodnim signalom PoPS in razmerje sinteze represorskega proteina. Konstanta k_d je stopnja zmanjšanja zaradi izgube R.

Prvi del enačbe (5.2) ovrednoti proces od PoPS do sinteze proteina, drugi del enačbe pa zajema razmerje izgube R zaradi degradacije mRNK in proteina zaradi cepitve celice.

Vrednost konstante k_d je odvisna od proteina, ki ga uporabimo v naši napravi. Nekateri proteini so izredno stabilni in se ne razgradijo hitro(e.g.GFP), drugi pa izginejo precej hitro, če jih znova ne sintetiziramo. Stabilnost proteina opisuje pojem “half-life”,

označen kot $t_{1/2}$, ki nam pove, koliko časa potrebuje protein za degradacijo. Vrednost stabilnega proteina je $t_{1/2} = \infty$, nestabilnega pa $t_{1/2} \approx 10min$. Vrednost k_d lahko enostavno izračunamo po enačbi (5.3).

$$k_d = \frac{0.69}{t_{1/2}} \tag{5.3}$$

Povedali smo, da je prenosna funkcija razmerje med vhodom in izhodom. Funkcija $f(P_{in})$ opisuje vhod in k_d opisuje izhod sinteze proteina. Splošno formulo prenosne funkcije R_{SS}, lahko definiramo z enačbo (5.4).

$$R_{ss} = \frac{(P_{in})}{k_d} \tag{5.4}$$

Naslednja stvar, ki jo moramo analizirati, je latenca generatorja proteina. Izražanje gena je izredno počasen proces, če ga primerjamo z latenco v silicijevih računalniških strukturah.

Zanima nas latenca prehoda iz visokega v nizko stanje L_{HL} in prehoda iz nizkega v visoko stanje L_{LH}. Za oceno latence si lahko pomagamo z enačbo (5.2).

Približno vrednost L_{LH} dobimo tako, da privzamemo, da na začetku R ni prisoten v celici in se vprašamo koliko časa potrebujemo da nivo R-ja v celici doseže stabilno vrednost R_{SS} ob spremembi P_{in} od 0 do najvišje vrednosti. Upoštevati moramo tudi pribitek časa zaradi degradacije na novo kreiranih R molekul.

Za oceno L_{HL} pa privzamemo da je koncentracija R v celici najvišja in nas zanima koliko časa potrebuje za totalno degradacijo, če P_{in} pade z najvišje vrednosti na 0.

PoPS regulator

Tako kot v prejšnjem primeru moramo pripraviti model za določanje prenosne funkcije in latence PoPS regulatorja [6]. Vhodni signal v PoPS regulator je v tem primeru koncentracija represorskega proteina R, izhod pa bo standarden PoPS signal, P_{out}.

Najprej si poglejmo kaj se zgodi v PoPS regulatorju. Protein R se veže na operator (glej razlago v dodatku A) na DNK-ju, kar fizično onemogoči vezanje RNK polimeraze in v tem primeru je P_{out} enak 0. Ko proteina R ni v sistemu, je polimerazi omogočena vezava na operator in s tem se inicializira transkripcija, kar ustvari visok P_{out} signal.

Za izračun latence PoPS regulatorja nas bo zanimal delež operatorja, ki ga zasede R protein ob poljubnem času. Uporabimo lahko diferencialno enačbo s katero ugotovimo

razmerje med represorjem in DNK-jem kot funkcijo časa:

$$\frac{d[R:D]}{dt} = +k_{on} * [D] * [R] - k_{off} * [R:D] \tag{5.5}$$

V enačbi (5.5) smo vpeljali nova koeficienta k_{on} in k_{off} ter spremenljivko D. Koeficient k_{on} je procentualna vrednost, s katero se R veže na operator DNK-ja, k_{off} pa koeficient razdruževanja R in DNK. Vedeti moramo tudi kakšna je koncentracija DNK znotraj celice, ta vrednost je označena z D.

Z znanimi podatki, ki jih potrebujemo za enačbo 5.5, lahko ocenimo latenco PoPS regulatorja.

Za določitev prenosne funkcije PoPS regulatorja bomo tako kot v generatorju proteina potrebovali funkcijo, ki bo ovrednotila P_{out} kot funkcijo vhoda R. Vemo da, če sistem ne vsebuje proteina R, potem bo P_{out} maksimalen in če je vsebovanost proteina R neskončna, bo P_{out} enak 0. Če lahko določimo delež časa, ko je R vezan na operator, bomo lahko ugotovili maksimalen P_{out} in z njim bomo lahko ugotovili dejanski P_{out} kot funkcijo od R. Iz tega dobimo enačbo 5.6, kjer je $K_D = k_{off}/k_{on}$ in N število potrebnih molekul za vezavo operatorja na DNK v primeru, da se veže kot N-mer.

$$P_{out} = P_{out}^{max} * \frac{K_D^N}{K_D^N + [R]^N} \tag{5.6}$$

Ob vezavi zgornjih delov v celoto, dobimo enačbo prenosne funkcije (5.7), ki nas pripelje iz $PoPS_{in}$ do $PoPS_{out}$. Latenca je običajno odvisna od generatorja proteina, tako da lahko privzamemo, da jo grobo lahko ocenimo z enačbo (5.2).

$$P_{out} = P_{out}^{max} * \frac{K_D}{K_D + \frac{f(P_{in})}{k_d}} \tag{5.7}$$

5.3.2 Bio NOT vrata

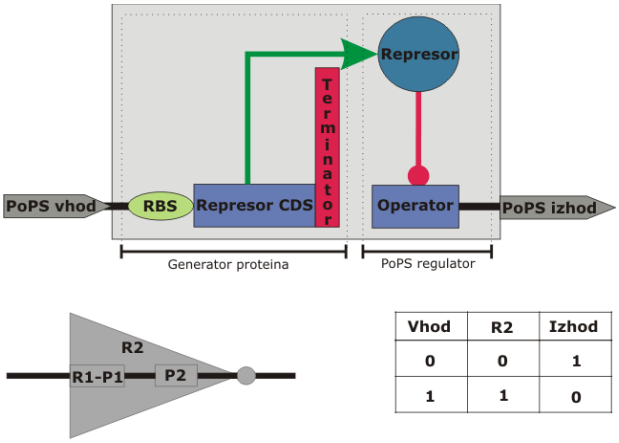
Logična NOT vrata (i.e. inverter) predstavljajo logično funkcijo negacije in za razliko od ostalih logičnih vrat ima NOT le en vhod. V skladišču biokock lahko vidimo, da v skupini naprav obstaja vrsta gradnikov z imenom inverterji in to je hkrati tudi edina vrsta biokocke logičnih gradnikov, ki ima lasten simbol.

Biološki inverter je sestavljen iz mesta vezave ribosoma, zaporedja kodirajočega proteina željenega represorja, terminatorja in promotorja, ki regulira kodiran represorski protein. Regulator kodiranega represorskega proteina imenujemo tudi operator oziroma PoPS regulator (glej poglavje 5.3.1).

Na sliki 5.4 je zajet koncept gradnje bioloških NOT vrat, simbol ki se uporablja v digitalnih vezjih in ustrezna pravilnostna tabela delovanja.

Iz pravilnostne tabele razberemo, da mora v primeru visokega vhodnega signala biti izhodni signal nizek in obratno. Da bi to lahko zgradili iz biokock, si pomagamo z izbiro ustreznega zaporedja kodirajočega proteina. Izbiramo med aktivatorji in represorji. Razlika med njima je funkcionalne narave, aktivator aktivira delovanje; represor zavira delovanje.

Za izgradnjo NOT vrat smo izbrali represorski protein. Ob aktivnem vhodnem signalu, ko je v visokem stanju, se prične kodiranje izbranega represorskega proteina. Ob interakciji represorja z regulatornim elementom operator, slednji prepreči vezavo RNK polimeraze nase in s tem je transkripcija onemogočena, izhodni signal pa je enak nič. Če v sistemu ni ustreznega represorskega proteina, se lahko RNK polimeraza veže na operator in prične s transkripcijo, kar postavi izhodni signal v visoko stanje.



Vhod	R2	Izhod
0	0	1
1	1	0

Slika 5.4 Koncept biološke zgradbe logičnih vrat NOT z PoPS vhodom, ki aktivira izražanje represorskega proteina. PoPS izhod je odvisen od količine vsebovanega represor proteina. Na sliki sta tudi pravilnostna tabela delovanja NOT vrat in simbol, ki se za NOT vrata uporablja v digitalnih vezjih.

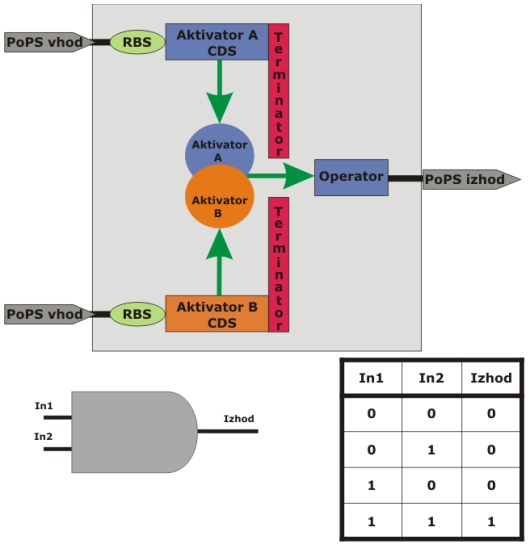
5.3.3 Bio AND vrata

Logična AND vrata imajo lahko za razliko od NOT vrat dva ali več vhodov, nikakor pa samo enega. Delovanje vrat je odvisno od sočasne aktivnosti vseh vhodnih signalov, saj je izhod aktiven le kadar so aktivni vsi vhodni signali. Na sliki 5.5 imamo pravilnostno tabelo, ki opisuje delovanje AND vrat, simbol ter koncept gradnje bioloških AND vrat z dvema PoPS vhodoma in PoPS izhodom.

Gradniki, ki sestavljajo AND vrata so mesto vezave ribosoma, zaporedje kodirajočega proteina in terminator. Zaporedje omenjenih gradnikov je na vsakem od vhodov v logična vrata. V zgradbi je tudi regulatorni element operator, ki bo glede na aktivnost vhodnih signalov izrazil izhodni signal.

Vsak posamezen vhod mora inicializirati kodiranje različnega proteina. V našem primeru sta to aktivator A in aktivator B. Da bi se na operator vezala RNK polimeraza in inicializirala transkripcijo, morata biti v sistemu vsebovana oba proteina. Ko pride do transkripcije operatorja, se izhodni PoPS signal postavi v visoko stanje. Na ta način smo oblikovali ustrezno delovanje logičnih AND vrat.

V [17] so avtorji uspešno realizirali AND logična vrata z dvema vhodoma vezanima na različna aktivatorja in izhodom vezanim na reporter. Vrata so realizirana v bakteriji *E.coli*.



In1	In2	Izhod
0	0	0
0	1	0
1	0	0
1	1	1

Slika 5.5 Koncept biološke zgradbe logičnih vrat AND z dvema PoPS vhodoma, ki neodvisno aktivirata izražanje različnih proteinov aktivatorjev. Hkratno vsebovanje obeh proteinov postavi PoPS izhod v visoko stanje. Na sliki sta tudi pravilnostna tabela delovanja AND vrat in simbol, ki se uporablja za AND vrata v digitalnih vezjih.

5.3.4 Bio OR vrata

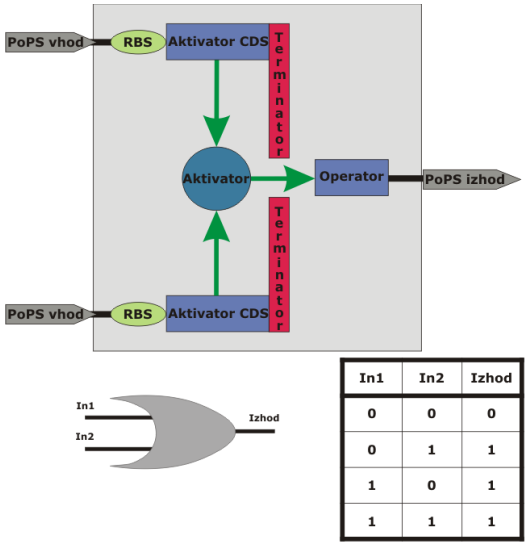
Izhod logičnih OR vrat lahko krmilita dva ali več vhodnih signalov. Izhod iz OR vrat je v visokem stanju kadarkoli je aktiven eden od vhodnih signalov.

Na sliki 5.6 imamo koncept gradnje bioloških OR vrat z dvema PoPS vhodoma in PoPS izhodom, simbol ki se uporablja v digitalni vezjih in ustrezno pravilnostno tabelo

delovanja.

Gradniki, ki sestavljajo biološka OR vrata so mesto vezave ribosoma, zaporedje kodirajočega proteina in terminator. Omenjeno zaporedje je vezano na vsakega od vhodnih signalov. Tako kot v vsakem od prejšnjih primerov, potrebujemo tudi tukaj regulatorni element, ki se bo ustrezno odzval glede na aktivnost vhodov.

Za izdelavo OR vrat smo na vsakega od vhodnih signalov vezali isti aktivatorski protein. Aktivnost kateregakoli od vhodnih signalov se bo izrazila s kodiranjem aktivatorja. S tem se bo omogočilo RNK polimerazi vezavo na operator in rezultat tega bo visok izhodni signal. Izhodni signal bo v visokem stanju tudi, če bosta hkrati aktivna oba vhodna signala, kar ustreza delovanju glede na pravilnostno tabelo OR vrat.



In1	In2	Izhod
0	0	0
0	1	1
1	0	1
1	1	1

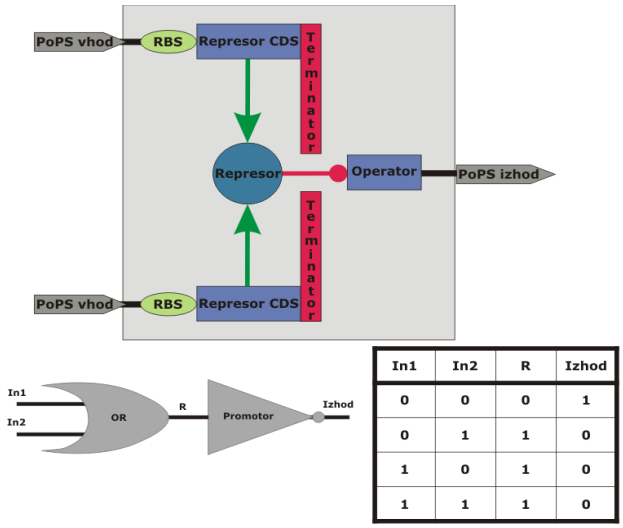
Slika 5.6 Koncept biološke zgradbe logičnih vrat OR z dvema PoPS vhodoma, ki aktivirata neodvisno izražanje identične kopije aktivatorskega proteina. Vsebovanost aktivatorskega proteina postavi PoPS izhod v visoko stanje. Na sliki sta tudi pravilnostna tabela delovanja OR vrat in simbol, ki se uporablja za OR vrata v digitalnih vezjih.

5.3.5 Bio NOR vrata

Logična NOR vrata so sestavljena iz zaporedno vezanih OR in NOT vrat in s tem tvorijo množico logičnih vrat polnega nabora. Iz pravilnostne tabele na sliki 5.7 razberemo, da je izhod iz NOR vrat aktiven le ko vhodi niso aktivni. Lahko bi sestavili NOR vrata z zaporedno vezavo bioloških OR vrat in bioloških NOT vrat, kot to predstavljajo simboli na sliki 5.7, vendar se nam zaradi raznolikosti dinamike proteinov ponuja možnost sestaviti podoben, minimaliziran sistem, ki bo deloval po pravilih NOR vrat.

Na v prejšnjem odstavku omenjeni sliki lahko vidimo koncept gradnje NOR vrat z dvema PoPS vhodoma in PoPS izhodom. Prav tako kot v prejšnjih primerih, je na vsakega od vhodov vezano mesto vezave ribosoma, zaporedje kodirajočega proteina in terminator. Z aktivacijo posameznih vhodov kodiramo enako vrsto represorskega proteina, ker nas bo zanimalo le, če je katerikoli od vhodov aktiven.

Zgradna logičnih vrat NOR vsebuje tudi regulatorni element, ki bo ob interakciji represorja z njim preprečil vezavo RNK polimeraze nase, kar je nakazano z inhibicijsko povezavo. Do interakcije bo prišlo, kadarkoli bo vsaj eden od vhodnih signalov v visokem stanju in bo kodiranje represorja inicializirano.



In1	In2	R	Izhod
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

Slika 5.7 Koncept biološke zgradbe logičnih vrat NOR z dvema PoPS vhodoma, ki neodvisno aktivirata izražanje identične kopije represorskega proteina. Vsebovanost represorja postavi izhodni signal PoPS v nizko stanje. Na sliki sta tudi pravilnostna tabela delovanja NOR vrat in simbolna vezava OR in NOT vrat.

5.3.6 Bio NAND vrata

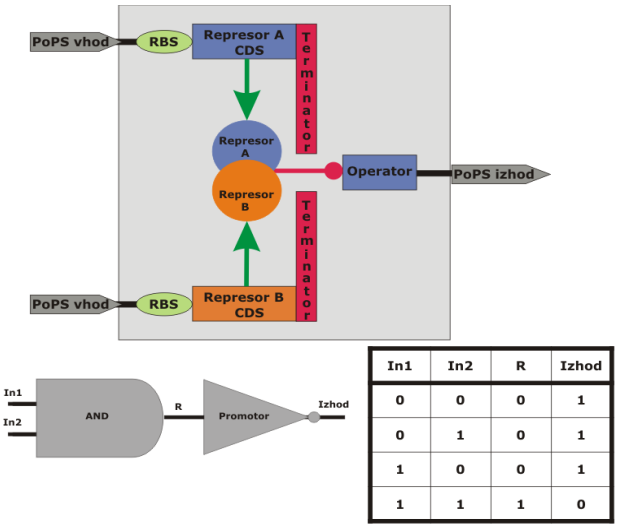
Logična vrata NAND so sestavljena iz zaporedne vezave logičnih vrat AND in logičnih vrat NOT. S tem tvorijo svojo množico polnega nabora.

Gradniki, ki sestavljajo NAND vrata, so mesto vezave ribosoma, zaporedje kodirajočega proteina in terminator. Ti gradniki so zaporedno vezani na vsakega od vhodnih signalov, regulatorni element, ki je pravtako del logičnih vrat, pa je povezan na izhodni signal.

Na sliki 5.8 je koncept zgradbe bioloških logičnih vrat NAND z dvema PoPS vhodnima

signaloma in PoPS izhodnim signalom, povezava z osnovnimi simboli in ustrezna pravilnostna tabela. Kot vidimo, tudi tokrat ne bo potrebno zaporedno vezati AND in NOT vrat, ker lahko izrazimo funkcijo, kot jo predlaga pravilnostna tabela na enostavnejši način.

Za razliko od AND vrat, kjer smo uporabili aktivatorje, smo tukaj na vsakega od vhodov vezali zaporedje, ki bo kodiralo različne represorske proteine. Pogoj za nizek izhodni signal je vsebovanost vseh ustreznih represorjev, ki se začnejo kodirati ob inicializaciji z posameznim vhodnim signalom. Izhodni signal bo neprenehoma v visokem stanju, dokler ne bo v sistemu toliko ustreznih različnih represorjev, kot je vhodnih signalov.



Slika 5.8 Koncept biološke zgradbe logičnih vrat NAND z dvema PoPS vhodoma, ki neodvisno aktivirata izražanje različnih represor proteinov. Hkratno vsebovanje obeh represorskih proteinov postavi PoPS izhod v nizko stanje. Na sliki sta tudi pravilnostna tabela delovanja NAND vrat in simbolna vezava AND in NOT vrat.

5.4 Ostale logične naprave

V prejšnjem poglavju smo si pogledali koncept gradnje logičnih vrat in pravila, ki pripomorejo k gradnji poljubnega vezja za odločanje ali pomnjenje. Za sestavo kompleksnejšega funkcionalnega tako računalniškega kot biološkega vezja, potrebujemo tudi druge dodatne gradnike, ki vsiljujejo ritem delovanja.

Na letnem tekmovanju iGem se zberejo študenti iz različnih univerz in poskušajo uresničiti različne ideje, ki že obstajajo v obliki digitalnega elektronskega vezja. V nadal-

jevanju je naštetih nekaj zanimivejših idej, ki so jih želele realizirati različne ekipe na iGem-u:

- oscilator (Imperial College London, iGem 2006),
- Schmitt trigger (University of Bologna, iGem 2007),
- števec hopov konjugacije celic (Peking University, iGem 2007),
- stikalo (Peking University, iGem 2007),
- PoPS ojačevalnik (University of Cambridge, iGem 2007),
- učenje celic, spomin in prepoznavanje okolja (ETHZ, iGem 2007),
- Delilnik PoPper (angl. *Division PoPer*) (University of Edinburgh, iGem 2007).

Večina naštetih idej je realizirana le delno, vse pa imajo gradnike izredno dobro okarakterizirane. Dokumentiran imajo potek dela ter dosedanje rezultate. Že v uvodu smo omenili pomembost protislovja, zato so zelo pomembne informacije o nezdružjivosti določenih gradnikov in o nedelovanju predvidene sestave gradnikov. To so podatki, ki jih lahko najdemo v dobro dokumentiranih projektih. Med drugim lahko posameznik, ki ni del idejne ekipe, najde v dokumentaciji vse potrebne informacije za izhodišče k nadaljnjem raziskovanju.

Eden najpomembnejših uspehov v sestavljanju bioloških logičnih vezij je uspešno izdelan *oscilator*. Funkcija elektronskega oscilatorja je pretvorba enosmerne napetosti izvora napajanja v izmenično napetost določene frekvence. S tem dobimo valovanje oziroma ponavljajoči se signal katerega uporabimo za sinhronizirano delovanje oziroma srčni utrip vezja.

Model oscilatorja, temelji na modelu plenilec-plen (angl. *Predator-Prey*), ki sta ga podrobno opisala Alfred J. Lotka in Vito Volterra. Velik del potrebnih biokock je bilo izdelanih, modul plena je karakteriziran, matematični model je do potankosti opisan, vendar za končen delujoč oscilator še ni bil izdelan modul plenilca. V članku [18] so avtorji konstruirali sistem dveh *E.coli*, ki delujeta po principu plenilec-plen in raziskali logiko ter dinamiko znotraj celic, kar nas je približalo končni realizaciji prej omenjenega oscilatorja.

Uspešno realiziran je biološki *Schmitt trigger*. Schmitt trigger se v digitalnih vezjih uporablja za stabiliziranje vhodnih signalov. Funkcija omenjenega elementa je primerjava

praga (angl. *threshold*) z vrednostjo na vhodu. Na izhod vrne nizko stanje, če je ta vrednost pod pragom in visoko stanje, če je nad pragom. V primeru, da je vrednost vhodnega signala enaka vredosti praga, se obdrži zadnja vrednost na izhodu. S tem je označen kot stikalo s spominom ali pametno stikalo.

Števec hopov konjugacije celic v *E.coli* je ideja delovanja omrežja znotraj bakterije, kjer naprava šteje število konjugacij med izvorno in ponorno celico. Rezultat projekta sta uspešen matematični model in meritve učinkovitosti posameznih sestavnih gradnikov. Rezultati kažejo na izvedljivost projekta, kjer bo v sistemu več različnih celic.

Projekt *stikala* temelji na ideji stikala na pritisk, le da je v tem primeru pritisk UV svetloba. Sestavljen je iz logičnih vrat NOR, bistabilnega stikala in vhodnega signala. Tudi ta projekt ima obilico dokumentacije in je pripravljen na uspešno realizacijo.

Za zanesljivo delovanje katerekoli naprave v sistemu je potebno pravilno interpretiranje vrednosti PoPS signala. Za jasno vrednost PoPS-a je bil razvit *ojačevalnik PoPS signala*.

Najbolje dokumentiran je projekt, kjer so poskusili učiti celice, jih prilagajati na okolje in jim v ta namen "vgraditi" spomin. Omenjen projekt je sicer teoretično naravnan, vendar je izjemno bogato izhodišče za nadaljnje razvijanje pametnih naprav. Primer takšne naprave je naprava, ki bi temeljila na nevronskih mrežah. Vsak dobro dokumentiran projekt z dobro karakteriziranimi gradniki je tudi dobro izhodišče za posameznike, ki bi nadaljevali raziskovanje in sčasom uspešno sestavili takšen sistem.

Naprava *delilnik PoPper* je projekt, ki bi signal PoPS predstavila na drugačen način. Generirala bi impulz vsakič, ko bi se zgodila delitev celice in bi s tem predstavljala uro sistema. Za štetje v sistemu bi napravo vezali s števcem. Ob vsakem impulzu, bi se števec povečal za ena. S hitrostjo delitve celic lahko dobimo tudi frekvenco delovanja sistema.

5.5 Povzetek poglavja

Cilj tega poglavja je bil predstaviti biološke vhodno-izhodne naprave in njihove signale. Ugotovili smo zakaj potrebujemo logična vrata v računalniških in bioloških napravah ter koncept gradnje teh z uporabo biokock in PoPS signalov. Navedli smo nekaj projektov, ki so jih poskusili realizirati študenti z različnih univerz na tekmovanju iGEM in so zgledno dokumentirani za nadaljnje raziskovanje.

V naslednjem poglavju si bomo podrobneje ogledali koncept medcelične komunikacije, saj je bistvena pri kreiranju večjih struktur zaradi omejitve števila plazmidov znotraj celice, kot tudi zaradi večje svobode pri izbiranju funkcije sistema.

6 Sistemi z večjim številom celic

Realizacija delujočega večceličnega sistema, ki deluje v interakciji s sosednjimi celicami, je temelj za nadaljnji razvoj inteligentnih naprav.

Porazdelitev genetskega omrežja na več celic, na module, in omogočen sistem medcelične komunikacije, poveča možnost kreiranja poljubnega funkcijsko kompleksnega sistema. Prav tako se lahko na ta način preseže zanesljivost delovanja individualne celice. V večceličnih sistemih, kjer je sinhronizacija med celicami bistvenega pomena, je aplikativnost medcelične komunikacije zaželjena.

Z željo po sestavljanju večceličnih sistemov pridemo tudi do neizogibnih zahtev in s tem tudi izzivov:

- za gradnjo kompleksnega sistema, bomo potrebovali različne signalne strukture, ki nam bodo omogočale pošiljanje in sprejemanje,
- komunikacija med različnimi vrstami celic in tudi med populacijami celic zahteva razlikovanje med posameznimi komunikacijskimi kanali,
- potrebovali bomo več različnih gostujočih celic,

- zaželjena je večsmerna komunikacija,
- potrebno je izboljšati natančnost in zanesljivost.

Za snovanje komunikacijskega sistema znotraj večceličnega okolja je pomembno, da je občutljivost medceličnih elementov uravnovešena in presluh na medcelične signale čimbolj izničen.

6.1 Medcelična komunikacija

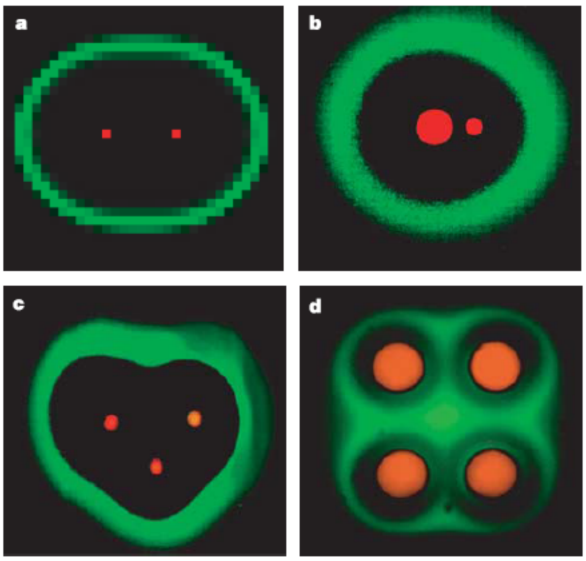
Prehod iz enoceličnih sistemov na večcelične odpre nove možnosti za raziskovanje. Dinamiko tokrat opazujemo in analiziramo v prostoru. Da lahko govorimo o dinamiki moramo omogočiti komunikacijo med celicami.

Medcelično komunikacijo omogočajo posebne strukture, ki so namenjene oddajanju in sprejemanju signalov. Celice, ki vsebujejo strukturo za oddajanje, imenujemo *oddajne celice* (angl. *sender cell*). Tiste, ki pa imajo strukturo za sprejemanje, imenujemo *sprejemne celice* (angl. *receiver cell*). Funkcija oddajne celice je generiranje impulza na katerega sprejemna celica ustrezno odgovori.

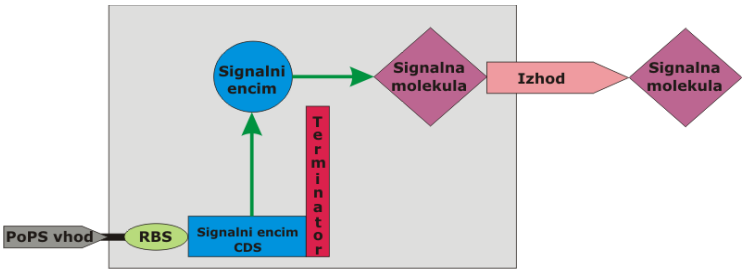
Z eksperimentom v netekočem mediju so raziskovalci Basu et al. [19], potrdili delovanje celic z vgrajenimi prej omenjenimi strukturami. Naloga oddajne celice je bila generiranje signala, ki sproži impulz v sprejemni celici. S tem eksperimentom so ugotovili, da generiran impulz najmočneje vpliva na celice, ki so v neposredni bližini. Bolj kot je bila sprejemna celica oddaljena od oddajne celice, šibkejši je bil generiran impulz. Na podlagi te informacije so ustvarili biološki pasovno prepustni sistem (angl. *band-pass*), kjer so z različnimi fluorescentnimi proteini opisali intenzivnost impulza. Rezultat lahko vidimo na sliki 6.1. Rezultat tega projekta je zelo pomemben za pospešeno razmišljanje o kreiranju večceličnih sistemov, kjer so celice odvisne od stanja sosednih celic.

Na tekmovanju *iGem* je šele letos prišlo do konkretne ideje o večceličnih sistemih, in s tem tudi do kreiranja potrebnih biokock. Tako bo po novembru 2008 skladišče bogatejše za nekaj gradnikov, ki bodo namenjeni delovanju večceličnih sistemov.

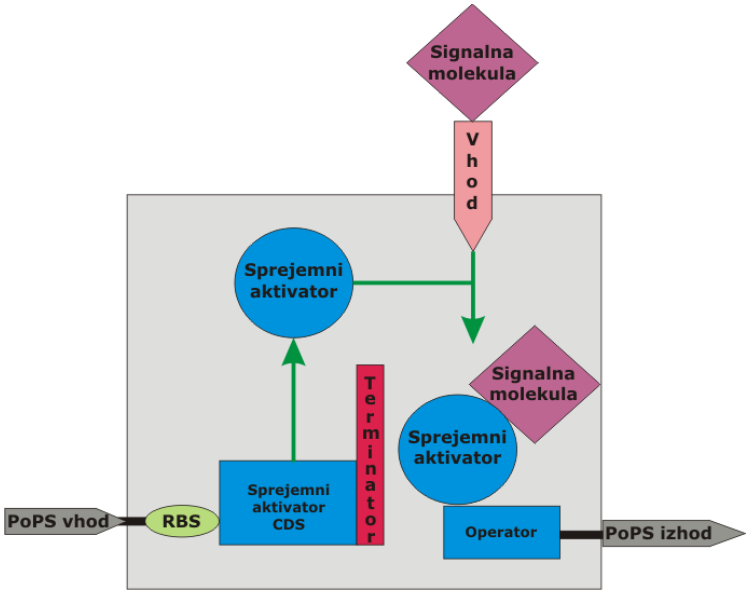
V skupini *naprav* iz skladišča biokock pa že lahko najdemo vrsto gradnikov imenovanih *signalizatorji*. Ti gradniki so strukture, ki definirajo oddajne in sprejemne celice. Posamezna struktura je sestavljena po principu, ki ga vidimo na sliki 6.2, če je sprejemna struktura in na sliki 6.3, če je oddajna struktura. Molekula, ki jo biokocke v skladišču oddajo oziroma sprejmejo, je 3OC6HSL.



Slika 6.1 Rezultat pasovne detekcije. Na slikah vidimo različne vzorce, ki so bili kreirani z oddajnimi in sprejemnimi celicami v netekočem mediju. Na slikah **a** in **b** sta dve oddajni celici, na sliki **c** so tri in na sliki **d** so štiri oddajne celice. Slika je rezultat projekta kreiranja vzorcev [19].



Slika 6.2 Koncept zgradbe oddajne strukture.



Slika 6.3 Koncept zgradbe sprejemne strukture.

Ko govorimo o komunikaciji med celicami se srečamo tudi s pojmom "quorum sensing". To je metoda medcelične komunikacije, ki se odziva na zasičenost populacije in v skladu z zasičenostjo regulira izražanje gena. S kombinacijo programirane celične smrti v *E.coli* bakteriji lahko ustvarimo sistem za kontrolo populacije in na ta način vzdržujemo velikost populacije. Če bi vse celice v populaciji delovale sinhrono, bi ob zasičenosti prišlo do hkratne smrti celotne populacije. Vendar se posamezne celice zaradi šuma drugače odzovejo in zato preživijo. Tako se vzpostavi ustrezna nizka zasičenost populacije.

6.2 Uporabnost večceličnih sistemov

Cilj snovanja sistemov, tako računalniških kot bioloških, je kreiranje novih, inteligentnih sistemov. Večcelične sisteme lahko povežemo z realizacijo sistemov za paralelno procesiranje, komunikacijo, celularnimi avtomati, učečimi avtomati, nevronskimi mrežami itd..

Celularni avtomati so diskretni dinamični sistemi, katerih delovanje je odvisno le od lokalnih interakcij. Predstavljeni so s prostorsko mrežo celic, pri čemer je vsaka celica v nekem stanju v določenem časovnem koraku. Dinamičnost sistema omogočajo pravila, ki definirajo stanje vsake celice v naslednjem časovnem koraku v odvisnosti od stanj celic v soseščini. Na ta način se oblikujejo skupine, ki branijo svojo odločitev in skupaj

prispevajo h končni klasifikaciji. Najpomembnejša značilnost celularnega avtomata je odsotnost človeka pri njegovem delovanju.

Ciljna uporaba celularnega avtomata je odvisna od algoritma, katerega uporabimo za spreminjanje stanj. Primer takšnega algoritma je evolucijski algoritem, ki temelji na načinu spreminjanja dednega zapisa živih organizmov tako, da se prilagodijo okolju. Operacije, ki se vršijo, so mutacija, rekombinacija in selekcija.

Glede na podobnost predstavitve celularnega avtomata s predstavitvijo večceličnih sistemov v netekočem mediju lahko rečemo, da je primeren za realizacijo s pomočjo sintezne bilogije. Enostavnost komunikacije, saj poteka le med sosednjimi celicami, bi lahko realizirali na način, kot je predstavljen v prejšnjem poglavju. Željen algoritem, ki bi ga naj biološki celularni algoritem izvajal, pa lahko izvedemo s pomočjo predstavljenih polnih naborov logičnih vrat.

Uspeh bi bil tudi realizacija nevronske mreže s pomočjo sintezne biologije in biokock. Nevronska mreža je način za obdelavo informacij, ki predstavlja matematični model človeških možganov. Obstaja več različnih nevronskih mrež, ki se med seboj razlikujejo po strukturi in po načinu učenja. Osnovni gradniki nevronskih mrež so nevroni, ki predstavljajo pragovno funkcijo. Slednja določa, kdaj bo nevron oddal signal na izhod. Vsak nevron ima določeno število obteženih vhodnih signalov in izhod. Nevroni si med seboj pošiljajo signale. Ko je vsota vhodnih signalov dovolj velika, se nevron vžge in odda signal na izhod.

Značilnost nevronskih mrež je, da v času učenja same izoblikujejo pravilo, ki povezuje izhodne in vhodne podatke. Ko je proces učenja končan, je nevronska mreža pripravljena sprejeti tudi primer problema s katerim se še ni srečala. Rešuje lahko tudi probleme za katere ne obstaja algoritem, ki bi po korakih vodil do rešitve.

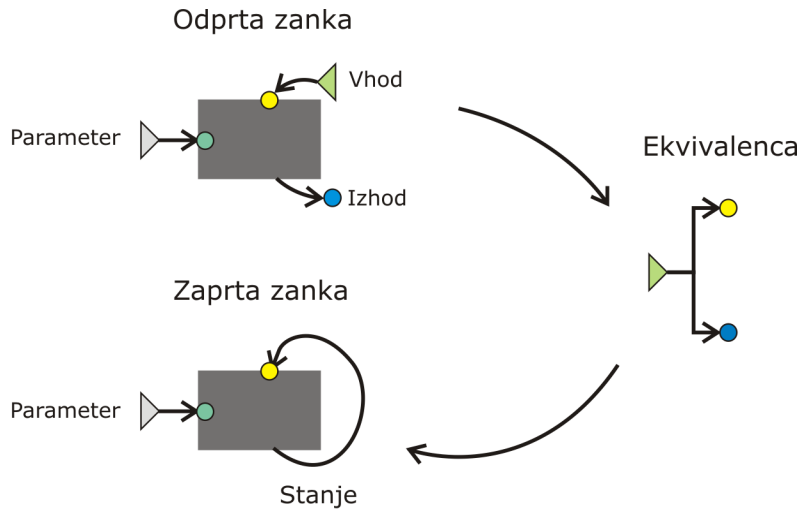
6.3 Primer

Zaradi finančnih in časovnih razlogov, se nismo mogli lotiti lastnega projekta, s katerim bi potrdili delovanje koncepta biokock. Zato si bomo za vpogled v nastajanje dejanskega sistema ogledali projekt skupine iz Indije (National Centre for Biological Sciences, Bangalore). Projekt je bil predstavljen na tekmovanju iGEM 2007 in nagrajen je bil z zlato medaljo.

6.3.1 Ideja projekta

Skupina je želela dokazati, da lahko z analizo, meritvami in karakterizacijo enostavnega sistema natančno predvidevamo delovanje kompleksnega sistema z enakimi parametri. S tem so želeli dokazati "bottom up" princip sestavljanja.

Izbran primer za karakterizacijo je bil enostaven sistem odprte zanke (angl. *open-loop*). Na podlagi rezultatov, ki so jih dobili pri meritvah delovanja sistema odprte zanke, so poskusili določiti delovanje kompleksnejšega sistema, kot je zaprta zanka (angl. *closed-loop*).



Slika 6.4 Ideja kako nastaviti model, da bo delovanje zaprte zanke predvidljivo.

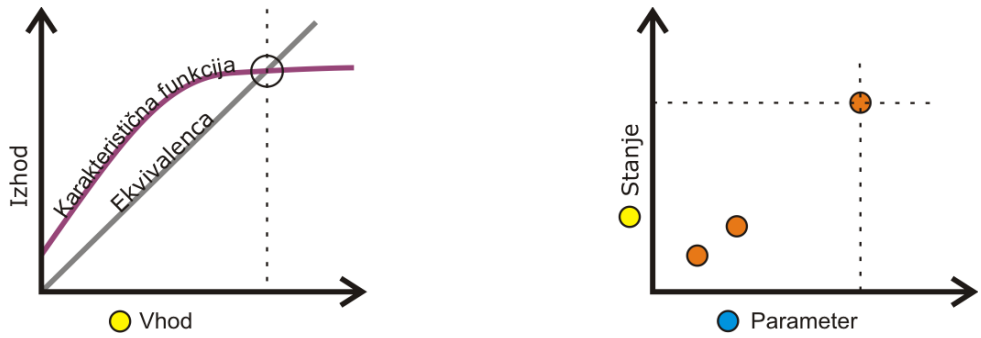
Na sliki 6.4, lahko vidimo teoretično zastavljen problem projekta.

Odprta zanka je nastavljena tako, da je koncentracija parametra fiksna. Ob spreminjanju koncentracije vhoda, pa so opazovali obnašanje na izhodu in s tem karakterizirali delovanje sistema odprte zanke.

Z ekvivalenco je definirana enakost vhoda in izhoda. S presečiščem karakteristične funkcije odprte zanke in funkcije ekvivalence so točno določili okolje v katerem dejansko pride do ekvivalence vhoda in izhoda.

Zaprt sistem dobimo s povezavo izhoda na vhod. Zaprta zanka je vedno le v enem od možnih stanj. Da bi določili več možnih stanj so spreminjali vrednost parametra in merili odziv zaprte zanke.

Na sliki 6.5 sta graf presečišča funkcije ekvivalence in karakteristične funkcije ter graf stanj ob spreminjanju parametra.



Slika 6.5 Na prvi sliki z leve je graf, ki ponazarja presečišče karakteristične funkcije in funkcije ekvivalence ob fiksnem parametru. Na drugi sliki z leve je graf stanj ob spremenljivem parametru in upoševanju ekvivalence.

Meritev presečišč ob spremenljivim parametrom, pripomore pri napovedovanju delovanja sistema zaprte zanke s karakteristiko sistema odprte zanke.

6.3.2 Eksperiment

Za aktivacijo in delovanje obeh zank so uporabljeni trije regulatorni moduli:

- IPTG z inducibilnim lac promotorjem,
- aTc z inducibilnim tet promotorjem,
- AHL z inducibilnim lux promotorjem.

Meritve so opravili na treh konstruktih in sicer na odprti zanki, zaprti zanki in na ekvivalenci.

Odprta zanka

Sistem odprte zanke je sestavljen iz dveh struktur in sicer iz oddajne strukture in sprejemne strukture. Vsaka od struktur je v svoji celici.

Oddajno strukturo so poimenovali "Sen-TIC" in jo lahko najdemo v skladišču pod oznako *BBa_I726041*. Biokocka oddajne strukture ima vlogo parametra omenjenega v 6.3.1.

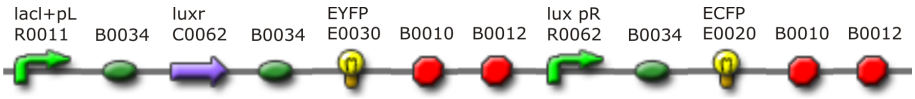
Sestavljena je iz tetR represibilnega promotorja, dveh mest vezave ribosoma, AHL autoinducerja, modro-zelenega reporterja in dveh terminatorjev, kar je skupaj 1551 baznih parov. Na sliki 6.6 sta zaporedje vezave in oznake izbranih gradnikov omenjene biokocke.

Slika 6.6 Biokocka z oznako *BBa_I726041*. aTc vsili izražanje LuxI proteina in tega ugotovimo z modro-zelenim reporterjem. Sintetizira se AHL, ki se odda in uporabi v kombinaciji z promotorjem LucpR v katerikoli AHL sprejemni napravi.

Sprejemno strukturo so poimenovali "Rec-LRY.RC" in najdemo jo lahko v skladišču pod oznako *BBa_I726081*. Na sliki 6.7 sta zaporedje vezave in oznake izbranih gradnikov, ki tvorijo omenjeno biokocko. Velikost te biokocke je 2697 baznih parov.

Sestavljena je iz dveh delov in sicer iz zaporedja za vhod in iz zaporedja, ki generira izhod. Vhod se regulira z induciranjem lac proteina na IPTG, ki sproži sintezo luxR proteina. LuxR protein ob sprejetem AHL iz oddajne strukture sproži izraz reporterskega modrozelenega proteina.

Na sliki 6.8 je celotna zgradba odprte zanke.



Slika 6.7 Sprejemna struktura z vhodom in izhodom. Na vhodu IPTG vsili izražanje luxR proteina katerega ugotovimo z rumenim reporterjem. Izhodna stran se začne z regulatornim elementom lux pR, ki ob sprejemu AHL vsili izraz modro-zelenega reporterja.

Zaprta zanka

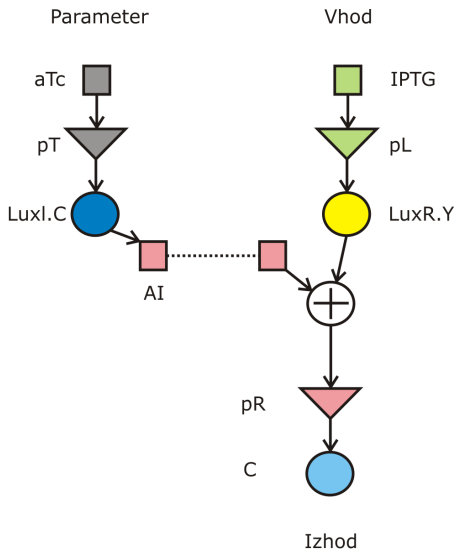
V zaprti zanki je uporabljena identična oddajna biokocka "Sen-TIC", kot v odprti zanki, medtem ko je na sprejemni strani večina vhodnega zaporedja neuporabljenega.

Sprejemni del v zaprti zanki je sestavljen le iz promotorja, ki vsili izražanje proteina luxR ob vsebnosti AHL, ki ga dobi iz oddajne celice. Velikost biokocke je 1748 baznih parov. Zaporedje in oznake posameznih gradnikov sprejemne biokocke *BBa_I726071*, ki so jo poimenovali "Rec-RRY", lahko vidimo na sliki 6.9.

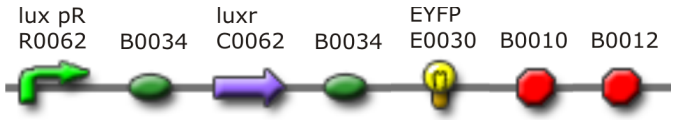
Celotno zgrabo zaprte zanke lahko vidimo na sliki 6.10.

Ekvivalenca

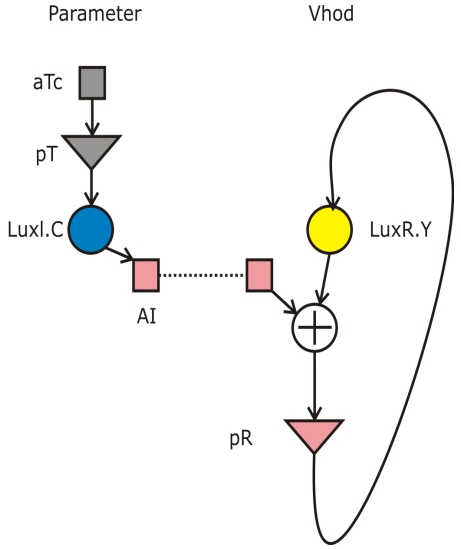
Ker morata biti vhod in izhod odprte zanke enaka, je potrebno primerjati dve strukturi. Na izhodu je to struktura, ki so jo poimenovali "Trc-LC" in na vhodu je to struktura "Trc-LRY". Ekvivalenca se določi na podlagi primerjave teh dveh struktur.



Slika 6.8 Konstrukt oddajne in sprejemene celice. Autoinducer, AHL se sintetizira v oddajni celici in pošlje. Ob sprejemu AHL v sprejemni celici in ob vsebovanosti luxR proteina se sintetizira fluorescentni protein C na izhodu.



Slika 6.9 V prisotnosti AHL, luxR protein prične sam sebe sintetizirati na promotorju pR.

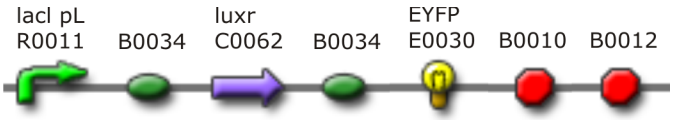


Slika 6.10 Vidimo lahko, da je zgradba oddajne celice popolnoma enaka kot v odprti zanki. Razlika je v sprejemni celici, kjer promotor, ki je prej sprožil sintezo reporterja, zdaj sproži sintezo luxR proteina.

Strukturo "Trc-LC" najdemo v skladišču pod oznako *BBa_I726005*. Ta biokocka je sestavljna iz IPTG promotorja, ki sproži sintezo modro-zelenega fluorescentnega proteina. Zaporedje gradnikov in njihove oznake lahko vidimo na sliki 6.11. Velikost biokocke je 941 baznih parov.

Slika 6.11 Biokocka z oznako *BBa_I726005* se uporablja za kalibriranje fluorescentnega branja. V našem primeru je uporabljena za ugotavljanje ekvivalence.

Struktura "Trc-LRY" je del sprejemne celice odprte zanke. IPTG promotor sproži sintezo luxR proteina, ki se uporabi za indukcijo sinteze reporterskega proteina ob sprejemu AHL iz oddajne celice. Velikost biokocke je pravtako 1748 baznih parov, zaporedje, ki jo sestavlja, pa lahko vidimo na sliki 6.12.



Slika 6.12 Biokocka z oznako *BBa_I726061* je del sprejemne celice odprte zanke. IPTG promotor sproži sintezo luxR proteina, ki se uporabi za vsilitev sinteze reporterskega proteina ob sprejemu AHL iz oddajne celice.

6.3.3 Rezultati

Meritve so potekale na treh sestavnih napravah sistema odprte zanke. Te naprave so:

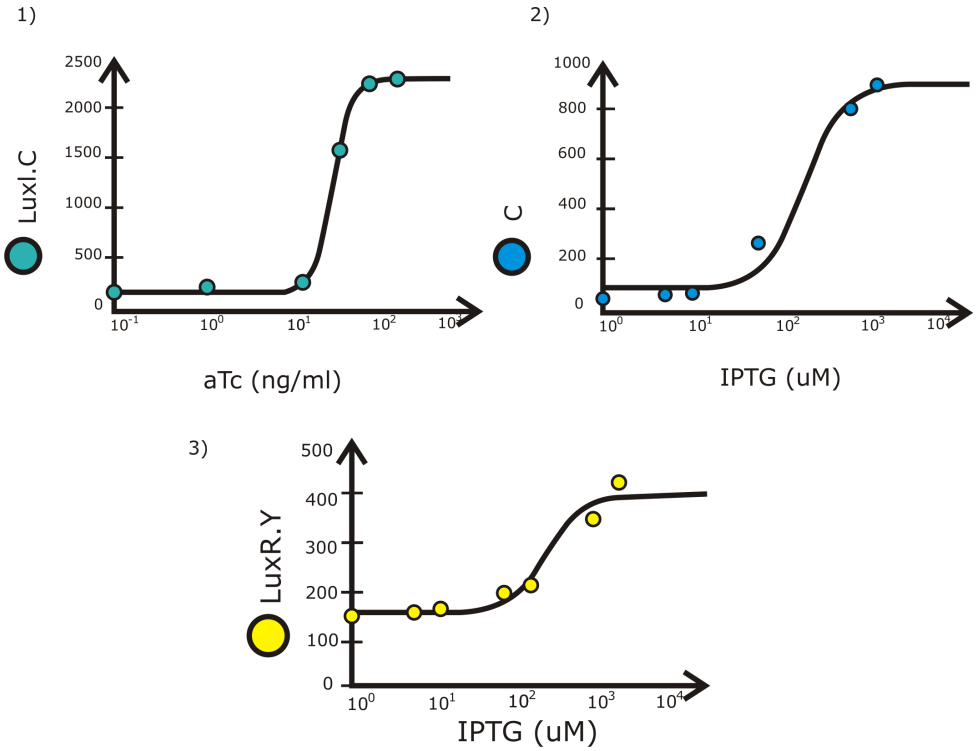
- naprava v oddajni celici,
- naprava za ugotavljanje ekvivalence na izhodu,
- naprava za ugotavljanje ekvivalence na vhodu.

Celice z oddajno napravo so gojili v mediju LB in jih nato prestavili v medij Glu-M9, ki je vseboval željeno koncentracijo aTc. V mediju Glu-M9 so izvajale svojo prenosno funkcijo in rasle dvanajst ur.

Meritev je potekala na množici približno 500 celic, kjer so merili fluorescenco na enoto celice. Vrednost fluorescence odraža intenzivnost sintetiziranja AHL. Meritev so v enakih pogojih večkrat ponovili. Povprečje rezultatov večkratnih meritev lahko vidimo na prvem grafu na sliki 6.13.

$$y = a_0 + a_1 \frac{x^n}{K^n + x^n}, \tag{6.1}$$

Funkcija na prvem grafu na sliki 6.13 se sklada s sigmoidno funkcijo oziroma enačbo (6.1), kjer so vrednosti $a_0 = 212$, $a_1 = 2027$, $K = 21.5$ in $n = 4.3$.



Slika 6.13 1) Graf rezultata meritev fluorescence in s tem intenzivnosti izražanja luxI proteina za sintezo AHL ob različnih koncentracijah aTc. 2) Graf rezultata meritev fluorescence modro-zelenega reporterja na izhodu. 3) Graf rezultata meritev fluorescence rumenega reporterja in s tem tudi izražanje lacI represorja ob izbrani količini IPTG proteina.

V sprejemnih celicah sta napravi za meritev intenzivnosti fluorescence na izhodu in izhodu. Celice z obema napravama so bile vzgojene v LB mediju in nato prestaviljene v medij Glu-M9, ki je vseboval željeno koncentracijo IPTG-ja. V mediju Glu-M9 so izvajale svojo prenosno funkcijo in rasle dvanajst ur.

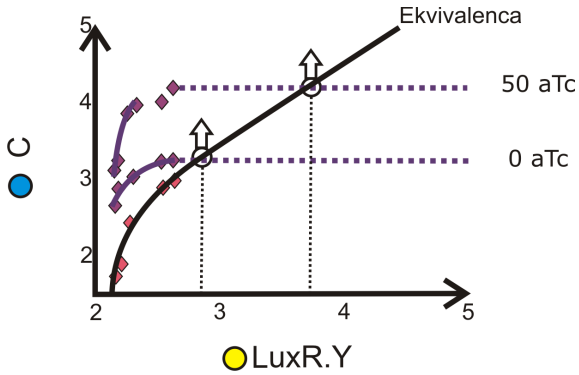
Meritev intenzivnosti izhoda je potekala na množici približno 500 celic, kjer so merili fluorescenco na enoto celice. Meritev so v enakih pogojih večkrat ponovili. Povprečje rezultatov večkratnih meritev lahko vidimo na drugem grafu na sliki 6.13.

Funkcijo na drugem grafu na sliki 6.13 lahko opišemo z enačbo (6.1), kjer so vrednosti $a_0 = 96$, $a_1 = 818$, $K = 158$ in $n = 2$.

Meritev intenzivnosti vhoda je potekala na istih množicah 500 celic, kjer so merili fluorescenco na enoto celice. Povprečje rezultatov večkratnih meritev vhoda lahko vidimo na tretjem grafu na sliki 6.13.

Funkcijo na tretjem grafu na sliki 6.13 lahko opišemo z enačbo (6.1), kjer so vrednosti $a_0 = 151$, $a_1 = 241$, $K = 158$ in $n = 2$.

Ekvivalenco med vhodom in izhodom so ugotovili s primerjanjem grafov vhoda in izhoda sistema odprte zanke ob izbrani koncentraciji aTC in dobili graf na sliki 6.14. Z grafa so razbrali koliko mora biti koncentracija aTc za posamezna stanja zaprte zanke. S tem so definirali vse potrebne vrednosti s katerimi je delovanje zaprte zanke uravnovešeno in predvidljivo.



Slika 6.14 Primerjava fluorescence vhoda in izhoda ob izbrani koncentraciji aTc kot parametra. S tem so dokazali, da je lahko zaprta zanka v dveh stanjih in sicer pri koncentraciji aTc = 0 in aTc = 50.

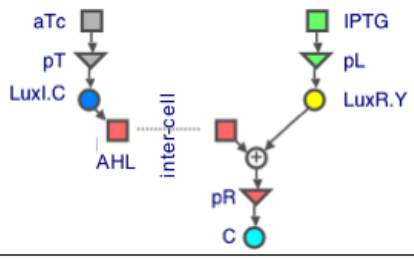
Rezultati meritev in karakteristike posameznih fragmentov v oddajni in sprejemni celici so predstavljeni v obliki kataloga podatkov. Kot primer si poglejmo sliko 6.15, na kateri je del kataloga, ki opisuje karakteristike sistema odprte zanke, poimenovanega "Sen-TIC+Rec-LRY.RC".

V katalogu so podatki o organizmu v katerem je sistem realiziran, o strukturi sistema, poteku realizacije in karakteristiki. V našem primeru lahko vidimo, da je gostiteljski organizem *E.coli* in da za realizacijo potrebujemo dve celici, ker je sistem sestavljen iz dveh delov. Opis v katalogu nam tudi pove kakšna je funkcija sistema, kako se ga uporablja in kakšni so pogoji za kompatibilnost z ostalimi konstrukti. Za vsak posamezen del imamo podane proteine in promotorje, ki jih potrebujemo ter protokol po katerem se sistem dejansko realizira. V izbranem kataloškem opisu je tudi graf, ki ponazarja odvisnost fluorescence od induciranih proteinov IPTG in aTc v sistemu odprte zanke.

[Sen-TIC+Rec-LRY.RC] Open-Loop System

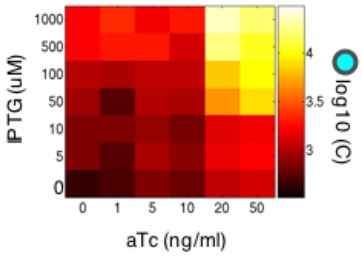
Author: Mukund Thattai (thattai@ncbs.res.in) Date: October 16th, 2007

Schematic



Characteristics

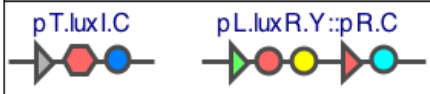
Protocol: We grew sender cells overnight in LB, then transferred them to Glu-M9 minimal medium containing the desired final concentration of aTc, and allowed them to grow for ~12h, until they reached OD600=0.2. We then removed the sender cells by filtration, retaining the medium (which now contained AHL) and replenishing it with an equal volume of fresh Glu-M9 containing double the standard glucose concentration, and the desired final concentration of IPTG. Receiver cells, which had previously been grown overnight in LB, were added to this medium. These cells were allowed to grow for ~12h, and harvested at OD600 < 0.1. Both sender and receiver cells were concentrated by centrifugation, and imaged on an epifluorescence microscope. We calculated the fluorescence per unit area of single cells, obtaining data from ~500 cells for each run. We then averaged these values in log space to obtain the final estimate of protein expression. The aTc mesh was [0 1 5 10 20 50] ng/ml, and the IPTG mesh was [0 5 10 50 100 500 1000] uM. The results shown below are the average of two replicates.



Measurements and analysis were carried out by members of the NCBS iGEM 2007 team: Kiran, Krishna, Mukund, Navneet, Nilesh, Senthil, Shashanka, Sugat, Sushant, Varun, Vini, Vivek

Host: Ecoli K12Z1

Parts: [Sen-TIC+Rec-LRY.RC]



Small molecules

- IPTG: Isopropyl β-D-1-thiogalactopyranoside.
- aTc: anhydrotetracycline.
- AHL: acyl homoserine lactone.

Promoters

- p1: Constitutive LacIq promoter.
- p2: Constitutive N25 promoter.
- pT: Tet promoter
- pL: Lac promoter.
- pR: *V. fischeri* LuxpR promoter.

Proteins

- LacI: Lac repressor, negative regulator of pL.
- TetR: Tet repressor, negative regulator of pT.
- LuxI: *V. fischeri* LuxI protein, synthesizes AHL.
- LuxR: *V. fischeri* LuxR protein, positive regulator of pR.
- C: Cyan fluorescent protein
- Y: Yellow fluorescent protein

Description

Sender: aTc drives the expression of the LuxI protein (monitored using polycistronic CFP). This in turn drives the synthesis of AHL, which is exported. Receiver: IPTG drives the expression of the LuxR protein (monitored using polycistronic YFP). LuxR, when bound imported AHL, drives the expression of CFP.

Usage and compatibility

This is a stand-alone system.

Slika 6.15 Primer opisa lastnosti odprte zanke v katalogu.

6.4 Povzetek poglavja

V tem poglavju smo na celično dinamiko pogledali s porostorskega vidika in ugotovili, da če želimo govoriti o dinamiki med celicami potrebujemo način medcelične komunikacije. Spoznali smo strukturi, ki omogočata oddajanje in sprejemanje signalov in izzive s katerimi se soočimo, če želimo graditi zanesljiv večcelični sistem. Navedeno je le nekaj možnih sistemov, kjer bi nam prišel prav večcelični sistem, vsekakor pa ni omejen le na te sisteme.

Na primeru smo si ogledali kako poteka gradnja sistema od ideje do realizacije in pokazali "bottom-up" princip sestavljanja.

7 Zmogljivost in varnost

7.1 Zmogljivost

Namen iskanja alternativ današnjemu načinu procesiranja na siliciju je zmanjšati velikost naprav in povečati hitrost procesiranja. Zanesljivost pa mora ostati enaka ali boljša od konvencionalnih načinov.

Sama standardizacija gradnikov loči proces konstruiranja posameznih gradnikov od procesa sestavljanja teh gradnikov v funkcionalno celoto. S tem je omogočena ponovitev eksperimenta in tudi razvijanje daljše in kompleksnejše biokocke, kjer pa princip konstruiranja ostaja enak. Vsak projekt je biokocka, ki se lahko uporabi pri sestavi nove biokocke. Biokocke pa se ponujajo kot vmesnik za žive organizme in s tem dobimo žive organizme, ki izvajajo predviden algoritem oziroma funkcijo.

Nad eno verigo DNK lahko hkrati deluje več encimov na različnih segmentih, kar lahko primerjamo s paralelnim izvajanjem operacij. Z upoštevanjem te lastnosti je hitrost izvajanja operacij lahko večja, kot pri današnjih superračunalnikih [20, 21].

Prednostna značilnost same DNK verige je potencial za shranjevanje podakov. DNK molekula potrebuje le en kubični nanometer za shrambo enega bita informacije, kar

pomeni, da lahko z DNK dosežemo visoko gostoto zapisa. V primerjavi z konvencionalnimi računalniki, ki lahko shranijo informacijo le v binarni obliki, (0 ali 1), lahko v DNK uporabimo štiri različne vrednosti. DNK sestavljajo štiri baze (A,T,C in G), s čimer dobimo 64 diskretnih kombinacij, s štirimi znaki v binarni obliki pa lahko dobimo le 6 diskretnih kombinacij.

Učinkvitost pri izrabi energije je tudi pomembna lastnost. Pri tem bomo omenili samoreplikacijo in samoreprodukcijo. S samoreplikacijo naredimo en prototip in imamo čez noč lahko več milijonov enakih celic.

Velika slabost bioloških računalnikov je ta, da potrebujemo veliko časa za postavitev in zagon sistema. Za postavitev enostavnega eksperimenta potrebujemo tudi do več dni. Naslednja pomembna slabost je ta, da so biološki sistemi izredno občutljivi na šume in so s tem tudi bolj nagnjeni k napakam.

Omejitev s katero se srečamo pri sestavljanju logičnih struktur je prostorska vidnost med posameznimi logičnimi vrati. Vhodno\izhodni signali logičnih vrat v biološkem sistemu niso izolirani, tako da bi signal potoval samo do željenega cilja, ampak prihaja do difuzije izraženega gena in s tem do presluha. Iz tega razloga moramo za vsaka posamezna logična vrata uporabiti drug transkripcijski faktor, kar omeji število logičnih vrat, ki jih lahko v nekem biološkem sistemu uporabimo.

Ker smo šele na začetku raziskovanja in razvoja bioloških sistemov, lahko v prihodnosti pričakujemo spremembe v tehniki konstruiranja sistemov zaradi boljše učinkovitosti in s tem zanesljivosti živih sistemov.

7.2 Varnost

Aplikativni potencial tehnologije rekombiniranja DNK je izredno visok. Področij kjer so genetsko modificirani organizmi uporabni je veliko, najbolj pa izstopajo v konstruiranju novih zdravil, biogorivih, konstrukciji novih naprav in v bioloških zdravilih, ki bi v prihodnosti lahko zdravila najbolj smrtonosni bolezni kot sta rak in virus HIV.

Visoka aplikativnost se izraža v povečanem raziskovanju in številu eksperimentov v labaratorijih.

Z vidika varnosti ali nevarnosti sintezne biologije bomo omenili probleme *bio-varnosti*, *zlorabe* in *etike*.

V primeru manjše nesreče v laboratoriju lahko pomislimo na več različnih scenari-

jev. Okužba laboratorijskega osebja s sintetičnimi virusi in mikroorganizmi je lahko prvi rezultat nesreče. Zaradi prepozne ugotovitve okužnosti osebja lahko virusi in mikroorganizmi nepričakovano zapustijo *in vitro* (glej razlago v dodatku A) okolje in razširijo novo bolezen na živali in ljudi. Bakterija se lahko zelo hitro množi in s tem nastajajo kopije modificirane bakterije. Najhuje bi bilo v primeru pobegle patogene bakterije, ker lahko pride do epidemije ali še huje do pandemije, če narava nima razvitega ustreznega obrambnega mehanizma.

Tudi če bakterija ne ogroža višjih življenskih oblik neposredno, lahko uniči ostale bakterije, ki so nam potrebne za življenje in nas ogrozijo posredno.

Naslednja omenjena nevarnost je *zloraba*. V namen nadzora raziskovanja v sintezni biologiji je bilo osnovanih veliko različnih nadzornih teles. Zaradi nevarnosti uporabe rezultatov sintezne biologije za izdelavo bio orožja, je glavni cilj nadzornikov ugotavljanje zlorabe dokumentacije raziskav.

In še zadnji, vendar nič manj pomemben vidik varnosti je *problem etike*. Etične probleme s katerim se bomo srečali mi, ki smo sestavili v celoti sintetičen biološki sistem, lahko zapišemo v obliki naslednjih vprašanj.

- Kakšne so posledice posredovanja v naravno življenje?
- Kakšna je definicija življenja?
- Kje je meja, do koder je umetno sintetiziran biološki sistem še vedno živeč material?
- Po čem se umetno sintetizirana celica razlikuje od stroja?
- Ali so umetno sintetizirani organizmi vredni zaščite tako kot naravni organizmi?
- itd.

Ključni pojmi, ki so povezani z varnostjo sintezne biologije, so avtonomnost, pravičnost in naravnost. Pri ustvarjanju umetnega biološkega sistema moramo razmišljati o posledicah, ki jih lahko biološki sistem pusti pri izdelovalcu in pri uporabniku.

8 Zaključek

Tehnologija rekombinantne DNK je v medicini prisotna že vrsto let, kot način oziroma ideja zdravljenja težkih bolezni. Da bi takšno zdravilo delovalo, morajo biti vsi koraki delovanja determinirani, kar je v bistvu izvajanje algoritma.

Koncept biokock je bil v našem primeru temeljna opora za razumevanje delovanja in delitve DNK zaporedij na funkcije, katere lahko uporabimo pri gradnji kompleksnih sistemov. Sistematiziranost skladišča biokock je idealna za začetniško seznanjanje in vpogled v funkcije najmanjših funkcionalov DNK, vendar pomanjkljiva dokumentacija gradnikov, ki so bistveni, zelo oslabi njihovo funkcionalno rabo. Nadaljevanje tega dela, bo torej potrebovalo svoje skladišče gradnikov z dokumentacijo, ki jo lahko primerjamo z dokumentacijo elektronskih elementov. Večina boljših svetovnih univerz že ima lastno zalogo funkcionalnih DNK gradnikov in z aktivnim sodelovanjem profesorjev in študentov v raziskovanju jo hkrati tudi povečujejo.

V nalogi smo predstavili elementarne DNK gradnike ali biokocke, s katerimi lahko sestavimo funkcionalne naprave in jih v skladu z željeno funkcijo uporabimo za sprejemanje in oddajanje signalov, odločanje, pomnjenje, prikazovanje in s tem gradnjo sistemov, ki

bodo delovali po željenem končnem algoritmu. Hkrati je predstavljen *bottom-up* pristop realizacije sistema, kjer so osnovni elementi do potankosti analizirani in dokumentirani. Z medsebojnim povezovanjem in analiziranjem teh, pridemo do večjih in kompleksnejših sistemov.

Postopek od ideje do dejanske realizacije smo opisali v obliki algoritma, v katerem smo predstavili vse korake s potencialnimi problemi in rešitvami. Z modeli osnovnih logičnih operatorjev, ki smo jih definirali kot poln nabor, smo predstavili idejo o gradnji računalniških struktur. Obravnavali smo tudi mnogocelične sisteme, kjer smo definirali osnovno oddajno in sprejemno strukturo, ki sta bistveni za večje sinhronizirane sisteme.

Primarni cilj nadaljnjega dela je realizacija računalniško biološkega sistema, sestavljenega iz DNK gradnikov, ki bo zanesljivo izvajal kakšnega od procesov, kot jih poznamo v računalništvu. Glede na zelo aktivno raziskavanje dinamike v fluidih, v katerih so lahko realizirani biološki sistemi, bi lahko bil eden od možnih nadaljevanj dela tudi realizacija sistema v različnih fluidih v nanoceveh (angl. *nanotubes*).

Dodatek A

Kratka razlaga omenjenih bioloških pojmov

Razlage bioloških pojmov so povzete po [22] in [9], kjer lahko najdemo tudi bolj natančne pojasnitve lastnosti in funkcij omenjenih pojmov.

Bralni okvir, (angl. *reading frame*) je sekvenca kodonov (skupin treh zaporednih nukleotidov), v okviru katerih lahko pride do prevajanja (translacije). Segment mRNK, na primer: 5'AUCCGAGGC3' se lahko prevede v treh možnih okvirih: 5'AUC..., 5'UCC... ali 5'CCG..., kar je odvisno od lokacije začetnega kodona.

De novo, je latinska fraza, ki pomeni znova, od začetka. Sinteza *de novo*, pomeni sinteza kompleksnih molekul iz enostavne molekule kot je aminokislina ali sladkor.

Ekson (angl. *exon*) je del z introni razdeljenega gena, ki je del nastajajoče mRNK v celičnem jedru in kodirajoče RNK v citoplazmi ter je na ta način vključen v transkript. Eksoni na splošno zasedajo tri različna področja genov, ki kodirajo proteine. Prvo področje, ki ni translirano v protein, signalizira začetek RNK transkripcije in ima sekvence, ki usmerjajo mRNK k ribosomom, kjer prihaja do sinteze proteinov. Eksoni drugega po-

dročja imajo informacije, ki so translirane (prevedene) v sekvenco aminokislin proteina. Eksoni tretjega področja pa so transkribirani v tisti del mRNK, ki signalizira končanje translacije ter pripenjanje poliadenilnega "repa".

Encim, (angl. *enzyme*) je protein, ki katalizira specifično kemično reakcijo, ne da bi bil v procesu reakcije porabljen.

Fenotip, (angl. *phenotype*) so morfološke, fiziološke ali opažane lastnosti nekega organizma, ki jih določa genotip v povezavi z okoljem, je torej rezultat delovanja genetskih in okoliških dejavnikov.

Genom, (angl. *genom*) je skupek vseh genov v celici.

Genotip (angl. *genotype*), je skupek vseh podedovanih lastnosti nekega organizma, razpoznavnega po njegovi zunanjosti ali fenotipu.

In silico, je besedna zveza, ki se zlasti v naravoslovnem izrazoslovju nanaša na procese in poskuse, ki virtualno potekajo računalniku. Naprimer simulacija procesov.

In vitro, je besedna zveza, ki se zlasti v naravoslovnem izrazoslovju nanaša na procese in poskuse, ki potekajo v nadzorovanem okolju zunaj živega organizma.

In vivo, je izraz, ki se zlasti v naravoslovnem izrazoslovju uporablja za procese, ki potekajo v živem organizmu.

Intron (angl. *intron*) je segment v genih, ki so z njim razdeljeni. Intron se sicer prepisuje v jedrno RNK, a se naknadno iz nje izloči in degradira.

Kazaminska kislina je mešanica amino kislin, ki se uporabaljajo kot dodatek mediju gojenja (AB). Proizvaja se iz mleka (caseus).

Kodon, (angl. *codon*) je nukleotidni trojček ali zaporedje treh nukleotidov v mRNK, ki določa, katera aminokislina bo vstavljena na določeno mesto v nastajajočem peptidu s pomočjo translacije (prevajanja).

Kromatin, (angl. *chromatin*), tudi jedrovina, je sestavljena iz nukleinskih kislin (DNK, RNK) ter proteinov, ki predstavljajo kromosome evkariontov.

Kromosom, (angl. *chromosome*) je pri prokariontih krožna molekula DNK, ki vsebuje celotno genetsko informacijo, potrebno za življenje celice. V evkariontskem celičnem jedru pa so to nitaste tvorbe iz kromatina in so nosilci genetskih informacij, ki so zapisane v linearno razporejenih genih.

Lac gen, je potreben pri transportu in metabolizmu laktoze v bakteriji *E.coli* in v nekaterih drugih bakterijah. Za reguliranje Lac gena se upošteva vsebovanost glukoze in laktoze. Sestavljen je iz treh sosednjih genov, promoter, terminator in operator.

Mejoza, (angl. *meiosis*) je proces delitve spolnih celic, ki je pri organizmih s spolnim razmnoževanjem namenjen razpolovitvi števila kromosomov v gametah, ki se s svojimi jedri spojijo, čemur pravimo singamija. Proces poteka tako, da se kromosomi najprej podvojijo, nato pa se celica dvakrat razdeli.

Mitoza, (angl. *mitosis*) je delitev celičnega jedra. Poteka v štirih značilnih fazah (profaza, metafaza, anafaza in telofaza) in rezultat mitoze sta dve hčerinski celici, ki imata enaki jedri in enako količino citoplazme.

Odprt bralni okvir, (*open reading frame*) je del v molekuli DNK, kjer zaporedni nukleotidni trojčki v kodonih določajo aminokisline in niso prekinjani s stop kodoni.

Operator je segment DNK na katerega se veže represor. Vzajemno lahko deluje z represorjem in na ta način nadzoruje delovanje odgovarjajočega gena.

Polimeraza, (angl. *polymerase*) se imenujejo encimi, ki katalizirajo nastajanje DNK in RNK molekul iz dezoksiribonukleotidov in ribonukleotidov (DNK- ter RNK-polimeraze).

Stoihiometrija, (angl. *stoichiometry*) je kvantitativen izračun relacije med reaktanti in produkti v kemičnih reakcijah.

Slike

2.1 Simboli biokock. 7
2.2 Analogija sintezne biologije. 10

3.1 Kloniranje plazmida. 18
3.2 Rezanje biokock in sestavljanje nove biokocke. 19

5.1 Primer zgradbe vhodnih naprav s prefiksno dodanim konvertorjem. 33
5.2 Biokocka z oznako BBa_J09855, konvertor AHL v PoPS. 34
5.3 Biokocka z oznako BBa_I732083, konvertor aTc v PoPS. 34
5.4 Koncept biološke zgradbe logičnih vrat NOT. 40
5.5 Koncept biološke zgradbe logičnih vrat AND. 41
5.6 Koncept biološke zgradbe logičnih vrat OR. 42
5.7 Koncept biološke zgradbe logičnih vrat NOR. 43
5.8 Koncept biološke zgradbe logičnih vrat NAND. 44

6.1 Rezultati sistema pasovne detekcije. 51
6.2 Koncept zgradbe oddajne strukture. 51
6.3 Koncept zgradbe sprejemne strukture. 52
6.4 Ideja kako nastaviti model, da bo delovanje zaprte zanke predvidljivo. . . 54
6.5 Graf presečišča karakteristične in ekvivalenčne funkcije ter graf stanj. . . . 55
6.6 Biokocka z oznako *BBa_I726041*. 56
6.7 Sprejemna struktura z vhodom in izhodom. 56
6.8 Konstrukt oddajne in sprejemene celice. 57
6.9 Sprejemni del zaprte zanke. 57
6.10 Zgradba zaprte zanke. 57
6.11 Biokocka za kalibriranje fluorescentnega branja. 58

6.12 Del sprejemne celice odprte zanke. 58
6.13 Grafi rezultatov meritev fluorescence. 59
6.14 Primerjava fluorescence vhoda in izhoda. 60
6.15 Primer opisa lastnosti odprte zanke v katalogu. 61

Literatura

[1] M. Dubash, Moore's law is dead, says gordon moore, Techworld.

[2] I. L. Bajec, N. Zimic, M. Mraz, Towards the bottom-up concept: Extended quantum-dot cellular automata, Microelectronic engineering vol. 83 (4/9) (2006) pp. 1826–1829.

[3] I. L. Bajec, N. Zimic, M. Mraz, The ternary quantum-dot cell and ternary logic, Nanotechnology 17 (8) (2006) 1937–1942.

[4] P. Pečar, M. Mraz, N. Zimic, M. Janež, I. L. Bajec, Solving the ternary quantum-dot cellular automata logic gate problem by means of adiabatic switching, Japanese Journal of Applied Physics 47 (6) (2008) 5000–5006.

[5] MIT, The registry of standard biological parts, `http://partsregistry.org/`.

[6] The biobrick fundation, `http://openwetware.org/wiki/The_BioBricks_Foundation`.

[7] M. Marchisio, J. Stelling, Computational design of synthetic gene cirsuits with composable parts, Bioinformatics.

[8] E. Andrianantoandro, S. Basu, D. K. Karig, R. Weiss, Synthetic biology: new engineering rules for emerging discipline, Molecular System Biology 2 (2006) pp. 1–14.

[9] B. R. Glick, J. J. Pasternack, Molecular biotechnology: principles and applications of recombinant DNA, 3rd Edition, ASM Press, Washington, 2003.

[10] P. Marguet, F. Balagadde, C. Tan, L. You, Biology by design: reduction and synthesis of cellular components and behaviour, Journal of the Royal Society Interface vol. 4 (2007) pp. 607–623.

[11] L. You, Toward computational systems biology, Cell Biochemistry and Biophysics vol. 40 (2004) pp. 167–184.

[12] P. R. Patnaik, External, extrinsic and intrinsic noise in cellular systems: analogies andimplications for protein synthesis, Biotechnology and Molecular Biology Review vol. 1 (2006) pp. 121–127.

[13] Z. Aydin, Y. Altunbasak, H. Erdogan, Bayesian protein secondary structure prediction with near-optimal segmentations, IEEE Transactions and Signal Processing vol. 55 (7) (2007) pp. 3512–3525.

[14] X. Xia, Bioinformatics and the Cell: Modern Computational Approaches in Genomics, Proteomics and Transcriptomics, Springer Science, New York, 2007.

[15] A. J. Viterbi, A personal history of the viterbi algorithm, Signal Processing Magazine, IEEE vol. 23 (4) (2006) pp. 120–142.

[16] A. Levskaya, A. Chevalier, J. Tabor, Z. Simpson, L. Lavery, M. Levy, E. Davidson, A.Scouras, A. Ellington, E. Marcotte, C.A.Voigt, Synthetic biology engineering escherichia coli to see light, Nature vol. 438 (2005) pp. 441–442.

[17] C. Anderson, C. A. Voigt, A. P. Arkin, Environmental signal integration by a modular and gate, Molecular Systems Biology vol. 3 (133).

[18] F. K. Balagadde, H. Song, J. Ozaki, C. H. Collins, M. Barnet, F. H. Arnold, S. R. Quake, L. You, A synthetic escherichia coli predator–prey ecosystem, Molecular Systems Biology vol. 4 (187).

[19] S. Basu, Y. Gerchman, C. H. Collins, F. H. Arnold, R. Weiss, A synthetic multicellular system for programmed pattern formation, Nature vol. 434 (2005) pp. 1130–1134.

[20] J. Parker, Computing with dna, EMBO reports vol. 4 (1) (2003) pp. 7–10.

[21] Top 500 supercomputer sites, `http://www.top500.org/`.

[22] F. Ločniškar, Katalog znanj, `http://www.bfro.uni-lj.si/zoo/publikacije/katalogznanj/`.

IZJAVA

Izjavljam, da sem diplomsko nalogo izdelala samostojno pod vodstvom mentorja izr. prof. dr. Mihe Mraza in somentorja prof. dr. Romana Jerale. Izkazano pomoč drugih sodelavcev sem v celoti navedla v zahvali.

Ljubljana, September 2008 Božidara Cvetković