Formal Modeling and Analysis of Pancreatic Cancer Microenvironment

Qinsi Wang

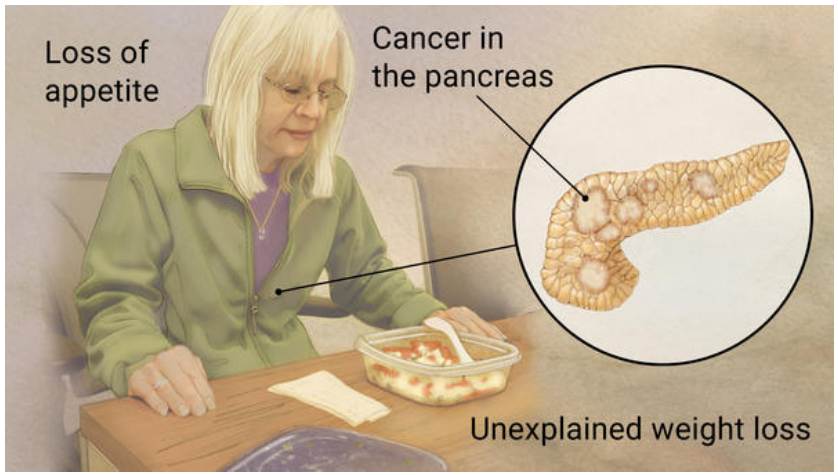
Computer Science Department, Carnegie Mellon University

Natasa Miskov-Zivanov, Bing Liu, James R. Faeder, Michael Lotze, Edmund M. Clarke

CMSB 2016

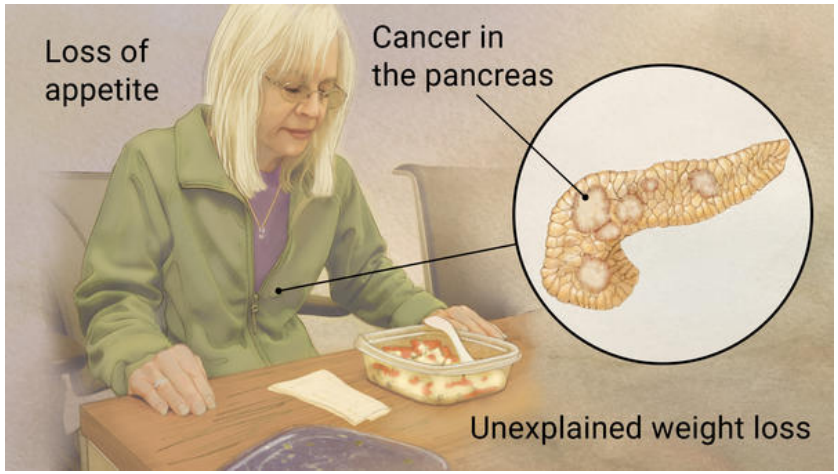
Pancreatic Cancer

- the **7th** most common cause of cancer deaths globally, and
- the **4th** in US



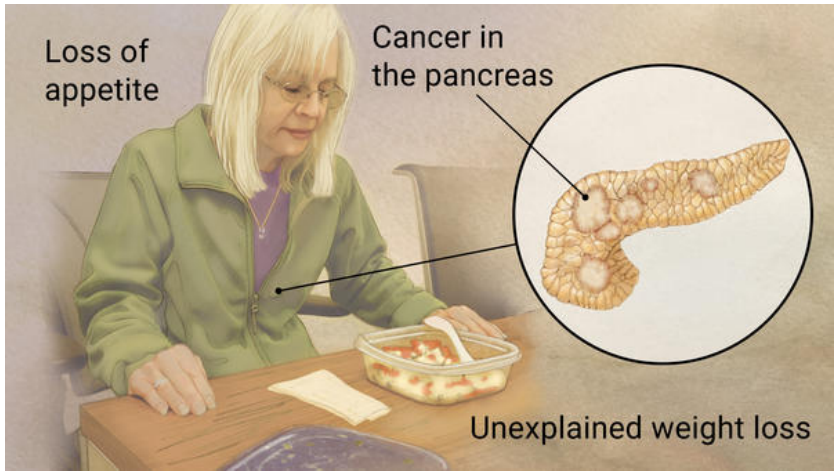
Pancreatic Cancer

- the **7th** most common cause of cancer deaths globally
- the **4th** in US
- hard to diagnose in the early stages
  - no symptoms
  - the lack of biomarkers allowing early screening

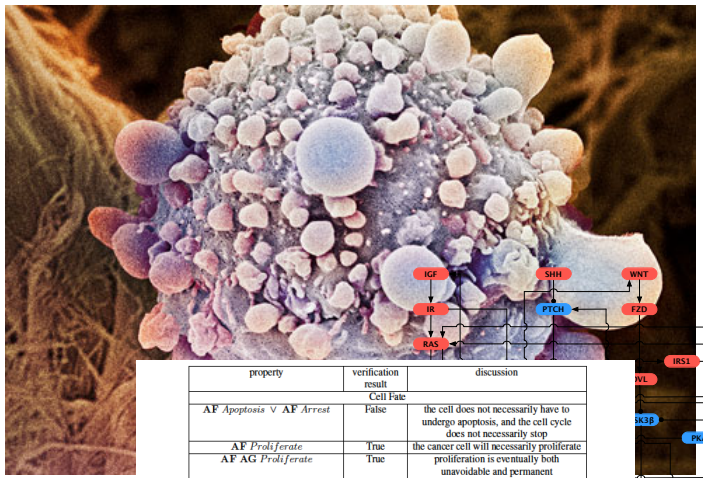


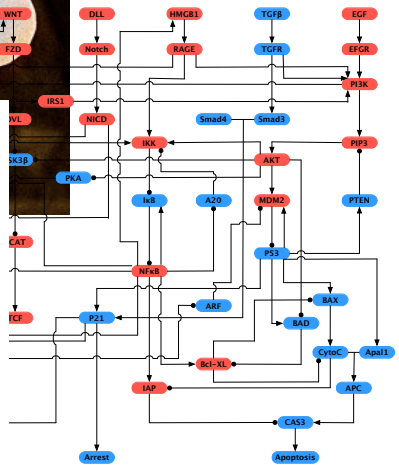
Pancreatic Cancer

- the **7th** most common cause of cancer deaths globally
- the **4th** in US
- hard to diagnose in the early stages
  - no symptoms
  - the lack of biomarkers allowing early screening
- very poor prognosis



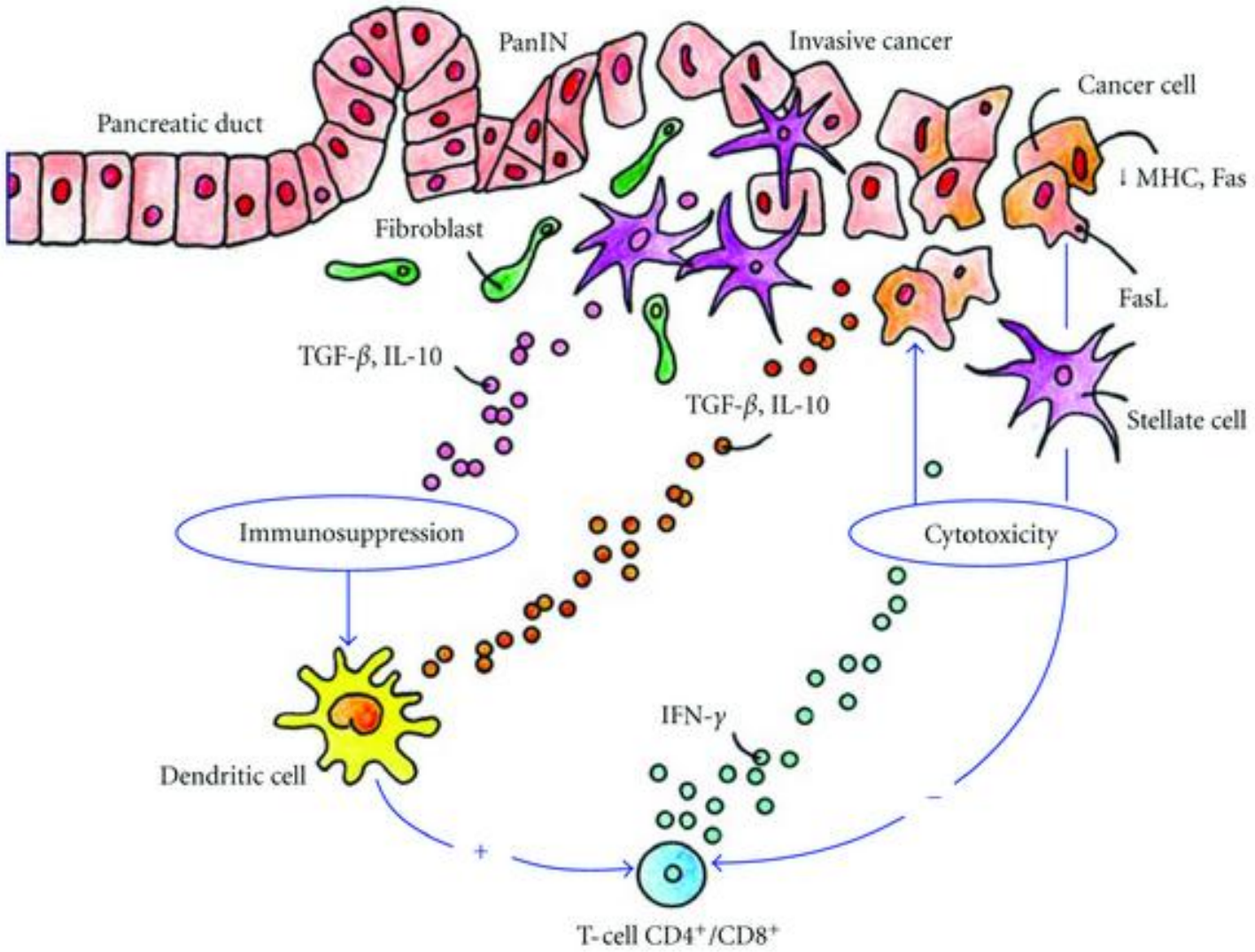
Studies on Pancreatic Cancer Cells



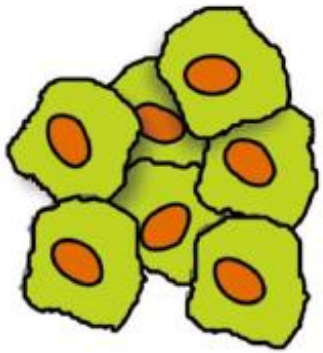


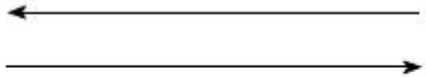
property	verification result	discussion
Cell Fate		
AF *Apoptosis* ∨ **AF** *Arrest*	False	the cell does not necessarily have to undergo apoptosis, and the cell cycle does not necessarily stop
AF *Proliferate*	True	the cancer cell will necessarily proliferate
AF AG *Proliferate*	True	proliferation is eventually both unavoidable and permanent
AF *!Apoptosis* ∧ **AF** *!Arrest*	True	it is always possible for the cancer cell to reach states in which Apoptosis and Arrest are OFF, thereby making cell proliferation possible
AF (*!Apoptosis* ∧ *!Arrest* ∧ *Proliferate*)	False	the model cannot always eventually reach a state in which apoptosis and cell cycle arrest are not inhibited and cell proliferation is active
AF AG *!Apoptosis* ∨ **AF AG** *!Arrest*	False	inhibition of apoptosis and cell cycle arrest are not unavoidable and permanent
Cell Cycle		
A (*!Proliferate* **U** *CyclinD*)	True	it is always the case that cell proliferation does not occur until Cyclin D is expressed (or activated)
AF AG *CyclinD*	False	in our model the activation of Cyclin D is not a steady state
!E (*!P53* **U** *Apoptosis*)	False	apoptosis can be activated even when P53 is not
Oscillations		
$TGF\beta \rightarrow$ **AG** ((*!NFκB* → **AF** *NFκB*) ∧ (*NFκB* → **AF** *!NFκB*)	True	an initial overexpression of TGFβ always leads to oscillations in NFκB's expression level
$PIP3 \rightarrow$ **AG** ((*!NFκB* → **AF** *NFκB*) ∧ (*NFκB* → **AF** *!NFκB*))	True	PIP3 has the similar impact on NFκB's expression level
AG ((*P53* → **AF** *MDM2*) ∧ (*MDM2* → **AF** *!P53*))	True	overexpression of P53 will always activate MDM2, which will in turn inhibit P53

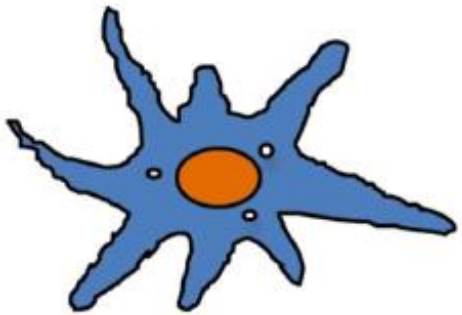
Pancreatic Cancer Microenvironment / Stroma



Pancreatic Cancer Cells and Stellate Cells







Pancreatic cancer cells

- ↑Proliferation
- ↓Apoptosis

] ↑Survival

- ↑Migration/invasion
- ↑Metastasis
- ↑Stem cell niche

Activated pancreatic stellate cells

- ↑Proliferation
- ↑Fibrosis/ECM synthesis
- ↑Angiogenic factors, MMPs
- ↑Migration and metastasis

Motivation and Contributions

Motivation **Contributions**

- Study the interplay between PCCs and PSCs, and identify major pathways and molecules in PSCs

Motivation and Contributions

Motivation

- Study the interplay between PCCs and PSCs, and Identify major pathways and molecules in PSCs

Contributions

- Construct the first multicellular and multiscale model

Motivation and Contributions

Motivation

- Study the interplay between PCCs and PSCs, and Identify major pathways and molecules in PSCs
- Appropriate modeling formalism (multiple cells, cell populations, both cellular and molecular dynamics, $\cdots$)

Contributions

- Construct the first multicellular and multiscale model

Motivation and Contributions

Motivation

- Study the interplay between PCCs and PSCs, and Identify major pathways and molecules in PSCs
- Appropriate modeling formalism (multiple cells, cell populations, both cellular and molecular dynamics, $\cdots$)

Contributions

- Construct the first multicellular and multiscale model
- Propose a multiscale hybrid rule-based modeling language

Motivation and Contributions

Motivation

- Study the interplay between PCCs and PSCs, and Identify major pathways and molecules in PSCs
- Appropriate modeling formalism (multiple cells, cell populations, both cellular and molecular dynamics, $\cdots$)
- Validate our model, and then predict possible targets for PC treatments

Contributions

- Construct the first multicellular and multiscale model
- Propose a multiscale hybrid rule-based modeling language

Motivation and Contributions

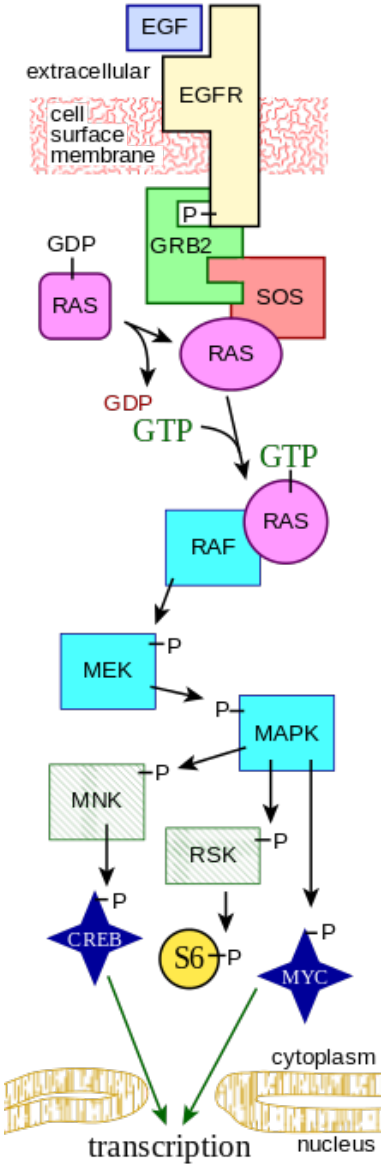
Motivation

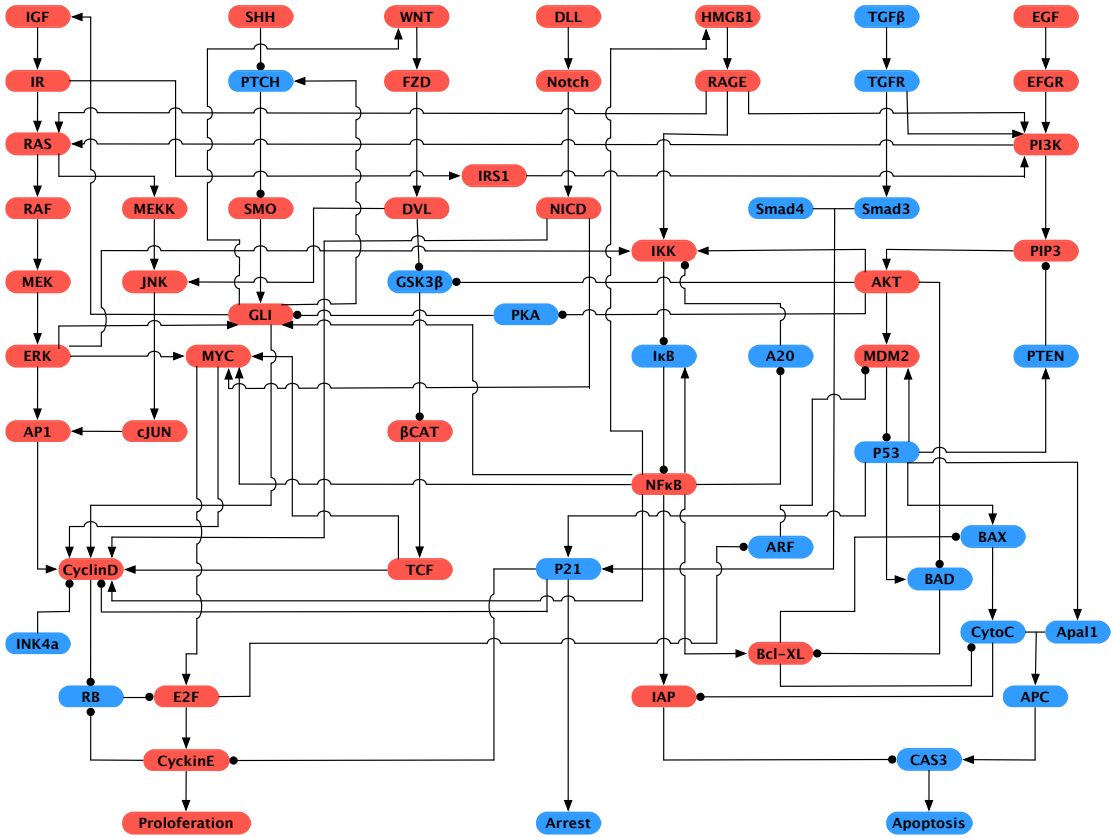
- Study the interplay between PCCs and PSCs, and Identify major pathways and molecules in PSCs
- Appropriate modeling formalism (multiple cells, cell populations, both cellular and molecular dynamics, $\cdots$)
- Validate our model, and then predict possible targets for PC treatments

Contributions

- Construct the first multicellular and multiscale model
- Propose a multiscale hybrid rule-based modeling language
- Statistical model checking is used to carry out model validation and prediction

Cell Signaling Pathways





Our Pancreatic Cancer Microenvironment Model

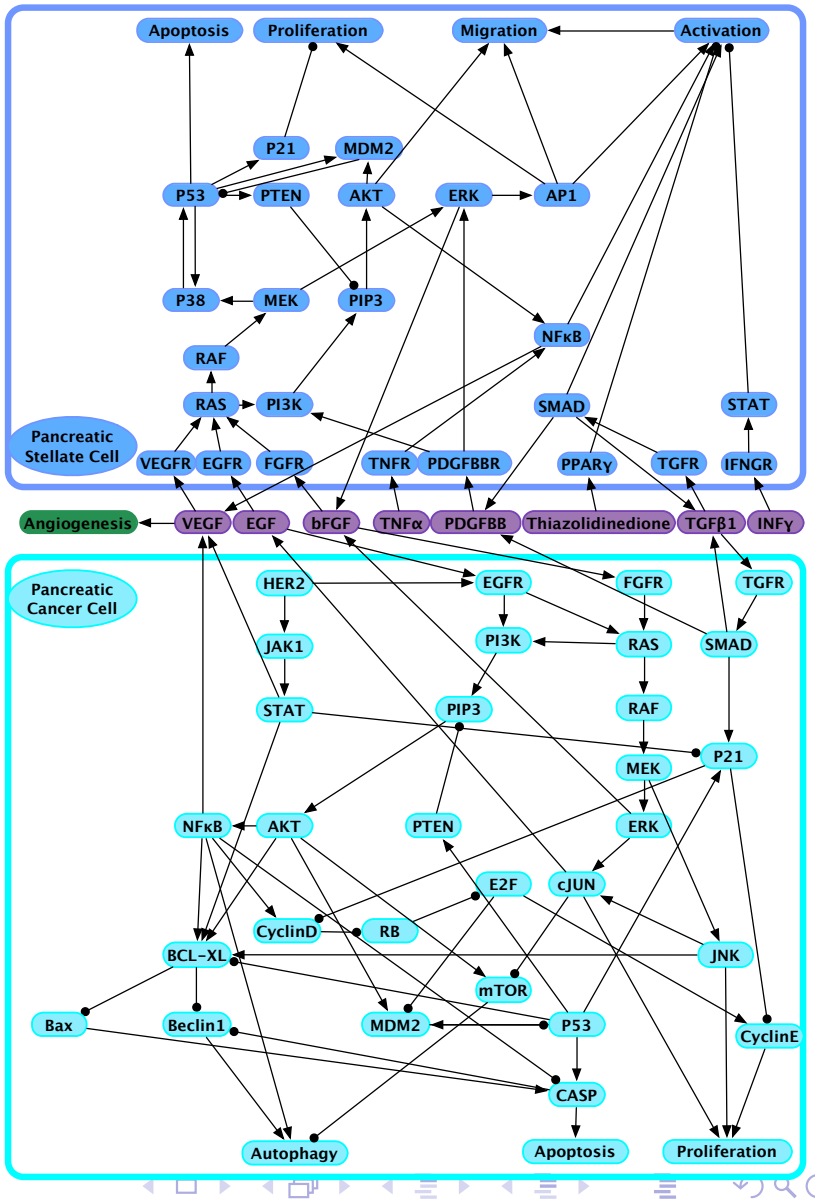
Pancreatic cancer cell (PCCs):
Pathways regulating

- Proliferation,
- Apoptosis, and
- Autophagy.

Pancreatic stellate cell (PSCs):
Pathways regulating

- Proliferation,
- Apoptosis,
- Activation, and
- Migration.

Interactions between PCCs and PSCs:
EGF, bFGF, VEGF, TGFβ1, and PDGFBB



Biological Background - Pancreatic Cancer Cells

Cell Function	**Promote (+) / Inhibit (-)**	**Pathway**
Proliferation	+	K-RAS mutation-induced RAS pathway
	+	HER2/neu mutation-induced EGFR pathway
	+	EGF-EGFR pathway
	+	bFGF pathway
Apoptosis	+	TGFβ1 pathway
	-	K-RAS mutation-induced PI3K pathway
	-	HER2/neu mutation-induced PI3K pathway
Autophagy	-	Pathways upregulating mTOR
	+	Overexpressed NFκB and Beclin1

Biological Background - Pancreatic Stellate Cells

Cell Function	**Promote (+) / Inhibit (-)**	**Pathway**
Activation	+	PDGFBB pathway
	+	TGFβ1 pathway
	+	TNFα pathway
Migration	+	MAPK pathway upregulated by EGF, bFGF, and VEGF
	+	PDGFBB regulated PI3K pathway
	+	PDGFBB regulated ERK-AP1 pathway
Proliferation	+	ERK-AP1 pathway upregulated by growth factors
	-	Pathways upregulating tumor suppressers
Apoptosis	+	MAPK pathway via P53

Biological Background - Extracellular Molecules

Pancreatic Cancer Cells

Autocrine and paracrine involving EGF
Autocrine and paracrine involving bFGF
Paracrine involving VEGF
Autocrine and paracrine involving TGFβ1
Paracrine involving PDGFBB

Pancreatic Stellate Cells

Multiscale Hybrid Rule-based Modelling

(Traditional) Rule-based Modeling (i.e. BioNetGen)

Multiscale Hybrid Rule-based Modelling

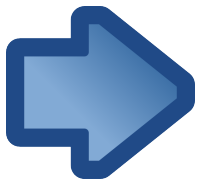
(Traditional) Rule-based Modeling (i.e. BioNetGen) aims at:

- Modeling reactions involving intracellular signaling molecules
- Describing dynamics continuously

Multiscale Hybrid Rule-based Modelling

(Traditional) Rule-based Modeling (i.e. BioNetGen) aims at:

- Modeling reactions involving intracellular signaling molecules
- Describing dynamics continuously



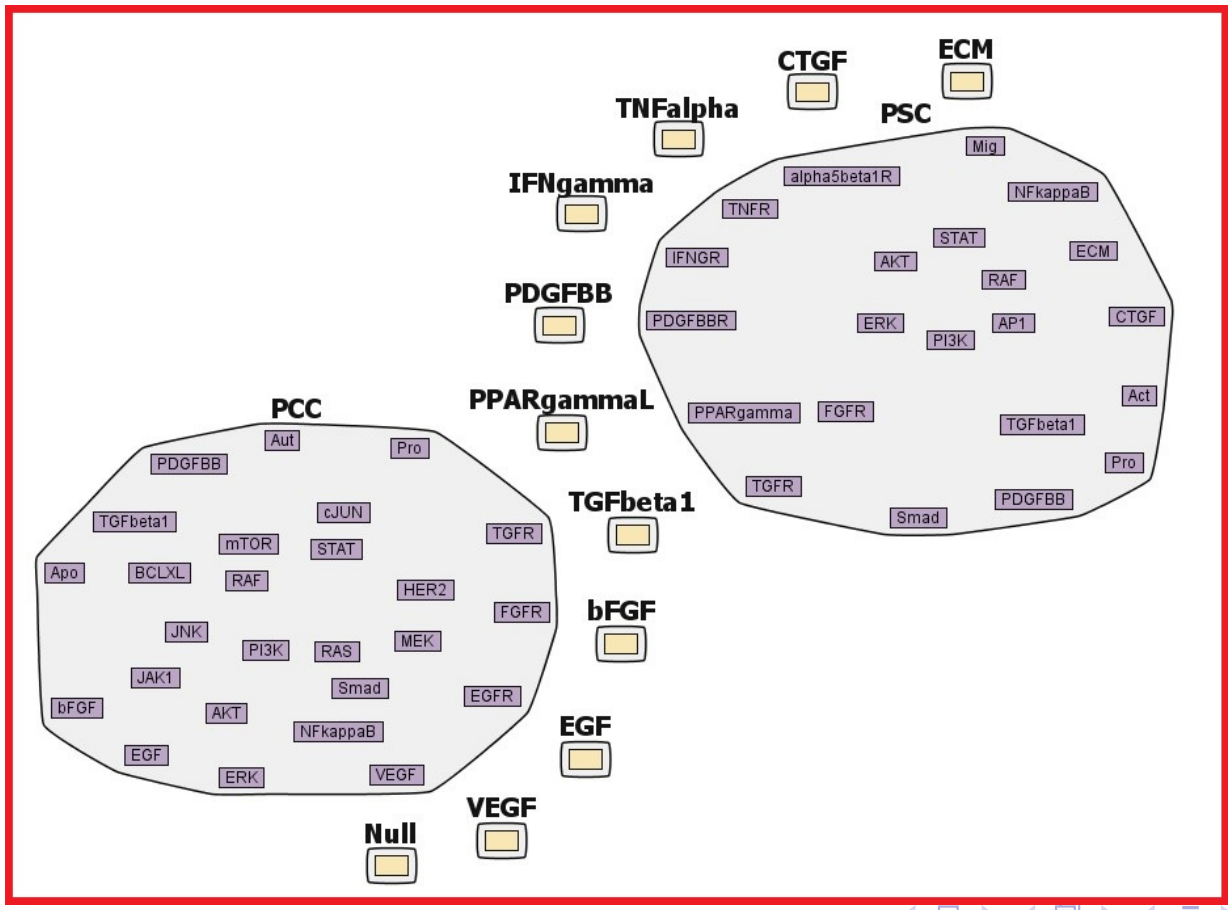
Multiscale Hybrid Rule-based Modeling can also:

- Describe intercellular interplay together with intracellular reactions
- In a hybrid way: continuously for intercellular, and discrete for intracellular

Multiscale Hybrid Rule-based Modelling

The basic building blocks

- Cells (with subunits as intracellular molecules), or
- Extracellular molecules (with no subunits)



The basic building blocks (Con.)

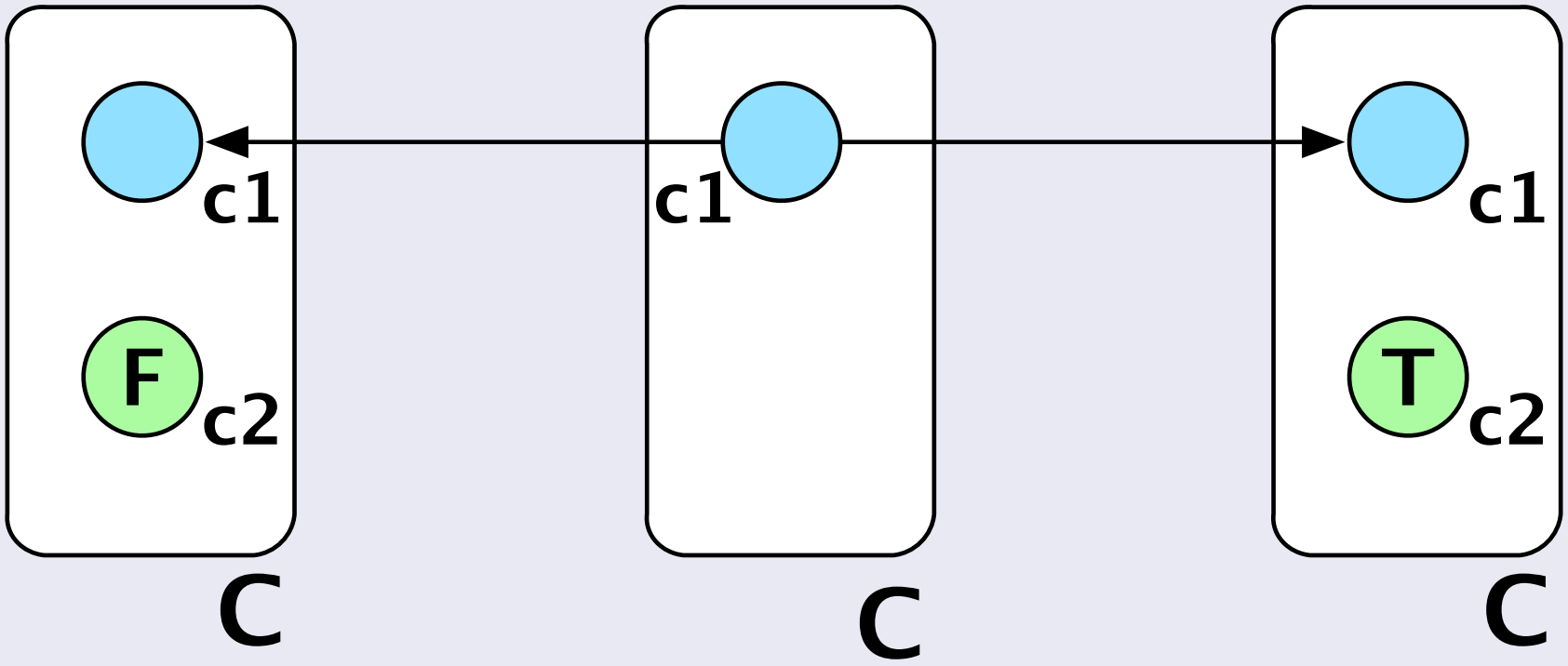
- Boolean values (**T** or **F**), easy to extend to discrete values
- Different biological meanings

Subunit	**T**	**F**
cell function / secretion	being triggered	not being triggered
receptor	being bounded	being free
protein	high concentration	low concentration
...	...	...

Multiscale Hybrid Rule-based Modelling

Patterns

- To identify **a set of species that share a set of features**
- Provides a rich yet concise description



Multiscale Hybrid Rule-based Modelling - Rules

Rules
Rule 1: Ligand-receptor binding
Rule 2: Mutated receptors form a heterodimer
Rule 3: Downstream regulation: Encoding Logical Functions as Rules
Rule 4: Cell functions
Rule 5: Secretion
Rule 6: Degradation of extracellular molecules
Rule 7: Mutation
Rule 8: Constantly over-expressed extracellular molecules
Rule 9: Human/treatment intervention

Multiscale Hybrid Rule-based Modelling - Rules

Rule 1: Ligand-receptor binding

$$Lig + Cell(Rec \sim F) \rightarrow Cell(Rec \sim T) \quad brate$$

Multiscale Hybrid Rule-based Modelling - Rules

Rule 3: Downstream regulation: Encoding Logical Functions as Rules

Given a logical updating function $Mol_3^{(t+1)} = \neg Mol_1^{(t)} \times (Mol_2^{(t)} + Mol_3^{(t)})$ where "Mol_1" is the inhibitor and "Mol_2" is the activator of "Mol_3".

$$Cell(Mol_1 \sim F, Mol_2 \sim T, Mol_3 \sim F) \rightarrow Cell(Mol_1 \sim F, Mol_2 \sim T, Mol_3 \sim T) \quad trate$$

$$Cell(Mol_1 \sim T, Mol_3 \sim T) \rightarrow Cell(Mol_1 \sim T, Mol_3 \sim F) \quad trate$$

Multiscale Hybrid Rule-based Modelling - Rules

Rule 7: Mutation

$$Cell(Mol \sim F) \rightarrow Cell(Mol \sim T) \quad mrate$$

$$Cell(Mol \sim T) \rightarrow Cell(Mol \sim F) \quad mrate$$

Multiscale Hybrid Rule-based Modelling - Rules

Rule 9: Human/treatment intervention

$$Cell(Mol \sim T) \rightarrow Cell(Mol \sim F) \quad intrate$$

$$Cell(Mol \sim F) \rightarrow Cell(Mol \sim T) \quad intrate$$

$$CancerEnv \rightarrow extraMol \quad intrate$$

$$extraMol \rightarrow Null() \quad intrate$$

Statistical Model Checking

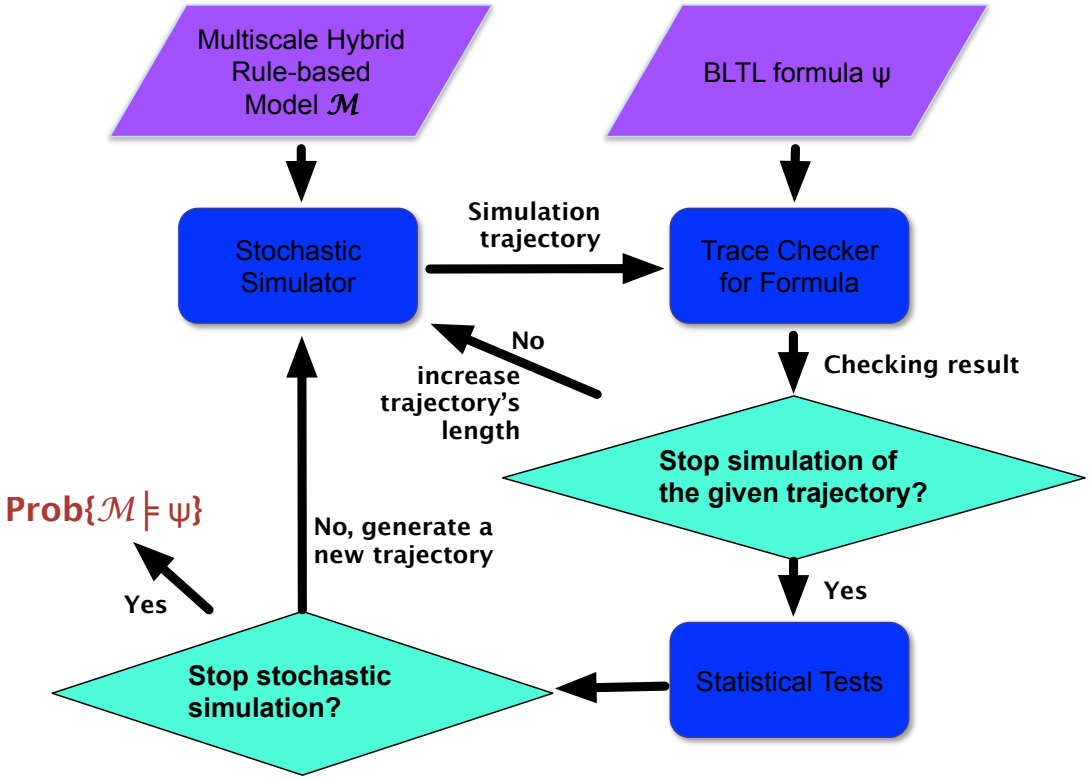
Bounded Linear Temporal Logic (BLTL)

The syntax of BLTL is given by: $\psi ::= x \sim v | \neg\psi | \psi_1 \vee \psi_2 | \psi_1 U^t \psi_2$
Example BLTL formula: $\neg F^5 G^{10}(Ras = 1 \wedge P53 = 0)$

F: eventually, G: always, U: until

Statistical Model Checking to Estimate $Prob_{=?}(\mathcal{M} \models \psi)$

- State Space Exploration unavoidable for complex systems
- Easier to simulate a complex system than to build its transition relation
- Goal: Provide probabilistic guarantees using fewer simulations
- Method: Trace Checker + Statistical Testing Methods



Results - Three Scenarios

Property	Estimated Prob	# Succ	# Sample	Time (s)	Note
Scenario I: mutated PCCs with no treatments					
1	0.4053	10585	26112	208.91	w.o. PSCs
	0.9961	256	256	1.83	w. PSCs
2	0.1191	830	6976	49.69	w.o. PCCs
	0.9961	256	256	1.75	w. PCCs
3	0.9961	256	256	5.21	-
4	0.9961	256	256	4.38	-
Scenario II: mutated PCCs with different exsiting treatments					
5	0.0004	0	2304	17.13	cetuximab and erlotinib
	0.0012	10	9152	68.67	gemcitabine
	0.7810	8873	11360	114.25	nab-paclitaxel
	0.8004	7753	9686	73.83	ruxolitinib
Scenario III: mutated PCCs with blocking out on possible target(s)					
6	0.0792	38363	484128	3727.99	w.o. inhibiting ERK in PSCs
	0.9822	2201	2240	17.37	w. inhibiting ERK in PSCs
7	0.1979	3409	17232	136.39	w.o. inhibiting ERK in PSCs
	0.9961	256	256	2.01	w. inhibiting ERK in PSCs
8	0.2029	2181	10752	92.57	w.o. inhibiting MDM2 in PSCs
	0.9961	256	256	2.18	w. inhibiting MDM2 in PSCs
9	0.0004	0	2304	15.77	w.o. inhibiting RAS in PCCs and ERK in PSCs
	0.9961	256	256	3.15	w. inhibiting RAS in PCCs and ERK in PSCs
10	0.9797	1349	1376	11.98	w.o. inhibiting STAT in PCCs and NFκB in PSCs
	0.1631	1476	9056	81.61	w. inhibiting STAT in PCCs and NFκB in PSCs

Results - Scenario I: with no treatments

Property 1: To estimate the probability that the population of PCCs will eventually reach and maintain in a high level.

$$Prob_{=?}\ \{(PCCtot = 10) \wedge F^{1200}\ G^{100}\ (PCCtot > 200)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.4053	10585	26112	208.91	w.o. PSCs
0.9961	256	256	1.83	w. PSCs

Results - Scenario I: with no treatments

Property 2: To estimate the probability that the number of migrated PSCs will eventually reach and maintain in a high amount.

$$Prob_{=?}\ \{(MigPSC = 0) \wedge F^{1200}\ G^{100}\ (MigPSC > 40)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.1191	830	6976	49.69	w.o. PCCs
0.9961	256	256	1.75	w. PCCs

Results - Scenario II: with existing treatments

Property 5: To estimate the probability that the population of PCCs will eventually drop to and maintain in a low amount.

$$Prob_{=?}\ \{(PCCtot = 10) \wedge F^{1200}\ G^{400}\ (PCCtot < 100)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.0004	0	2304	17.13	cetuximab and erlotinib
0.0012	10	9152	68.67	gemcitabine
0.7810	8873	11360	114.25	nab-paclitaxel
0.8004	7753	9686	73.83	ruxolitinib

Results - Scenario III: block out possible target(s)

Targeting at ERK in PSCs

Property 6: To estimate the probability that the number of PSCs will eventually drop to and maintain in a low level.

$$Prob_{=?}\ \{(PSCtot = 5) \wedge F^{1200}\ G^{400}\ (PSCtot < 30)\}$$

Property 7: To estimate the probability that the population of migrated PSCs will eventually stay in a low amount.

$$Prob_{=?}\ \{(MigPSC = 0) \wedge F^{1200}\ G^{100}\ (MigPSC < 30)\}$$

Property	Estimated Prob	# Succ	# Sample	Time (s)	Note
6	0.0792	38363	484128	3727.99	not inhibit
	0.9822	2201	2240	17.37	inhibit
7	0.1979	3409	17232	136.39	not inhibit
	0.9961	256	256	2.01	inhibit

Results - Scenario III: block out possible target(s)

Property 8: To estimate the probability that the number of PSCs entering the proliferation phase will eventually be less than the number of PSCs starting the apoptosis programme and this situation will maintain. (Target at MDM2 in PSCs)

$$Prob_{=?}\ \{F^{1200}\ G^{400}\ ((PSCPro - PSCApop) < 0)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.2029	2181	10752	92.57	not inhibit
0.9961	256	256	2.18	inhibit

Results - Scenario III: block out possible target(s)

Property 9: To estimate the probability that the number of bFGF will eventually stay in such a low level. (RAS in PCCs and ERK in PSCs)

$$Prob_{=?}\ \{F^{1200}\ G^{400}\ (bFGF < 100)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.0004	0	2304	15.77	not inhibit
0.9961	256	256	3.15	inhibit

Results - Scenario III: block out possible target(s)

Property 10: To estimate the probability that the concentration of VEGF will eventually reach and keep in a high level. (STAT3/4 in PCCs and NFκB in PSCs)

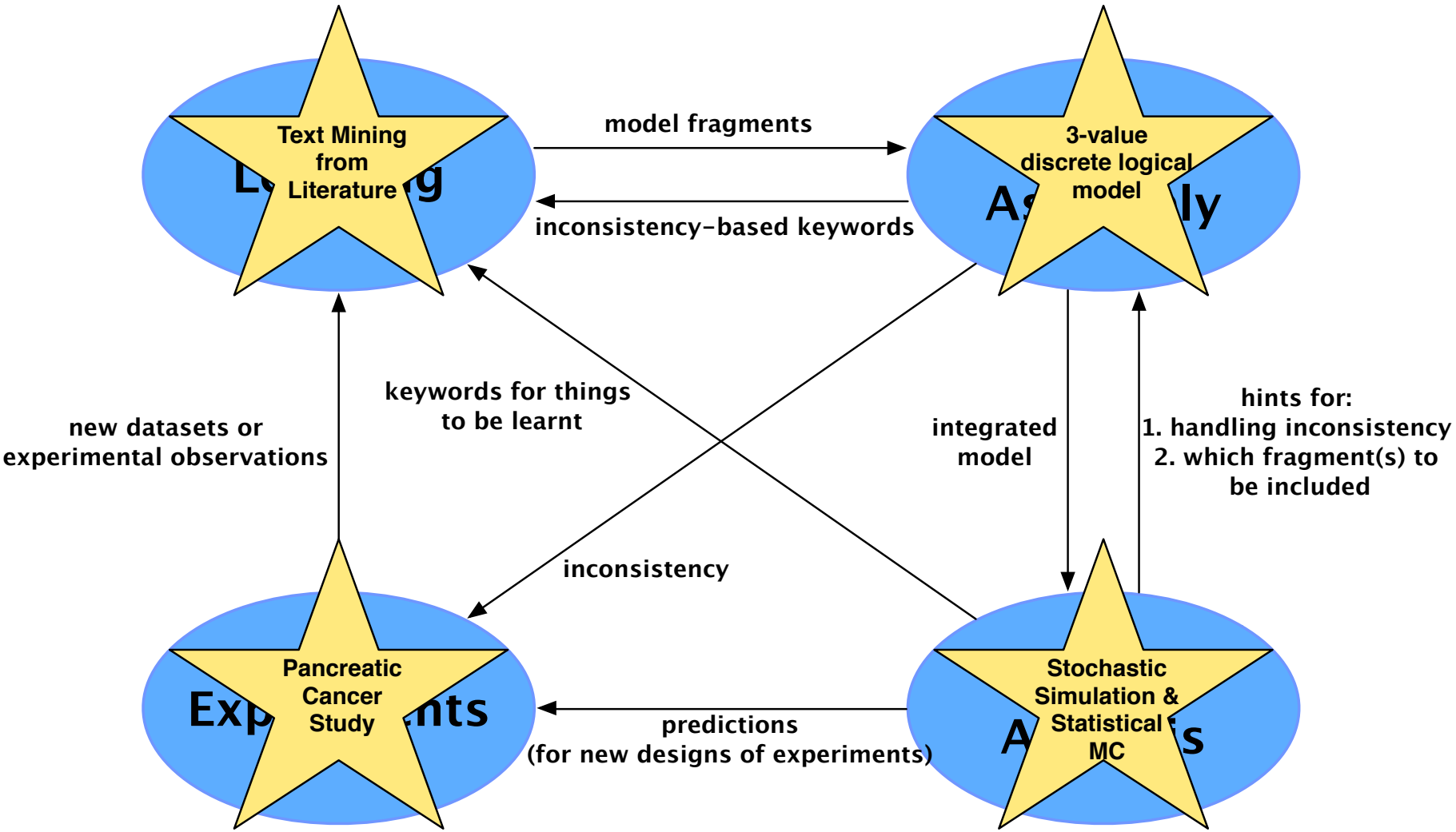
$$Prob_{=?}\{F^{400}\ G^{100}\ (VEGF > 200)\}$$

Estimated Prob	# Succ	# Sample	Time (s)	Note
0.9797	1349	1376	11.98	not inhibit
0.1631	1476	9056	81.61	inhibit

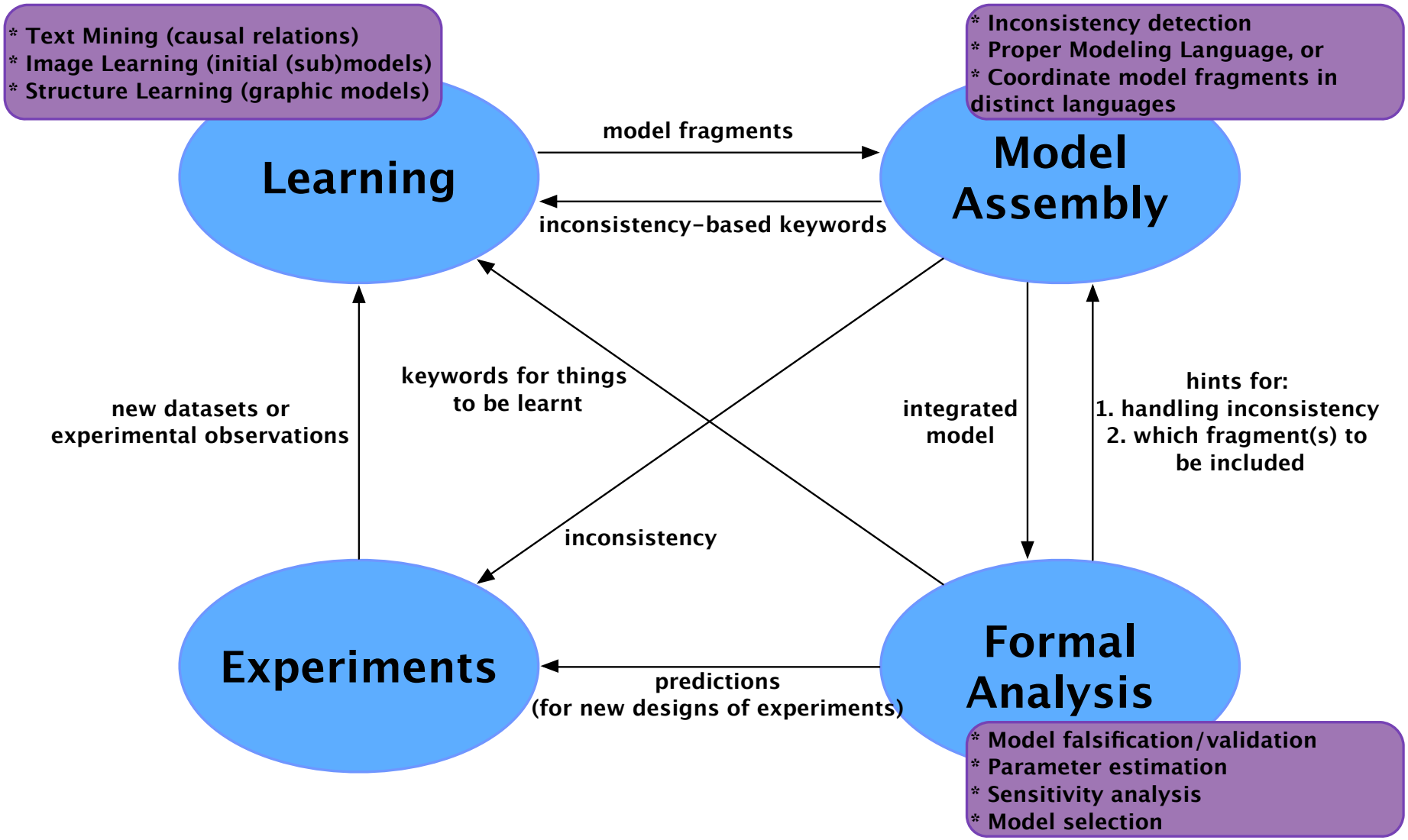
Conclusion

- Construct a multicellular and multiscale model
- Propose a language for multiscale biological systems using continuous and discrete rules
- Apply stochastic simulation and StatMC to analyze system behaviors under diffident conditions
- Confirm experimental findings
- Gain insights on how existing treatments latching onto different targets can lead to distinct outcomes
- Predict potential new targets aiming at depleting PSCs and inhibiting the PC development

Future Work


Text Mining from Literature
model fragments
inconsistency-based keywords
3-value discrete logical model
keywords for things to be learnt
new datasets or experimental observations
integrated model
hints for:
1. handling inconsistency
2. which fragment(s) to be included
inconsistency
Pancreatic Cancer Study
predictions
(for new designs of experiments)
Stochastic Simulation & Statistical MC

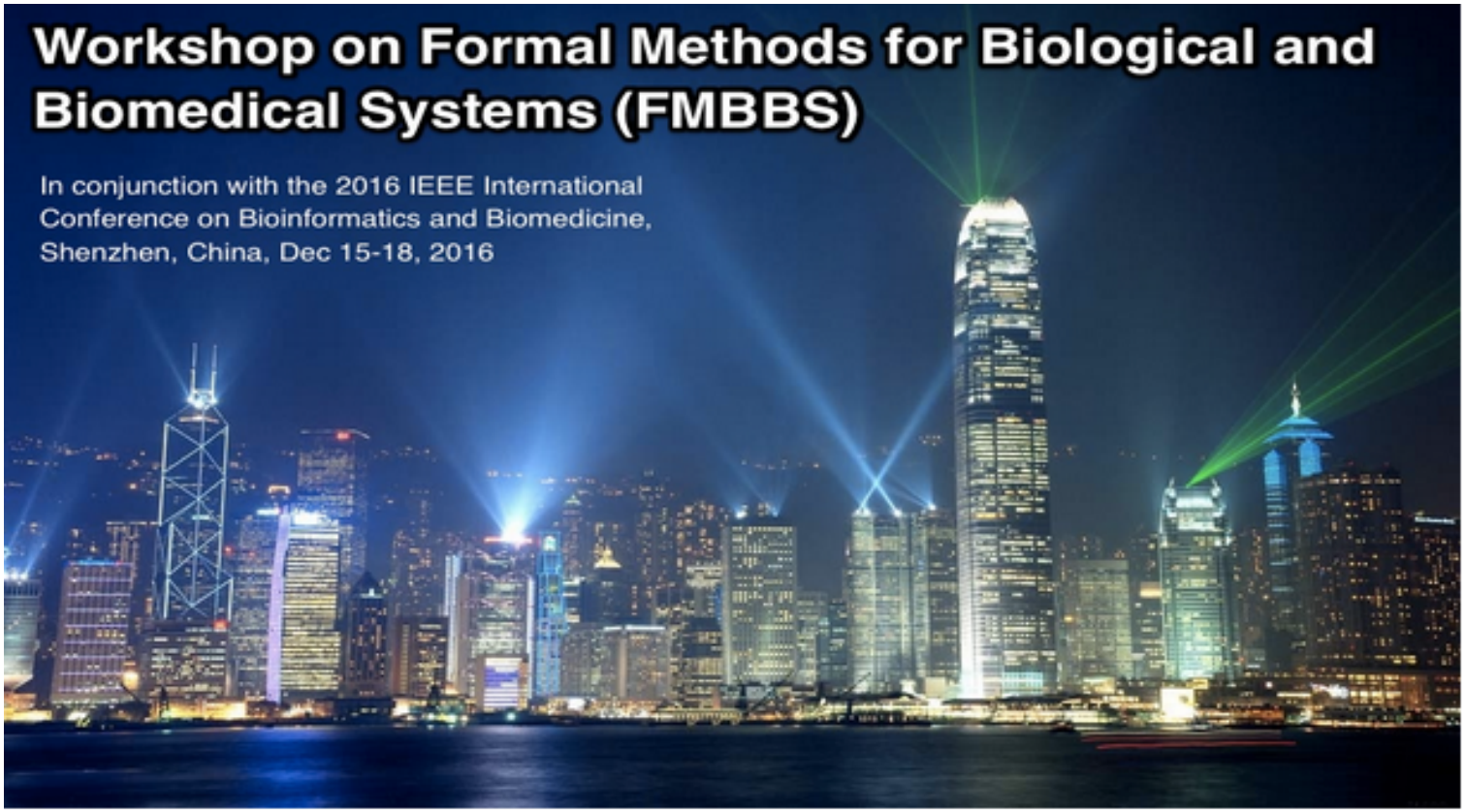
Future Work


* Text Mining (causal relations)
* Image Learning (initial (sub)models)
* Structure Learning (graphic models)
* Inconsistency detection
* Proper Modeling Language, or
* Coordinate model fragments in distinct languages
Learning
Model Assembly
model fragments
inconsistency-based keywords
new datasets or experimental observations
keywords for things to be learnt
integrated model
hints for:
1. handling inconsistency
2. which fragment(s) to be included
inconsistency
Experiments
Formal Analysis
predictions
(for new designs of experiments)
* Model falsification/validation
* Parameter estimation
* Sensitivity analysis
* Model selection

Q & A

Thanks for your time! ~~Questions?~~

FMBBS 2016

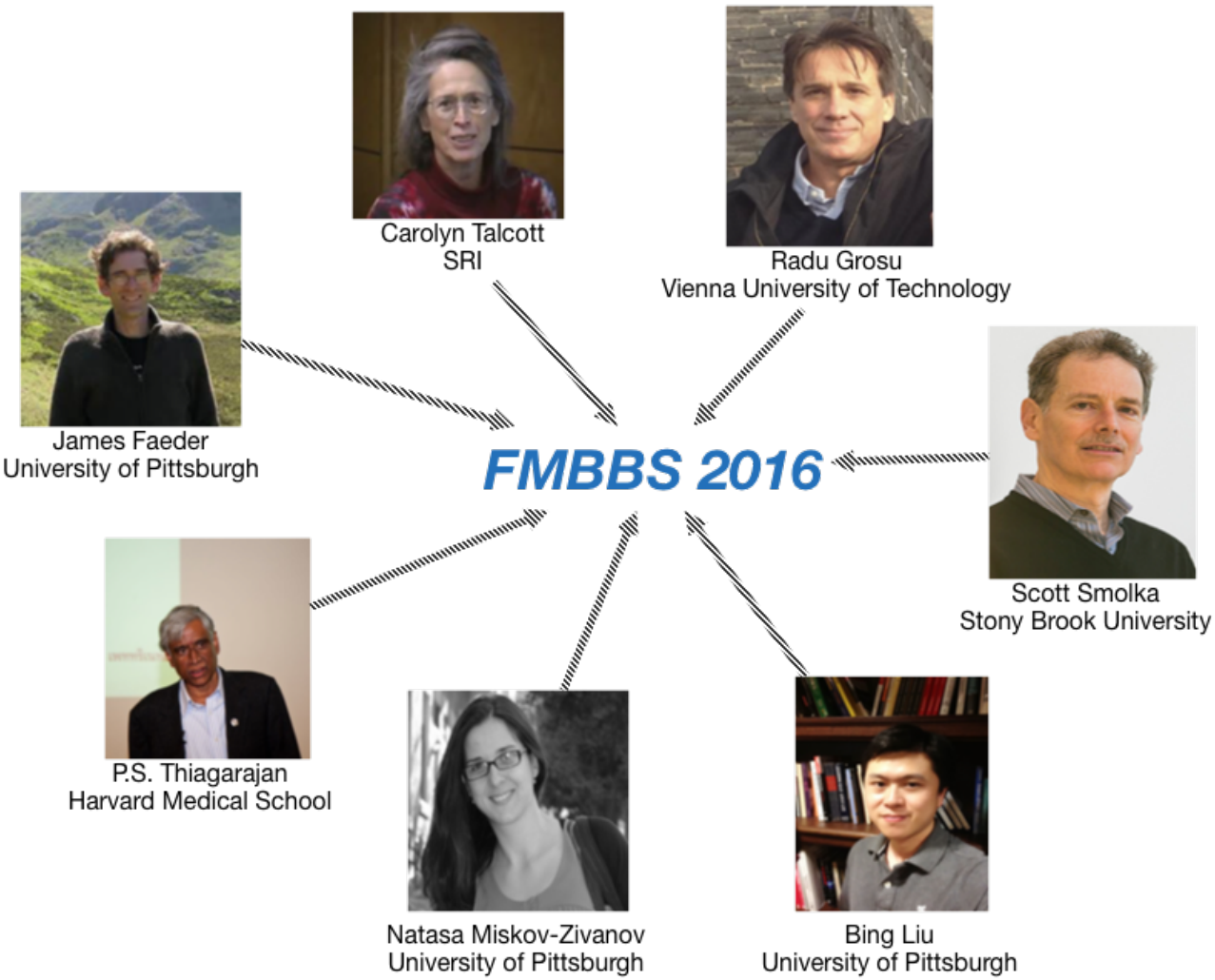


Home
Call for Papers
Organization
Accepted Papers
Invited Speakers
Program
Venue
Registration

FMBBS 2016

As biomedical research advances into more complicated systems, there is an increasing need to model and analyze these systems to better understand them. For decades, biologists have been using diagrammatic models to describe and understand the mechanisms and dynamics behind their experimental observations. Although these models are simple to build and understand, they offer only a rather static picture of the corresponding biological systems, and scalability is limited. Formal specification and analysis methods, such as model checking techniques, hold great promise in promoting further discovery and innovation for complicated biochemical systems. Models can be tested and adapted inexpensively in silico to provide new insights. However, development of accurate and efficient modeling methodologies and analysis techniques for biochemical systems is still an open challenge. This workshop will provide an opportunity for practitioners to present their work in this area to both computer scientists

FMBBS 2016


Carolyn Talcott
SRI
Radu Grosu
Vienna University of Technology
James Faeder
University of Pittsburgh
FMBBS 2016
Scott Smolka
Stony Brook University
P.S. Thiagarajan
Harvard Medical School
Natasa Miskov-Zivanov
University of Pittsburgh
Bing Liu
University of Pittsburgh

Q & A

Questions?