

El proyecto utiliza un Algoritmo Genético para la optimización de parámetros en un modelo de sistemas biológicos. Busca los parámetros para que un sistema de Ecuaciones Diferenciales Ordinarias (EDOs) que representa una red de genes se comporte de una manera específica deseada. Al encontrar los parámetros óptimos, se está descubriendo la configuración de la red regulatoria génica (GRN) que puede llevar a la célula de un estado inicial a un estado final deseado, simulando así un proceso de reprogramación. La base de conocimiento con la que se está trabajando es fundamentalmente interdisciplinaria, combinando tres áreas principales:

1- Optimización Computacional con Algoritmos Evolutivos: El núcleo del proyecto es un Algoritmo Genético Elitista (EGA).

2- Biología de Sistemas y Modelado Matemático: El problema biológico se está abordando a través de un modelo de GRN, representado por un sistema de Ecuaciones Diferenciales Ordinarias (EDOs). Este enfoque permite describir la dinámica de las concentraciones de proteínas a lo largo del tiempo, basándose en las interacciones regulatorias entre los genes.

3- Ingeniería Celular y Reprogramación (como dominio de aplicación): El objetivo final del proyecto está inspirado en la reprogramación celular.

La función objetivo evalúa un conjunto de parámetros y retorna un valor de fitness escalar. El fitness combina tres componentes:

1. Distancia L2: Qué tan cerca está el resultado final del objetivo.

2. Penalización por Complejidad: Favorece soluciones con parámetros más pequeños (parsimonia: simplicidad, eficiencia, economía en el uso de recursos).

3. Recompensa por Alcance: Premia a las soluciones que alcanzan el objetivo rápidamente.

Un valor de fitness más bajo indica una mejor solución.

Lo que la función de fitness realmente evalúa son los parámetros del individuo (cuya validez está limitada por bounds) y compara la simulación resultante contra el target (Distancia L2), aplica penalidad por complejidad (función de los parámetros mismos) y añade recompensa por alcance temporal (depende de t_span , dt y $noise_std$ en la forma en que se detecta el logro).

"bounds" representan los rangos biológicamente plausibles para parámetros (p.ej. tasas de producción/degradación, coeficientes Hill, fuerzas de interacción). Limitar bounds evita soluciones no fisiológicas.

"target" es el perfil de expresión condrogénico deseado (niveles relativos de TFs/genes). Es el "fenotipo" objetivo que queremos alcanzar.

" t_span ", " dt " y " $noise_std$ " definen cuánto tiempo de simulación corresponde a un proceso de reprogramación (ej. etapas de diferenciación), la influencia en la precisión numérica y en la detección temprana de alcanzar el objetivo (impacta recompensa por alcance), y variabilidad intrínseca/ruido en expresión génica (ruido transcripcional, heterogeneidad celular), respectivamente.

Cuando se dijo que se aplica penalidad por complejidad (función de los parámetros mismos) se hace referencia a una penalización numérica que depende directamente del vector de parámetros que representa cada individuo: en la implementación que se está usando (evaluador toy de 3 factores) cada individuo es un vector de 9 parámetros organizados por TF en bloques de 3. En una GRN, los TFs son proteínas que regulan la expresión de genes al unirse a regiones específicas del ADN.

Combinaciones específicas de TFs determinan si una célula es NCC, condrocito, mesenquimal, etc. Los bloques de tres parámetros representan la tasa de producción (síntesis) del factor de transcripción, la tasa de degradación (degradación/clearance) del factor, y la escala de interacción/activación (en el toy model modula la fuerza con que la actividad total de la red activa la producción: actúa como un parámetro de sensibilidad / "ganancia").

tasa de producción: representa control transcripcional+translacional; experimentalmente equivalente a over-expression (\uparrow prod) o knockdown (\downarrow prod).

tasa de degradación: controla la vida media del TF; manipularlo es distinto de cambiar producción (p. ej. inhibir proteasas).

fuerza/sensibilidad de interacción: representa afinidad, eficacia de activación, o coeficientes Hill aproximados; puede ser positivo (activación neta) o negativo (efecto inhibitorio global, si el modelo lo admite).