

Name:

Bitte schreiben Sie Ihren Nachnamen auf jede Seite!

Sessionsprüfung

August 2016

Systembiologie: 551-1174-00L

- *Es sind keine Hilfsmittel erlaubt.*
- *Die Prüfungssprache ist deutsch, aber Sie dürfen englische Ausdrücke verwenden.*
- *Bitte antworten Sie möglichst kurz.*
- *Insgesamt sind 44 Punkte mit 16 Fragen zu erreichen. 24 Punkte im Stelling/Sauer Teil; 20 Punkte im Borgwardt/Zamboni Teil.*
- *22 Punkte = Note 4.0; 40 Punkte = Note 6.0!*

Name:

Prüfungsteil Sauer/Stelling

1. Im Stoffwechselweg $A \rightarrow B \rightarrow C$ gibt es eine Feedbackhemmung von C auf das erste Enzym, die Umwandlung von A zu B. (*Stichworte reichen aus*)
 - a. Nennen Sie 2 Gründe wieso dieser Feedback für die Zelle notwendig sein könnte. **(1 P)**
 - b. Nennen Sie 3 Beispiele wie diese Hemmung mechanistisch erreicht werden könnte und geben Sie eine kurze Erklärung bezüglich der zu erwartenden Geschwindigkeit. **(2 P)**

Name:

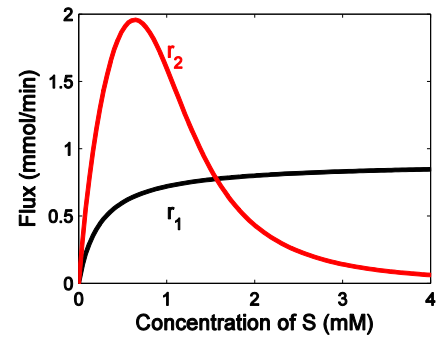
2. Sie möchten die ersten 5 Minuten der Stoffwechselreaktion von Leberzellen auf die Zugabe von Inhibitoren der Glykolyse vorhersagen. Dazu entwickeln Sie ein Modell welches Sie mit der Reaktion auf einen bekannten Inhibitor trainieren. (*Stichworte reichen aus*)
- a) Welche Art von Modell brauchen Sie für diese Aufgabe? **(0.5 P)**
 - b) Welche Informationen und welche Daten brauchen Sie dafür? **(1.5 P)**
 - c) Beschreiben Sie kurz wie Sie das Problem angehen, wie sie das Modell entwickeln werden, und was ihr Vorgehen ist falls das Modell die Daten nicht beschreiben kann. **(2 P)**

Name:

3. Die irreversible Reaktion $\mathbf{S} \rightarrow \mathbf{P}$ wird vom Enzym \mathbf{E} mit dem katalytischen Koeffizienten k_{cat} und der Substrataffinität k_m katalysiert. Die Reaktion folgt einer Michaelis Menten Kinetik und der Fluss wird durch die folgende Gleichung beschrieben: $v = [E] \cdot k_{cat} \cdot \frac{[S]}{k_m + [S]}$
- a) Sie geben einen kompetitiven Effektor \mathbf{m}_c zu. Welcher Teil der Gleichung ($[E]$, k_{cat} , k_m , $[S]$) wird sich ändern und wieso? **(1 P)**
- b) Sie geben einen nicht-kompetitiven Aktivator \mathbf{m}_{nc} zu, der allosterisch auf Enzyme \mathbf{E} wirkt. Welcher Teil der Gleichung wird sich ändern und wieso? Wird der Fluss v grösser, kleiner oder bleibt er unverändert, wenn die Konzentration von \mathbf{m}_{nc} verdoppelt wird? **(1 P)**
- c) Nehmen Sie an, dass System wäre transkriptionell reguliert. Welcher Teil der Gleichung wird sich ändern und wieso? Wird der Fluss v grösser, kleiner oder bleibt er unverändert, wenn sich das Expressionlevel auf 10% des Normalzustandes reduziert (unter der Annahme, dass von jedem Transkript ein Protein gebildet wird)? **(1.5 P)**

Name:

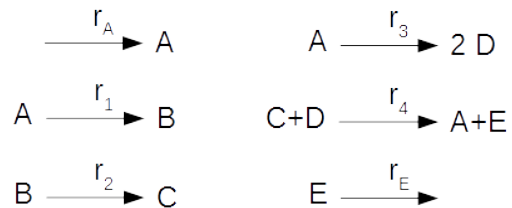
4. Ein Substrat **S** wird mit Rate r_0 von der Zelle aufgenommen und von zwei irreversiblen metabolischen Pathways weiter verwendet. Die Kinetiken der beiden ersten Enzyme der Pathways, die die Flüsse r_1 und r_2 generieren, sind in der Abbildung rechts gezeigt. Das System sei immer im steady state.



- a) Um welche Art von Kinetik handelt es bei dem Enzym für r_2 ? (0.5 P)
- b) Schreiben Sie die Bilanzgleichung für **S** als Funktion der Raten. (0.5 P)
- c) Durch das externe Substrat sei der Fluss $r_0 = 0.5$ mmol/min fixiert. Warum muss in diesem Fall die interne Substratkonzentration < 0.5 mM sein? (1 P)
- d) Welche interne(n) Substratkonzentration(en) sind möglich für $r_0 = 1$ mmol/min (Begründen Sie Ihre Antwort und geben Sie eine ungefähre Abschätzung)? (1 P)
- e) Falls in d) mehrere stationäre Zustände existieren, was ist der Mechanismus und wovon hängt ab, in welchem Zustand sich das System befindet? (1 P)

Name: _____

5. Sie analysieren ein metabolisches Netzwerk im stationären Zustand. Das Netzwerk hat fünf interne Metabolite A-E, vier interne Reaktionen mit Flüssen r_1 - r_4 , sowie zwei externe Reaktionen mit Flüssen r_A und r_E . Alle Reaktionen sind irreversibel und unten gegeben.



Hinweis: Die Lösung von a) ist nicht Voraussetzung für die anderen Teile der Aufgabe.

- a) Geben Sie die stöchiometrische Matrix des Netzwerkes an. **(2 P)**
b) Sie haben den Fluss $r_E = 1$ mmol/h gemessen. Welche anderen stationären Flüsse können Sie direkt bestimmen und welche Werte haben diese? **(1 P)**
c) Welche Reaktionen sind essentiell um einen stationären Zustand zu erhalten in dem nicht alle Flüsse null sind (Begründen Sie die Antwort)? **(0.5 P)**

Name:

6. Sie haben ein stöchiometrisches Netzwerkmodell für den Menschen mit m Metaboliten und n enzym-katalysierten Reaktionen gegeben. Sie wollen FBA anwenden, um metabolische Funktionen des Netzwerkes zu analysieren.
- a) Geben sie ein kleines aber effizientes Skript in Pseudocode an, das es Ihnen erlaubt, herauszufinden welche Doppelmutanten letal für Zellwachstum sind. **(2 P)**
- b) Welche Objective Function würden Sie **nicht** wählen, um Vorhersagen über Flussverteilungen zu erhalten (und warum)? **(0.5 P)**
- c) Aus Genexpressionsanalysen wissen Sie, dass bestimmte Enzyme in bestimmten Geweben nicht exprimiert werden. Wie könnten Sie dieses Wissen nutzen um das Modell zu verfeinern? **(0.5 P)**

Name: _____

7. Die Aktivität einer Signalkaskade im stationären Zustand kann mit Hilfe der sogenannten Hill-Kinetik beschrieben werden:

$$A_i = \frac{A_{i-1}^n}{K_i^n + A_{i-1}^n},$$

wobei A_i die Aktivität der i -iten Kinase ist, welche von der vorhergehenden Kinaseaktivität A_{i-1} abhängt, n der Hill-Koeffizient (Exponent) ist und K_i die Michaelis-Menten Konstante für die Aktivierung von Kinase i ist.

- a) Wie muss n sein um Ultrasensitivität jedes Signalschrittes zu erreichen? **(0.5 P)**
- b) Nehmen Sie an, dass eine MAPK Kaskade mit $i = 1$ K 3 Kinasen die folgenden Parameterwerte hat: $n = 8$, $K_1 = 0.2$, $K_2 = 0.4$ und $K_3 = 0.8$. Skizzieren Sie graphisch die Aktivitäten der einzelnen Kinasen als Funktion der vorhergehenden Kinaseaktivität bzw. des Inputs A_0 (alle Signale sind in $[0,1]$). **(1.5 P)**
- c) Welche Kinasekaskade wird den Output auf der letzten Ebene bei einem geringeren Inputsignal aktivieren, die Kaskade aus b) oder ein alternativer Signalweg mit $n = 8$, $K_1 = 0.8$, $K_2 = 0.4$ und $K_3 = 0.2$ (Begründen Sie Ihre Antwort)? **(1 P)**

Name:

Prüfungsteil Borgwardt/Zamboni

8. Was bedeutet Instabilität bei der Featureselektion auf Untermengen desselben Datensatzes? **(1 P)**
9. Ein Forscher versucht, mittels Genexpressionsdaten einen Krankheitsphänotypen vorherzusagen. Er trainiert ein Modell auf den Trainingsdaten und sagt dann auf den Testdaten voraus. Wenn die Vorhersagegenauigkeit nicht zufriedenstellend ist, trainiert er ein neues Modell auf den Trainingsdaten. Dieses Vorgehen wiederholt er Dutzende Male, bis ein erfolgreiches Modell gefunden wurde. Auf die Rückfrage, ob er nicht Overfitting betreibe, antwortet der Forscher, er separiere stets Trainings- und Testdatensatz gründlich beim Trainieren und Testen und vermeide damit Overfitting.
1. Begründen Sie, warum hier in Wahrheit doch ein Problem mit Overfitting vorliegt. **(2 P)**
 2. Begründen Sie, warum ebenfalls ein Problem mit multiplem Hypothesentesten vorliegt. **(2 P)**

Name:

10. Ein Krankenhaus hat für eine Forschungsstudie Genexpressionsprofile von 1000 Lungenkrebspatienten mit 20.000 Genen erstellen lassen. Man geht davon aus, dass höchstens 1% der Gene mit dem Auftreten von Lungenkrebs assoziiert sind. Das Krankenhaus möchte nun Untergruppen unter den Lungenkrebspatienten an Hand der Genexpressionsdaten entdecken.

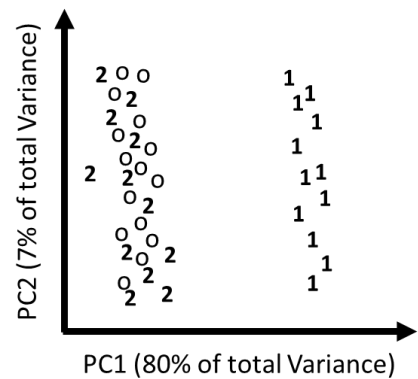
- a) Warum erscheint es sinnlos, diese Untergruppen durch Clustering der kompletten Genexpressionsprofile zu bestimmen? **(2 P)**
- b) Wie könnte man durch Data Mining, eventuell unter Hinzunahme zusätzlicher Daten, diesen Datensatz so vorverarbeiten, dass die anschließende Suche nach den Untergruppen mittels Clustering sinnvoll wird? **(2 P)**

11. Wie hoch ist die AUC eines Klassifikators, der keine Fehler macht? **(1 P)**

Name: _____

12. Nennen sie zwei Gründe für den Einsatz der Dimensionsreduktion bei der Analyse von grossen Datensätze? (*Stichwörter reichen aus*) (1 P)

13. Ein Spital hat eine Proteom-Analyse von Blutproben für drei im voraus bekannten Patientengruppen durchgeführt: Krankheit 1, Krankheit 2, und gesunden Kontrollen (o). Etwa 1000 Proteine wurden in jeder Probe gemessen. Eine Hauptkomponenten Analyse (PCA) liefert dieses Bild:



- a) Was interpretieren sie diese Ergebnisse? (1 P)
- b) Wie würden Sie vorgehen, um Proteom-Features zu identifizieren die zur Erkennung von Krankheit 1 dienen? (1 P)
- c) Aufgrund der PCA-Ergebnisse, ist es überhaupt sinnvoll nach Features zu suchen, die Patienten mit Krankheit 2 von Gesund und gleichzeitig von Krankheit 1 unterscheiden? Falls ja, wie würden Sie vorgehen um diese Features zu finden? Falls nein, warum? (2 P)

Name:

- 14.** Sie haben den Auftrag bekommen, die Funktion von etwa 100 Transkriptionsfaktoren (TFs) in einem *Escherichia coli* Stamm experimentell zu charakterisieren.
- a) Welches Experiment schlagen sie vor? Welche Mutanten lassen sie mit welcher Methode untersuchen? **(1 P)**
 - b) Welche Data Mining Methode würden Sie anwenden, um zu bestimmen welche zelluläre Prozesse von jedem einzelnen TF reguliert werden? Begründen Sie die Antwort. **(2 P)**
 - c) Beschreiben Sie (kurz) die Grundsätze der in b) gewählten Methode. **(1 P)**
 - d) Nachdem Sie die gesamte Analyse abgeschlossen haben, möchten Sie ihre Ergebnisse mit früheren Studien aus der Literatur vergleichen. Welcher Ansatz würden sie Verwenden um die Übereinstimmung zu quantifizieren? (*Stichwörter reichen aus*) **(1 P)**