

Programació i Algorismes 2 -- Grau en Intel·ligència Artificial

Professors: Jordi Delgado, José Luis Balcázar, Sergio Álvarez

Pràctica: Treball en equip, Maig de 2023

Enunciat per a grups de 3 persones.

Aquesta pràctica pretén que els estudiants, **en grups de tres persones**, realitzin un treball que els obligui a fer servir allò que han après durant el curs, principalment orientació a objectes i estructures dinàmiques de dades.

Entrega:

Els fitxers Python (es recomana un fitxer per cada classe) solució del problema plantejat i un informe detallant les decisions preses pels estudiants seran entregats no més tard del **divendres 16 de juny**. L'informe de l'estudiant és de format lliure, però hauria d'incloure: Justificació de les estructures de dades utilitzades, justificació de les corresponents implementacions (hi ha estructures de dades de les que els estudiants han vist diverses implementacions durant el curs), i jocs de prova creats per testar l'aplicació.

Enunciat:

En un laboratori de biologia estan fent experiments per intentar relacionar gens i trets ("*rasgos*" en castellà) d'individus pertanyents a espècies amb genomes relativament fàcils de manipular amb les eines d'edició genètica de les que disposa el laboratori.

Per simplificar els experiments, en comptes de tractar amb tot el genoma de cada individu, tractarem només amb un parell de cromosomes de cada individu. Tots els parells de cromosomes seran anàlegs, és a dir, ocuparan la mateixa posició (que no ens importa) dintre del genoma.

Tots els cromosomes d'un experiment correspondran a individus de la mateixa espècie, tindran la mateixa longitud m i constaran de m valors (zeros i uns) que ens indicaran si el gen corresponent és dominant o recessiu. Per a cada experiment, el laboratori ens proporcionarà el nombre d'individus de l'experiment $n \geq 3$, i el nombre de gens de cada cromosoma $m \geq 1$. Els individus estaran identificats pels valors $1, 2, \dots, n$. També ens indicaran les relacions de parentiu entre els diferents individus. Tots els individus tindran dos progenitors o cap. Els que no tenen progenitors seran els individus inicials, a partir dels quals es generaran la resta de d'individus de l'experiment. **Tots els individus tindran un únic fill, excepte un sol individu que no en tindrà cap.** Aquesta informació de parentiu es pot veure com l'arbre genealògic de l'únic individu sense fills.

Un cop que coneixem les relacions de parentiu entre els individus, ens proporcionaran la composició del parell de cromosomes que hem d'estudiar de cadascu d'ells. A continuació, s'aniran introduint els trets que es vagin observant en cada individu a mesura que el laboratori ens vagi indicant els resultats de les seves proves per a cada individu. De tant en tant, també es decidirà treure un tret d'algun individu perquè es creu que potser la seva inclusió estigui equivocada.

Un tret serà una paraula que necessàriament començarà amb una lletra i que podrà constar de lletres, nombres i el caràcter '_'. En tot moment s'ha de saber quins trets té cada individu i per cada tret s'ha de saber quins individus el presenten i quina és la combinació de gens que suposadament el produeix. Aquesta combinació de gens és la intersecció del parells de

cromosomes de tots els individus que presenten aquest tret. Quan s'afegeix el tret a un individu és possible que aquesta combinació sigui més restringida. Quan es treu un tret d'un individu és possible que passi tot el contrari.

Per entendre com va aquesta intersecció el més fàcil és un exemple. Estem tractant amb parells de cromosomes que tenen longitud 10. Suposem que només l'individu 5 i 8 tenen el tret *asdfg_23* i que el seus parells de cromosomes són pel que fa a l'individu 5:

```
1 0 1 0 1 0 1 0 1 0
1 0 1 0 1 0 1 0 1 1
```

i pel que fa a l'individu 8:

```
1 1 1 0 1 0 0 0 1 0
1 0 0 1 1 0 0 1 0 1
```

La intersecció dels dos parells de cromosomes queda així:

```
1 - - - 1 0 - - - 0
1 - - - 1 0 - - - 1
```

Això significa que pel moment, suposem que per tal que es manifesti el tret *asdfg_23* cal que el primer parell de gens sigui el 1 i 1; el cinquè també 1 i 1; el sisè 0 i 0 i el desè i darrer 0 i 1. La resta de parells de gens aparentment no intervenen en la manifestació d'aquest tret concret. Si més endavant es descobreix que algun individu té aquest tret és possible que aquesta intersecció sigui més restringida.

Quan s'afegeix per primer cop el tret, la combinació que el manifesta és el parell de cromosomes de l'individu a qui se li ha afegit. Si un tret només el té un individu i se li treu, aquest tret deixa d'estar present al sistema.

També s'ha de poder consultar les dades d'un individu a partir del seu identificador (un enter entre 1 i n) i les dades d'un tret a partir del seu identificador (que és el nom del tret).

Per últim, el laboratori ens demana poder estudiar la distribució d'un tret dins de l'arbre genealògic. Donat l'arbre genealògic i el nom d'un tret, volem el subarbre genealògic on apareguin tots els camins des de l'individu sense descendents fins a cada individu que manifesti aquest tret. Els nodes d'aquest subarbre que manifestin el tret hauran d'estar etiquetats amb l'identificador de l'individu i, els que no el manifestin, amb l'identificador de l'individu canviat de signe.

Detall de les operacions

Les operacions del nostre programa són les següents. Per veure els detalls concrets de format cal consultar el joc de proves públic.

- **experiment**: S'introduiran dos valors enters $n > 0$ i $m > 0$ que correspondran al nombre d'individus i al nombre de gens de cada cromosoma de cada parell de cada individu. Després s'introduirà en preordre l'arbre genealògic d'n nodes de l'únic individu que no té descendents¹. L'identificador d'aquest individu pot ser qualsevol valor entre 1 i n. Cada individu apareixerà un cop i només un cop. Per últim, s'introduiran les dades del parell de cromosomes de cada individu per ordre d'identificador d'individu.

¹ El format concret en que s'introdueix l'arbre genealògic i com llegir-lo s'explicarà a classe.

- **afegir_tret**: S'introduirà el nom d'un tret t i un enter $1 \leq id \leq n$. Si l'individu id ja té aquest tret s'escriurà **error**. Si no, quedarà constància de que l'individu id té el tret t i es calcularà la combinació de parells de cromosomes que fan que es manifesti el tret t .
- **treure_tret**: S'introduirà el nom d'un tret t i un enter $1 \leq id \leq n$. Si l'individu id no té aquest tret s'escriurà **error**. Si no, quedarà constància de que l'individu id ja no té el tret t . Si id era l'únic individu que tenia aquest tret, desapareixerà qualsevol informació relativa a aquest tret. Si hi ha altres individus que tenen aquest tret es calcularà la combinació de parells de cromosomes que fan que es manifesti el tret t .
- **consulta_tret**: S'introduirà el nom d'un possible tret. Si aquest no existeix s'escriurà **error**. Si existeix s'escriurà la combinació que suposadament fa que es manifesti i els identificadors dels individus que la manifestin per ordre creixent d'identificador.
- **consulta_individu**: S'introduirà un enter $1 \leq id \leq n$ i s'escriurà la informació de l'individu amb identificador id , és a dir, el seu parell de cromosomes i el nom dels seus trets per ordre alfabètic.
- **distribucio_tret**: S'introduirà el nom d'un possible tret. Si aquest no existeix s'escriurà **error**. Si existeix s'escriurà el subarbre resultat en inordre.
- **fi**: S'acaba l'execució del programa.

Observeu que en una execució del programa es poden realitzar molts experiments. A cada execució del programa es farà com a mínim un experiment. Per facilitar la llegibilitat s'escriuran totes les instruccions a la primera columna i tots els seus paràmetres separats per un espai. Qualsevol resultat s'escriurà a partir de la columna 3. Consulteu el joc de proves públic.