

# Demostraciones matemáticas

Federico Lebrón



A la Universidad de Buenos Aires.

# Índice

1	¿Qué es una demostración matemática?	4
1.1	Demostraciones en la historia	4
1.2	Contexto en el que están demostrando	7
1.3	Formalidad y rigor	9
2	¿Cuándo, qué, y para qué demostramos?	11
3	¿Cómo aprendemos a demostrar?	14
4	¿Cómo demostramos?	15
4.1	Formalizar la consigna	15
4.2	Comprender qué se nos pide	16
4.3	Considerar ejemplos	17
4.4	Estrategias de demostración	22
4.4.1	Inducción	22
4.4.2	Correctitud de ciclos en algoritmos	25
4.4.3	Correctitud de algoritmos recursivos	27
4.4.4	Definiciones equivalentes	28
4.4.5	Contrarecíproco y contradicción	29
4.4.6	Si y sólo si	31
4.4.7	Partir en casos	32
4.4.8	Unicidad	33
4.5	Pasar en limpio	34
5	Errores comunes	37
5.1	Ser informal	37
5.2	No decir nada	39
5.3	Empezar con la conclusión	40
5.4	No entender qué estamos asumiendo	41
6	Ejemplos	41
6.1	Conjuntos	42
6.2	Funciones y análisis asintótico	44
6.3	Sucesiones y series	53
6.4	Combinatoria	60
6.5	Divide and conquer y programación dinámica	66
6.6	Greedy	83
6.7	Árboles	99
6.8	Caminos mínimos	104
6.9	Planaridad	120
6.10	Coloreo	123
6.11	Homomorfismo e isomorfismo de grafos	127
6.12	Circuitos y caminos	129
6.13	Flujo	133
7	Ejercicios	135
	Bibliografía	136

# 1 ¿Qué es una demostración matemática?

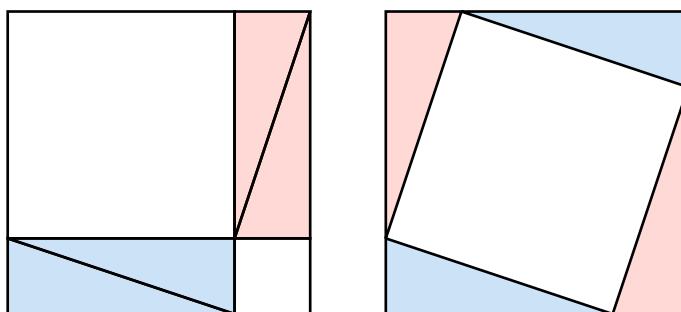
Una demostración matemática es un argumento convincente de la veracidad de una proposición matemática. Inmediatamente tenemos la pregunta, ¿convence a *quién*? Distintos argumentos van a convencer a distinta gente. Por ejemplo, el siguiente es un argumento de la veracidad del teorema de Pitágoras.

## Teorema 1 (Teorema de Pitágoras)

Dado un triángulo rectángulo, la suma de los cuadrados de los catetos es igual al cuadrado de la hipotenusa.



*Demostración.*



¿Esto es siquiera una demostración? ¿Los convence? ¿Convencería a sus amigos o a alguien que parase en la calle? ¿Cómo saben, o definen, si «está bien»?

## 1.1 Demostraciones en la historia

Históricamente, empezamos haciendo matemática sin generalidad, y sin demostraciones. Un matemático en el 2000 a.c.<sup>1</sup> hubiera afirmado una proposición, y listo. Con suerte hubiera dado ejemplos que mostraran la veracidad de esa proposición en casos particulares. No había un lenguaje formal para hablar en generalidades sobre, por ejemplo, «los enteros». Había tablas de recíprocos de enteros y soluciones a ecuaciones cuadráticas, pero no un algoritmo de división o una expresión como  $\frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ .

Los griegos son los primeros que intentan dar argumentos de algún tipo para proposiciones generales. Thales de Mileto demostró el siguiente teorema, aunque la primer demostración que sobrevive es la de Euclides.

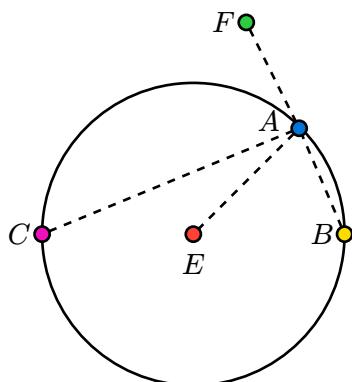
## Teorema 2 (Teorema de Thales)

Si tenemos tres puntos  $A$ ,  $B$ , y  $C$  en un círculo, donde la línea  $\overline{AB}$  pasa por el centro del círculo, entonces  $\angle ABC$  es un triángulo rectángulo.



<sup>1</sup>No había un concepto de «matemático» en ese momento. Los babilonios entendían conceptos sobre aritmética y álgebra desde el punto de vista de la geometría. Describían relaciones entre ángulos, lados, y perímetros de figuras dibujadas, y usaban esto para construir, arar, estimar distancias, etc.

Demostración.



Dada por Euclides[1].

Sea  $ABC$  un círculo,  $\overline{BC}$  su diámetro,  $E$  su centro. Unir  $\overline{BA}$ , y  $\overline{AC}$ .

Digo que el ángulo  $\angle BAC$  es rectángulo.

Unir  $\overline{AE}$ , y extender  $\overline{BA}$  hasta  $F$ . Como  $\overline{BE}$  es igual a  $\overline{EA}$ , el ángulo  $\angle ABE$  también es igual al ángulo  $\angle BAE$ . Como  $\overline{CE}$  es igual a  $\overline{EA}$ , el ángulo  $\angle ACE$  es igual al ángulo  $\angle CAE$ .

Luego, el ángulo  $\angle BAC$  es igual a la suma de los dos ángulos  $\angle ABC$  y  $\angle ACB$ . Pero el ángulo  $\angle FAC$  exterior al triángulo

$ABC$  también es igual a la suma de los dos ángulos  $\angle ABC$  y  $\angle ACB$ . Luego el ángulo  $\angle BAC$  también es igual al ángulo  $\angle FAC$ . Luego ambos son rectángulos. Luego el ángulo  $\angle BAC$  es rectángulo.

□

¿Los convence esa demostración? ¿Por qué  $\angle FAC$  es igual a la suma de  $\angle ABC$  y  $\angle ACB$ ? En su momento, esta demostración no sólo era convincente, sino que fue parte del libro de demostraciones de geometría más famoso y celebrado de la historia. ¿Por qué puede ser que hoy en día nos resulta confusa?

Viajamos luego al 1758, donde en «De numeris qui sunt aggregata duorum quadratorum» («Sobre números que son la suma de dos cuadrados») Leonhard Euler prueba el siguiente teorema.

### Teorema 3

Si  $p$  y  $q$  son dos números, cada uno de los cuales la suma de dos cuadrados, entonces el producto  $pq$  es también una suma de cuadrados.

♥

Demostración. Sea  $p = aa + bb$  y  $q = cc + dd$ . Tendremos que  $pq = (aa + bb)(cc + dd) = aacc + aadd + bbcc + bbdd$ , que se puede representar de manera tal que  $pq = aacc + 2abcd + bbdd + aadd - 2abcd + bbcc$  y por tal razón  $pq = (ac + bd)^2 + (ad - bc)^2$ , de donde el producto  $pq$  será la suma de dos cuadrados.

□

¿Qué les parece esa demostración? ¿Es más fácil de entender que las dos anteriores? ¿Los convence? ¿Les parece que convencería a *cualquiera*? ¿Qué asume esta demostración? ¿Qué significa «número» acá<sup>2</sup>? Si mis números son las horas del reloj, y estoy haciendo aritmética modular módulo 12, ¿sigue valiendo el teorema? ¿Y si mis «números» son funciones de  $\mathbb{R}$  a  $\mathbb{C}$ ? ¿Y si son elementos de punto flotante<sup>3</sup> en una computadora, sigue valiendo, o se rompe algo que está asumiendo?

Ahora veamos una demostración formal, escrita en el lenguaje de programación Lean 4.

<sup>2</sup>El concepto formal de número real que usamos hoy en dia se definió en 1872, en paralelo por Richard Dedekind y Georg Cantor. Euler y sus matemáticos contemporaneos usaban nociones no totalmente formales, llegando a veces a paradojas.

<sup>3</sup>Siguiendo el standard de IEEE 754, por ejemplo punto flotante de 32 bits.

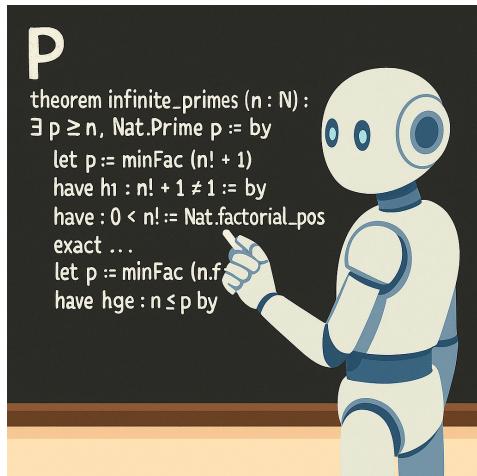
#### Teorema 4

Para todo  $n \in \mathbb{N}$ , existe un número primo más grande que  $n$ .



Demostración.

```
-- Primo (p : N) : Prop
theorem hay_infinitos_primos (n : N) : ∃ p, n ≤ p ∧ Primo p :=
let fac := n ! + 1
-- minFac (n : N) : N, el mínimo factor de n
let p := minFac fac
-- factorial_pos (n : N) : 0 < n !
-- succ_lt_succ {n m : N} : n < m → n + 1 < m + 1
-- ne_of_gt {a b : N} (h : b < a) : a ≠ b
have f1 : fac ≠ 1 := ne_of_gt <| succ_lt_succ <| factorial_pos n
-- minFac_es_primo {n : N} (n1 : n ≠ 1) : Primo (minFac n)
have pp : Primo p := minFac_es_primo f1
-- le_of_not_ge: {α: Tipo} {a b : α} : ¬(a ≤ b) → b ≤ a
have np : n ≤ p :=
le_of_not_ge fun h =>
-- dvd_factorial {m n : N} : 0 < m → m ≤ n → m | n !
-- minFac_pos (n : N) : 0 < minFac n
have h1 : p | n ! := dvd_factorial (minFac_pos fac) h
-- dvd_suma {k m n : N} (h : k | m) : k | n ↔ k | m + n
-- minFac_dvd (n : N) : (minFac n) | n
have h2 : p | 1 := (dvd_suma h1).2 (minFac_dvd fac)
-- no_divide_uno {p : N} (pp : Primo p) : ¬(p | 1)
no_divide_uno pp h2
(p, np, pp)
```



La correctitud de esta demostración es verificable automáticamente por una computadora. Esta demostración, ¿los convence? ¿Siquiera la pueden leer? Indudablemente es correcta, formal, y rigurosa (porque las computadoras no nos «creen» nada), ¿eso significa que «está bien»? ¿Ustedes escribirían algo así para comunicarle algo a un par? Si estas demostraciones se pueden verificar formalmente, ¿por qué no escribimos todas las demostraciones así? Esto nos dice que la formalidad no implica legibilidad para nuestra audiencia.

Por último, veamos un teorema moderno, traducido del libro Algebra de Serge Lang. Este es un libro de estudio clásico para álgebra abstracta, y el teorema se ve en la carrera de Matemática en la facultad.

### Teorema 5 (Teorema de Jordan-Hölder)

Sea  $G$  un grupo, y sea

$$G = G_1 \supset G_2 \supset \dots \supset G_r = \{e\}$$

sea una torre normal tal que cada grupo  $G_i/G_{i+1}$  es simple, y  $G_i \neq G_{i+1}$  para  $i = 1, \dots, r-1$ . Entonces cualquier otra torre normal de  $G$  teniendo las mismas propiedades es equivalente a esta.

♥

*Demostración.* Dado cualquier refinamiento  $\{G_{i,j}\}$  para nuestra torre, observamos que para cada  $i$  existe precisamente un único índice  $j$  tal que  $G_i/G_{i+1} = G_{i,j}/G_{i,j+1}$ . Luego esta secuencia de factores no-triviales para la torre original, o la torre refinada, es la misma. Esto prueba nuestro teorema. □

¿Qué tal esta última? ¿Los convence? ¿Les parece clara la demostración? Imagino que a la mayoría les va a resultar incomprendible. Si le está faltando algo para convencerlos, ¿significa que esta demostración «está mal»?

Con estos ejemplos concluimos que lo que se considera una demostración matemática ha evolucionado, como ha evolucionado la noción de matemática que usamos. También, lo que se considera una demostración correcta va a depender del contexto en que uno se encuentre, qué puede uno asumir del lector, y para qué propósito uno está demostrando.

## 1.2 Contexto en el que están demostrando

Ustedes se encuentran cursando una carrera universitaria en las ciencias formales. El objetivo de la carrera es formarlos como computadores científicos. Como tales, deberían salir de la carrera ambos con conocimientos sobre objetos y herramientas específicas como grafos, números racionales, programación dinámica, sistemas operativos, inducción, cálculo diferencial, teoría de lenguajes, y algoritmos numéricos; pero más importantemente **deberían adquirir la habilidad de razonar formalmente**, de demostrar a cualquiera que se los pida que sus conclusiones son correctas, de entender y criticar razonamientos de otros, de saber cuándo y cómo usar las herramientas que conocen, y cuándo y cómo desarrollar ideas y algoritmos nuevos.

Para la primer parte, el temario de la carrera incluye esos temas en distintas materias. Para la segunda es que se los entrena en demostrar formalmente la veracidad de proposiciones, y la correctitud<sup>4</sup> de argumentos. Se usan los objetos de estudio como grafos o números enteros como sujetos de las proposiciones y algoritmos que desarrollen.

Es por esto que la pregunta clásica de «¿pero cuándo voy a usar esto?» está mal planteada. Asume que se les enseña sobre grafos porque alguien les va a «dar un grafo» en su vida profesional. En cambio, se les enseña sobre grafos para que ustedes los introduzcan para modelar problemas que tengan, y para que los usen como sujetos en proposiciones al aprender a demostrar. Podrían estos sujetos ser otros, como fluidos y campos eléctricos en física, o anillos y ecuaciones diferenciales en matemática. En computación, se usan los objetos de ese campo de estudio.

<sup>4</sup>Nota pedante: La RAE no reconoce el término «correctitud», sino «corrección». Sin embargo, el término «corrección» tiene otros significados que pueden confundirlos en el contexto de una carrera universitaria (e.g. corrección de un parcial, corrección de una aproximación numérica). Es común, entonces, usar el término «correctitud» para referirse a la propiedad de ser correcto. Es la diferencia en Inglés entre «correctness» y «correction», que la RAE no hace.

Sus demostraciones, entonces, cumplen un doble propósito. Tienen que:

1. Convencer al lector de la veracidad de la proposición que afirman. Esto es común a todas las demostraciones matemáticas.
2. Convencer al docente que entienden cómo convencer a *cualquier persona*. El docente ya sabe que la proposición es cierta, no se les va a pedir demostrar proposiciones falsas. Luego, ya está convencido/a de la veracidad de la proposición antes de empezar a leer lo que escribieron. El docente va a evaluar si sus argumentos pueden convencer a *cualquiera*.

Como vimos, *cualquiera* está definido dentro de un contexto. Sus argumentos no van a convencer a un bebé de seis meses, porque no los puede leer. Muchas veces tampoco van a convencer a alguien que no hable español.



¿A quién están convenciendo, entonces? Al interlocutor para el cual la carrera los prepara: Un par suyo, de la comunidad científica. Estamos formando científicos, después de todo. Llegamos entonces a la siguiente definición.

#### Definición 1.2.1 (Demostración)

Una demostración es un argumento formal de la veracidad de una proposición, que puede convencer a *cualquier* par suyo en la comunidad científica.

Esto nos dice que tenemos que ser un poco paranoides, si nuestro objetivo es convencer a *cualquier* par. Si somos imprecisos, nuestro lector puede no comprender lo que quisimos decir, y luego no estar convencido. Si no demostramos una proposición que usamos, nuestro lector puede no creernos que esa proposición es cierta, y luego no estar convencido. Tenemos no sólo que estar seguros de la veracidad de lo que estamos probando, sino también saber comunicar las razones por las que *cualquier* par tiene que estar de acuerdo, inclusive si nuestro par nos odia, si nunca vio la proposición antes, si no sabe si es cierta o no antes de empezar a leernos, o si es su primera vez viendo el objeto de estudio sobre el cual estamos hablando (números naturales, programas imperativos, árboles, lo que sea).

Por otro lado, uno puede valerse de que el lector es un par. Sus conocimientos son similares a los que tienen ustedes, y al momento de estar cursando una materia, pueden asumir que el lector cursó y aprobó las materias correlativas. Por ende, si sabemos que  $m \leq \frac{n(n-1)}{2}$ , podemos concluir que  $m < n^2$ , porque asumimos que el lector aprobó álgebra del secundario. Debemos *decir* que estamos usando que  $m \leq \frac{n(n-1)}{2}$  para concluir que  $m \leq n^2$ , pero no hace falta demostrar esa conclusión. Si debemos decir quién es  $n$  y quién es  $m$ , y si estamos usando, por ejemplo, que  $n \geq 0$ , debemos afirmarlo explícitamente. De nuevo: Estamos siendo un poco paranoides, porque no queremos que quede ninguna duda en la mente de ningún par que nos lea (en particular, en la mente del docente que nos va a evaluar).

### 1.3 Formalidad y rigor

Cuando uno habla de «formalidad» en matemática, se está refiriendo a la *forma* en la que algo está escrito. Las demostraciones, entonces, existen en un continuo de formalismo. Ese formalismo va desde argumentos heurísticos, como « $n! + 1$  no es divisible por nadie debajo de  $n$ , luego es primo», hasta demostraciones verificables por un asistente de demostración, como la que vimos en el Teorema 4. El formalismo, en el contexto en el que están, se usa para intentar ser riguroso en nuestros argumentos. Un argumento informal como el primero puede ser cierto, como puede ser falso (en particular,  $n! + 1$  no siempre es primo, por ejemplo  $4! + 1 = 25 = 5^2$ ).

Razonar informalmente es parte de hacer matemática y computación, pero es sólo la primer parte. Al intentar argumentos, primero los vamos a pensar de forma vaga, no formal. Luego, si queremos asegurarnos de su veracidad, los formalizamos, para estar seguros de ser rigurosos.

Heurístico	El número de subconjuntos de un conjunto de $n$ elementos es $2^n$ , porque cada cosa puede o estar o no estar.
Poco formal	Como vemos en el siguiente dibujo, siempre vamos a tener suficientes aristas para que $u$ y $v$ tengan un vecino en común.
Razonablemente formal	Queremos ver que todo grafo con $n$ vértices tiene a lo sumo $\frac{n(n-1)}{2}$ aristas. Toda arista une dos vértices distintos, y a lo sumo hay una arista entre cada par de vértices distintos. Hay exactamente $\frac{n(n-1)}{2}$ pares de vértices distintos. Luego hay a lo sumo $\frac{n(n-1)}{2}$ aristas en total.
Obviamente formal	<p>Queremos ver que <math>f : \mathbb{R} \rightarrow \mathbb{R}</math>, <math>f(x) = 2x</math> es diferenciable en todo su dominio. Sea <math>x_0 \in \mathbb{R}</math>. Vemos que <math>\lim_{h \rightarrow 0} \frac{2(x_0+h)-2x_0}{h} = \lim_{h \rightarrow 0} \frac{2x_0+2h-2x_0}{h} = 2</math>. Luego, la derivada existe en <math>x_0</math>, y es igual a 2. Luego <math>f</math> es diferenciable en todo su dominio.</p> <hr/> <p>Sea <math>X</math> un conjunto con <math> X  = n</math>, y sea <math>\mathcal{P}(X) = \{A \subseteq X\}</math> su conjunto de partes. Para cada <math>k \in \{0, \dots, n\}</math>, sea <math>\mathcal{P}_k(X) = \{A \subseteq X \mid  A  = k\}</math>. Luego, <math>\mathcal{P}(X) = \bigsqcup_{k=0}^n \mathcal{P}_k(X)</math>. Entonces,</p> $\begin{aligned}  \mathcal{P}(X)  &= \sum_{k=0}^n  \mathcal{P}_k(X)  \\ &= \sum_{k=0}^n \binom{n}{k} \end{aligned}$ <p>Asimismo, <math> \mathcal{P}(X)  = 2^n</math> por la proposición 7. Luego, <math>\sum_{k=0}^n \binom{n}{k} = 2^n</math>.</p>
Extremadamente formal	Queremos probar que $1 + 1 = 2$ . Los naturales $\mathbb{N}$ están definidos inductivamente como:

- $0 \in \mathbb{N}$ .
- $\forall x. x \in \mathbb{N} \Rightarrow S(x) \in \mathbb{N}$ .

	<p>La suma entre naturales está definida inductivamente como:</p> $a + 0 = a$ $a + S(b) = S(a + b)$ <p>Notamos por conveniencia <math>1 = S(0)</math>, y <math>2 = S(S(0))</math>. Luego,</p> $\begin{aligned} 1 + 1 &= S(0) + S(0) \\ &= S(S(0) + 0) \\ &= S(S(0)) \\ &= 2 \end{aligned}$ <p>Luego <math>1 + 1 = 2</math>, que es lo que queríamos demostrar.</p>
Totalmente formal	<pre>inductive N where   Z : N   S (n : N) : N open N def suma (m n : N) : N := match n with   Z =&gt; m   S n =&gt; S (suma m n) def uno : N := S Z def dos : N := S (S Z) theorem teorema_de_lebron : suma uno uno = dos := rfl</pre>

Lo que se espera de ustedes es que puedan escribir y leer demostraciones entre los niveles «Razonablemente formal» y «Obviamente formal».

Mencioné dos veces la palabra «rigor», pero ¿qué significa? En el contexto de demostraciones matemáticas, el rigor significa que **el lector debería poder seguir la demostración paso a paso**, y no preguntarse ¿y esto por qué vale? a cada momento.

### ! Importante

Mientras que la formalidad se refiere a **la forma en la que escribimos**, el rigor se refiere a **la implicación lógica de nuestras oraciones**.

Acá vemos otra vez que el contexto es importante. Podemos asumir, al momento de escribir demostraciones en una materia universitaria, que el lector ha completado las materias correlativas a la que están cursando. Por ejemplo, si ya vieron que diferenciabilidad implica continuidad, pueden asumir que el lector sabe eso, y no tienen que demostrarlo.

Veamos un ejemplo de una demostración no rigurosa:

#### Ejercicio 1.3.1

Sea  $A \subseteq \mathbb{N}$ ,  $A \neq \emptyset$ . Demostrar que existe un  $a \in A$  tal que para todo  $b \in A$ ,  $a \leq b$ .

*Demostración.* Supongamos por contradicción que  $A$  no tiene un tal mínimo elemento. Sea  $x_0$  cualquier elemento de  $A$ . Como  $A$  no tiene un mínimo elemento, sea  $x_1 \in \mathbb{N}$  tal que  $x_0 > x_1$  (si  $x_1$  no existiera,  $x_0$  sería menor o igual a todo elemento en  $A$ , que asumimos no sucede).

Como  $A$  no tiene un mínimo elemento, sea  $x_2 \in \mathbb{N}$  tal que  $x_1 > x_2$ . Siguiendo de esta manera tenemos una sucesión infinitamente decreciente de naturales, que no puede suceder. Por lo tanto, lo que asumimos debe haber sido falso, y  $\mathbb{N}$  tiene un mínimo elemento.  $\square$

El lenguaje en el que está escrito esto es razonablemente formal. Sin embargo, esta demostración no es rigurosa. Un lector se va a preguntar qué exactamente está pasando cuando decimos «Siguiendo de esta manera». Estamos haciendo alusión a un proceso implícito, y usando alguna noción no definida de límite. Después de todo, todos los procesos que podríamos enumerar en una demostración son finitos, entonces hacer alusión a que primero elegimos  $x_0$ , luego  $x_1$ , luego  $x_2$ , «...», esconde la dificultad en explicitar exactamente qué es ese «...».

Una vez más, esto depende del contexto. Los antiguos Griegos usaban este tipo de argumentos todo el tiempo, y esa demostración hubiera sido considerada rigurosa. Fue sólo a principios del 1900, con el trabajo de Ernst Zermelo[2], que nos dimos cuenta que ese tipo de razonamientos, si no tenemos cuidado, llevan a paradojas.

## 2 ¿Cuándo, qué, y para qué demostramos?

Un computador científico escribe demostraciones cuando quiere establecer sin dudas la veracidad de una proposición lógica. Por ejemplo, si queremos probar que nuestro sistema no va a quedarse sin memoria independientemente de las consultas que arriven, si queremos probar que nuestro programa no se va a ralentizar desmedidamente a medida que se aumente el tamaño de su entrada, o si queremos probar que no va a haber «deadlock» en ninguna circunstancia.

No está muy lejos de la verdad el decir que **la ciencia de la computación sin demostraciones es ingeniería**<sup>5</sup>. Vamos a ver luego cómo podemos usar herramientas formales para referirnos a los algoritmos que escribimos, y demostrar propiedades sobre los mismos.

```

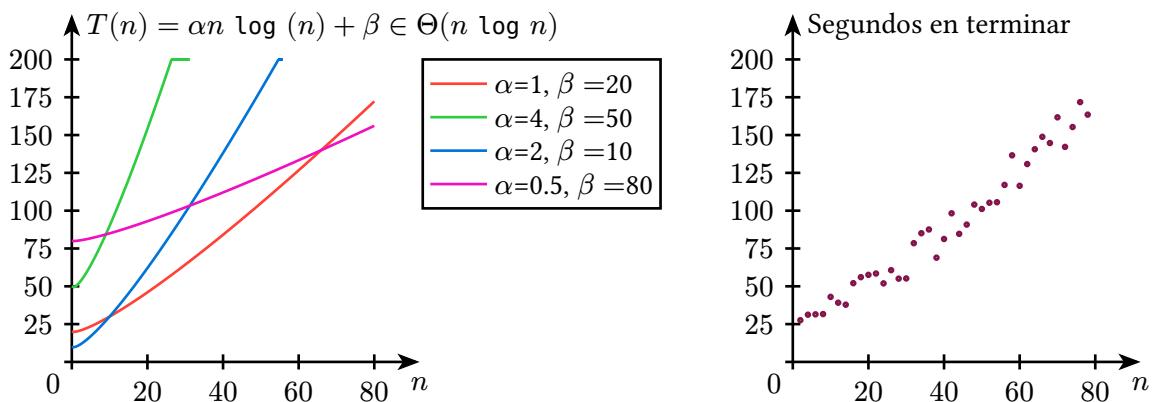
1: procedure FILOSOFO( $i, n$ , Palitos)
2:   while TRUE do
3:     PENSAR()
4:     AGARRAR-PALITO(Palitos[ $i$ ])
5:     AGARRAR-PALITO(Palitos[( $i + 1$ ) mod  $n$ ])
6:     COMER()
7:     SOLTAR-PALITO(Palitos[( $i + 1$ ) mod  $n$ ])
8:     SOLTAR-PALITO(Palitos[ $i$ ])
9:   end
10: end

```

No todas las cosas que querremos saber las vamos a demostrar formalmente. El motivo por el cual la computación es una ciencia, y no sólo matemática, es que hay propiedades de nuestro sistema que vamos a evaluar con argumentos prácticos (por ejemplo, tomando mediciones de nuestro sistema). A veces vamos a usar ambas cosas: Vamos a ver que nuestro programa se comporta bien en tamaños pequeños, y luego mostramos que su comportamiento asintótico es bueno. Al usar nociones asintóticas de complejidad, nos van a quedar constantes sin resolver analíticamente. Para encontrar esas constantes, vamos a medir el sistema real.

---

<sup>5</sup>Un ingeniero respondería, «la ingeniería irrelevante y pedante es ciencia de la computación».



Cuando decidimos demostrar algo formalmente, entonces, es porque realmente queremos concluir algo con total seguridad. No nos alcanza con argumentos heurísticos, como mirar qué pasa con «*n* chico», o verificar que es cierto para todos los casos que se nos ocurre.

Recordando el contexto en el que están como alumnos de una carrera universitaria, el otro motivo es, como dijimos antes, convencer a su docente que pueden convencer a cualquier par. Esto a veces va a requerir explicitar argumentos en más detalle que lo que esperan. Veamos un ejemplo de la diferencia. El siguiente es un ejercicio de la práctica 1, y la demostración dada por un alumno, verbatim.

### Ejercicio 2.1

Calcule la complejidad de un algoritmo que utiliza  $T(n)$  pasos para una entrada de tamaño  $n$ , donde  $T$  cumple:

$$T(n) = 2T(n - 4)$$

*Demostración.*

$$\begin{aligned} T(n) &= 2T(n - 4) = 2(2T(n - 4 - 4)) \\ &\dots = 2^i T(n - 4i) \end{aligned}$$

Como  $n - 4i = 1 \Leftrightarrow i = \frac{n-1}{4}$ . Luego,

$$= 2^{\frac{n-1}{4}} T(1) = \left(2^{\frac{1}{4}}\right)^n 2^{-\frac{1}{4}} = O\left(\left(2^{\frac{1}{4}}\right)^n\right)$$

□

¿Qué les parece esa demostración? ¿Los convence? ¿Les parece que sería marcado correcto en un parcial?

Esta demostración a mi me convenció al leerla. Si me la presentara un compañero en el trabajo, estaría de acuerdo que  $T \in O\left(\left(2^{\frac{1}{4}}\right)^n\right)$ . Sin embargo, alguien podría preguntar, «estás restando de a 4 a  $n$  hasta llegar a 1 pero, ¿y si  $n$  no es congruente con 1 módulo 4?», o «¿cómo sabés qué  $T(1) = 1$ ?». El alumno asumió que  $T(1) = 1$ , y no pensó realmente en una inducción formal, ni en una definición precisa de  $T$  (que hubiera requerido definir su dominio, y asignarle a cada elemento de tal dominio un valor del codominio).

Veamos otra demostración, un poco más formal.

*Demostración.* Sea  $T : \mathbb{N} \rightarrow \mathbb{N}$ . Asumo que  $T(n) = 2T(n - 4)$  para todo  $n \geq 4$ . Sea  $a = \max(T(0), T(1), T(2), T(3))$ .

Definimos  $P(n) : T(n) \leq a2^{\frac{n}{4}}$ . Vamos a probar que  $P(n)$  es cierta para todo  $n \in \mathbb{N}$ .

Como nuestra recurrencia usa un valor de  $n$  menor por 4 que el que recibe, tenemos que probar cuatro casos base, donde no tiene sentido restar 4 a  $n$ , pues saldríamos de  $\mathbb{N}$ .

Para todo  $0 \leq n \leq 3$ , tenemos que  $T(n) \leq \max(T(0), T(1), T(2), T(3)) = a$ . Como  $2^{\frac{n}{4}} \geq 1$  para  $0 \leq n \leq 3$ , y  $T(n) \leq a$ , tenemos que  $T(n) \leq a2^{\frac{n}{4}}$ , que prueba nuestros cuatro casos base.

Ahora el paso inductivo. Asumo  $P(k)$  para todo  $k < n$ , quiero probar  $P(n)$ . Si  $n \leq 3$ , ya probamos los casos base arriba, y vale  $P$  para ellos. Si  $n \geq 4$ , entonces como sabemos,  $T(n) = 2T(n - 4)$ . Como  $0 \leq n - 4 < n$ , podemos usar la hipótesis inductiva,  $P(n - 4)$ , obteniendo que  $T(n - 4) \leq a2^{\frac{n-4}{4}}$ . Entonces, como  $2^{\frac{n-4}{4}} = 2^{\frac{n}{4}}2^{-1}$ , vemos que  $T(n) = 2T(n - 4) \leq 2a2^{\frac{n}{4}}2^{-1} = a2^{\frac{n}{4}}$ . Esto es precisamente  $P(n)$ , que es lo que queríamos demostrar.

Como sabemos que  $T(n) \leq a2^{\frac{n}{4}}$  para todo  $n \in \mathbb{N}$ , podemos usar la definición de  $O$ , que es que  $f \in O(g) \Leftrightarrow \exists \alpha \in \mathbb{R}_{\geq 0}, n_0 \in \mathbb{N}$ , tal que  $\forall n \in \mathbb{N}, (n \geq n_0 \Rightarrow f(n) \leq \alpha g(n))$ .

Podemos acá elegir  $\alpha = a$ ,  $n_0 = 0$ ,  $g(n) = 2^{\frac{n}{4}}$ , y vemos que  $T \in O(g)$ .  $\square$

¿Qué les pareció esa demostración? Es más larga. Es más detallada. La anterior no «estaba mal», sólo no muestra que el escritor entiende los conceptos que se están evaluando (en este caso, inducción, comportamiento asintótico, y funciones recursivas). Al no tener en cuenta los detalles difíciles sobre inducción en esta demostración, el docente no puede tener confianza que el alumno sabe hacer esto bien, y no va a cometer un error por olvidárselos. De nuevo: Nunca se les va a pedir probar algo falso, por lo cual no importa que sea *cierto* que  $T \in O(2^{\frac{n}{4}})$ . Lo que importa es si lo saben demostrar.

Para un computador científico, la segunda demostración tiene otro «sabor». Al leerla, nos lleva de la mano, de paso a paso, explicitando cada uno. Uno llega a la conclusión con una seguridad de que cada paso está bien fundado, sin tener que adivinar qué quiere decir cada oración, y sin tener que preguntarse «¿y qué pasa si tal cosa no se cumple?». Las demostraciones que ustedes escriban tienen que dejarlos, y dejar al lector, con la misma sensación. Parece poco serio lo que estoy diciendo, pero realmente es una buena guía para saber cuándo están haciendo las cosas bien. Esa sensación la van a afilar practicando, haciendo demostraciones, recibiendo correcciones de sus docentes, cometiendo errores, viendo dónde les faltó definir algo, asumieron algo, no se acordaron de un caso, o no entendieron la consigna.

Finalmente, la longitud de la demostración no es algo a intentar emular en sí. De hecho, lo contrario es cierto: si pueden ser precisos y concisos, ¡mejor! Es muy común, sin embargo, que erren por el otro lado, haciendo demostraciones extremadamente escuetas, de una o dos oraciones, pretendiendo que eso le demuestre al lector lo que dijimos que una demostración tiene que mostrar. A veces es porque piensan que, como no saben probar algo, mejor ni lo mencionan y «si pasa pasa». Otras veces es porque no se dan cuenta que están asumiendo algo. Otras es porque no entendieron qué se les está pidiendo demostrar. En las próximas secciones vamos a ver errores clásicos y cómo no cometerlos.

### 3 ¿Cómo aprendemos a demostrar?

Nadie aprendió a andar en bicicleta viendo a otros andar en bicicleta. Tampoco van a aprender a demostrar leyendo cómo alguien más demostró. La **única** forma que van a aprender es escribiendo demostraciones ustedes mismos. Por cada segundo que pasan leyendo este documento, sugiero que pasen cinco pensando y escribiendo sus propias demostraciones.

Empiecen demostrando proposiciones simples. Si intentan ambos aprender un tema nuevo (como teoría de grafos) y *al mismo tiempo* aprender a demostrar, les va a resultar demasiado difícil, se van a confundir, y frustrar. Demuestren, primero, propiedades de conjuntos, de enteros, de racionales, de matrices, de objetos con los que ya saben trabajar.

#### Teorema 6 (Teorema de Euclides)

Hay un número infinito de primos.



*Demostración.* Esta demostración se encuentra en «Elementos» [1].

Sea  $L = [p_1, \dots, p_n]$  una lista finita de números primos distintos. Vamos a mostrar que existe algún número primo que no está en esta lista.

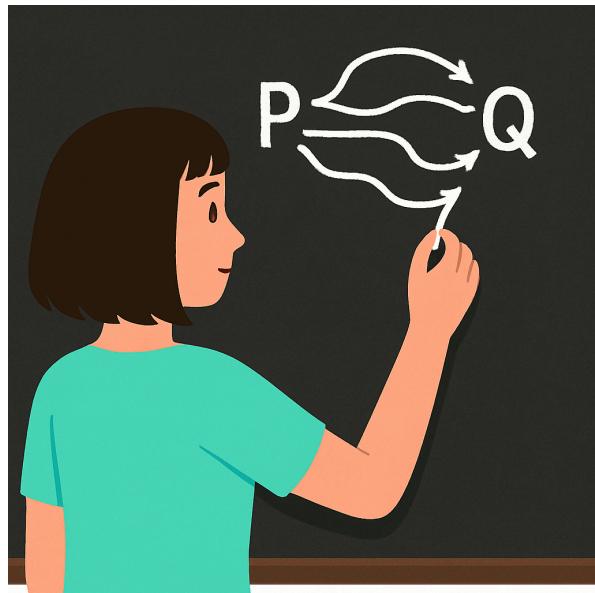
Sea  $P = \prod_{p \in L} p$  el producto de todos los números en  $L$ . Sea  $q = P + 1$ . Entonces o bien  $q$  es primo, o no es primo.

1. Si  $q$  es primo, entonces como  $q > p$  para todo  $p \in L$ , al menos hay un número primo ( $q$  mismo) fuera de  $L$ .
2. Si  $q$  no es primo, entonces hay algún factor primo  $p$  que divide a  $q$ . Entonces  $p \in L$ , o  $p \notin L$ .
  1. Si  $p \notin L$ , encontramos un número primo ( $p$ ) que no está en  $L$ .
  2. Si  $p \in L$ , entonces  $p \mid P$ , pues  $P$  es el producto de todos los números en  $L$ . Como  $p \mid P$  y  $p \mid q$ , entonces  $p \mid (q - P)$ , es decir  $p \mid 1$ . Como ningún número primo puede dividir a 1, este caso no puede suceder.

Concluimos que para toda lista finita  $L$  de números primos, existe algún número primo que no está en  $L$ , y luego hay infinitos números primos. □

En este momento, muchos de ustedes no tienen todavía una intuición sobre cuándo una demostración es correcta, versus cuándo les están mintiendo. Más aún, tienen sesgos entendibles por su educación: si leen una demostración en un libro, «debe estar bien», o si se los dice un docente, «debe estar bien». Lo que el practicar les va a dar es la confianza de decir «No, no creo que lo que esté diciendo el docente / libro esté bien.», así como también «Estoy seguro/a de que el argumento que acabo de escribir es correcto.» La matemática es el único lugar donde podemos tener esta certeza, ninguna ciencia puede tenerla. ¡Aprovechémolas! **No me crean, ni le crean a ningún docente, nada.** Si no lo pueden demostrar, no saben si es cierto, y no deberían usarlo.

La matemática que van a ver en la universidad no es como la que vieron en secundario. No hace falta seguir una receta particular, hacer cálculos larguísimos de memoria como derivar 50 polinomios de grado 9, o memorizar patrones de demostración y usar los que el docente quiera. Si se les pide demostrar que  $P$  implica  $Q$ , el objetivo es que ustedes produzcan una demostración clara y convincente de ese hecho. Si el docente pensó que la demostración tenía que ser por inducción, y ustedes la hicieron partiendo en casos, por contrarecíproco, o por absurdo; no importa. Tienen una enorme flexibilidad a la hora de argumentar, lo que da lugar a la **creatividad** en matemática.



Finalmente, al momento de empezar una demostración, sepan que les debería tomar tiempo. Si esperan que las demostraciones les salgan en 10 minutos, van a cometer errores. Es frecuente que lleven horas. No está *mal* que tomen ese tiempo, y no se deberían sentir mal cuando lo hacen. Tengan paciencia, practiquen con tiempo, entiendan que el objetivo *no* es «que salga el ejercicio». El objetivo es que aprendan a demostrar. Si llegan de  $P$  a  $Q$ , pero no están seguros si lo que hicieron está bien, *no terminaron el ejercicio*. Si la demostración ni siquiera los convence a ustedes, ¿cómo va a convencer a cualquier par? Vayan lento, vayan seguro, y van a aprender a hacerlo. No hay trucos, sólo sudor y tiza.

### Consejo

**El aprendizaje no sucede cuando terminan un ejercicio. El aprendizaje sucede cuando piensan, intentan, juegan, fallan, y reflexionan, durante horas.** Su objetivo debe ser este, y no el terminar el mayor número de ejercicios.

## 4 ¿Cómo demostramos?

Ahora miremos cómo se construye una demostración, proceduralmente. En el contexto de la carrera, los siguientes pasos van a ser necesarios cuando intenten escribir una demostración de algo.

### 4.1 Formalizar la consigna

En su vida profesional rara vez les van a dar un problema pre-formalizado, en términos de secuencias de enteros, o grafos. Esta parte es **crucial**: Si formalizan incorrectamente, todo lo que hagan después es totalmente irrelevante. Parte de esto es lectura y comprensión, y la otra parte es poder usar lenguaje formal. Consideremos la diferencia entre:

#### Ejercicio 4.1.1

Probar que en todo grupo de amigos, si cada par de amigos tiene sólo un amigo en común, entonces existe una persona que es amigo de todos.

La oración habla sobre grupos de amigos, no de algo que veamos directamente en la carrera. Para usar las herramientas que tenemos, lo traducimos a alguna estructura que nos sirva. En la carrera vemos varias, como ser números reales, listas, registros, números enteros, árboles, grafos, lenguajes

formales, matrices, redes, autómatas, interrupciones del procesador, etc. De todas esas, tenemos que fijarnos cuál es la que probablemente nos sirva para este problema. Como el enunciado habla sobre la relación «tener amigos», queremos algo que modele una relación de amistad. El enunciado no aclara que la amistad es simétrica y antireflexiva, así que es algo que deberíamos preguntar. Si podemos asumir eso, parecería que un grafo  $G = (V, E)$ , donde  $E$  es un subconjunto de pares sin orden de  $V$ , es un buen modelo. Podemos traducir el enunciado al siguiente enunciado sobre grafos.

#### Ejercicio 4.1.2 (Teorema de la amistad (Erdős et. al., 1966))

Sea  $G = (V, E)$  un grafo. Dado  $v \in V$ , definimos  $N(v) = \{w \mid \{v, w\} \in E\}$ . Probar que si para todo  $u, v \in V$  tenemos que  $|N(v) \cap N(w)| = 1$ , entonces existe un  $w \in V$  tal que  $|N(w)| = |V| - 1$ .

Ahora podemos usar las herramientas que nos da la carrera para atacar el problema. Notar cómo esta segunda versión nombra los objetos de los que habla ( $G, V, E, v, N$ , etc), explicita relaciones formalmente sobre los mismos ( $G = (V, E)$ ,  $v \in V$ ,  $N(v) = \dots$ ,  $|N(w)| = |V| - 1$ , etc), cuantifica las variables usadas («sea» / «para todo», «existe»), y usa conectores lógicos («si», «entonces»). Contrastar esto con algo como:

#### Ejercicio 4.1.3

Cada vez que dos personas siempre tienen un amigo en común, alguien es amigo de todos.

No sólo es difícil de leer, sino que distintas personas lo van a interpretar de distintas maneras, y eso aumenta el riesgo de no comprender la consigna correctamente. No se usan variables para referirse a la misma cosa, las relaciones entre los elementos no están explícitas, los cuantificadores son o inexistentes o imprecisos, y los conectores lógicos implícitos.

## 4.2 Comprender qué se nos pide

Una herramienta útil para entender qué hay que probar es pensarla como una conversación entre dos personas. A nosotros nos van a pedir que demostremos algo, y nosotros tenemos que convencer al interlocutor. Luego, una proposición como la siguiente:

#### Ejercicio 4.2.1 (Continuidad de $f(x) = e^x$ en $x_0$ )

Para todo  $\varepsilon > 0 \in \mathbb{R}$ , existe un  $\delta > 0 \in \mathbb{R}$ , tal que para todo  $x \in \mathbb{R}$ , si  $|x - x_0| < \delta$ , entonces  $|e^x - e^{x_0}| < \varepsilon$ .

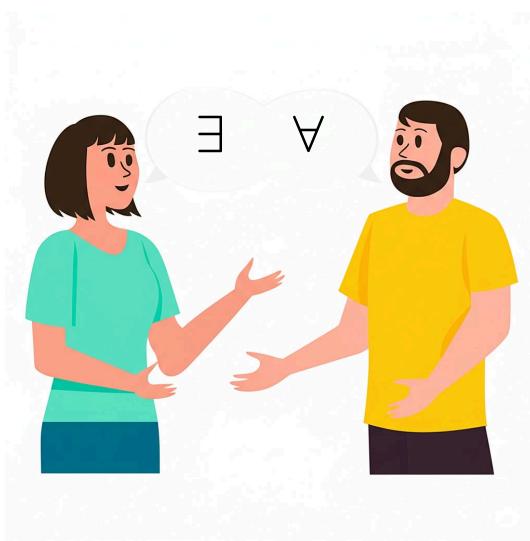
Probar un «para todo» es equivalente a la siguiente conversación, entre el que está probando algo (Alicia) y el que está pidiendo que le demuestren algo (Beto):

**Beto:** No te creo que es cierto, a ver, para  $\varepsilon = 0.2$ , quién es  $\delta$ ?

**Alicia:** Para  $\varepsilon = 0.2$ , te doy  $\delta = 0.4$ .

**Beto:** OK, ponele que  $\varepsilon = 0.2$ , y  $\delta = 0.4$ . Te doy un  $x$  tal que  $|x - x_0| < \delta$ , por ejemplo  $x = x_0 - 0.3$ . Por qué vale que  $|e^{x_0-0.3} - e^{x_0}| < 0.2$ ?

**Alicia:** [Demostración de que con ese  $\varepsilon$  y  $x$ , tenemos que  $|e^x - e^{x_0}| < \varepsilon$ .]



Nosotros vamos a tomar el rol de Alicia. Nos van a dar un  $\varepsilon$ , y tenemos que decir quién es  $\delta$ . Como nos están dando un  $\varepsilon$ , nuestro  $\delta$  puede (y en general va a) depender de  $\varepsilon$ . Por ejemplo, a veces vamos a concluir que  $\delta = \frac{\varepsilon}{8}$ . Este es el baile del «para todo / existe». El término «para todo» resulta en «se nos va a dar un». El término «existe» resulta en «tenemos que devolver».

Por el contrario, si tuvieramos probar «existe un  $x$  tal que para todo  $y \geq x$ ,  $x = y$ », entonces tenemos que dar un  $x$  explícito, y mostrar que sin importar cuál  $y$  elija Beto, podemos probar que  $x = y$ .

Luego de que devolvemos ese  $\delta$ , le sigue otro «para todo», «para todo  $x \in \mathbb{R}$ ». Entonces, nos van a dar otro  $x$ . A ese «para todo», le sigue un «si», «si  $|x - x_0| < \delta$ ». Un «si» nos deja asumir algo - para probar que  $X \Rightarrow Y$ , podemos *asumir*  $X$ , puesto que de otra manera no hay nada que probar («falso implica todo»). Entonces, esto se traduce en «nos van a dar un  $x$ , y podemos asumir que  $|x - x_0| < \delta$ ». El «entonces» de un «si» es lo que tenemos que probar. Luego, tenemos que probar que para ese  $x$  que nos dieron, vale  $|e^x - e^{x_0}| < \varepsilon$ .

Notemos cómo al igual que ocurre en la conversación, tuvimos que decir quién es  $\delta$  sin saber quién es  $x$ . Entonces, no puede ser que  $\delta$  dependa de  $x$ ! En la conversación, las únicas variables que vinieron antes de  $\delta$  fueron  $\varepsilon$  y  $x_0$ . Entonces, sólo de esas dos puede depender  $\delta$ .

### 4.3 Considerar ejemplos

En general, las cosas que probamos van a ser de la forma  $A \Rightarrow B$ , con  $A$  algo que podemos asumir, y  $B$  algo que queremos demostrar. Para demostrar esto, es frecuentemente útil considerar ejemplos de cosas que cumplen  $A$ , y ver «por qué» se tiene que cumplir  $B$  para ellas. Podemos empezar con ejemplos pequeños, si nuestra estructura tiene alguna noción de «tamaño» (la longitud de una lista, el valor absoluto de un número real, el número de vértices mas aristas de un grafo, el numero de líneas de un programa, el número de reglas de una gramática, etc).

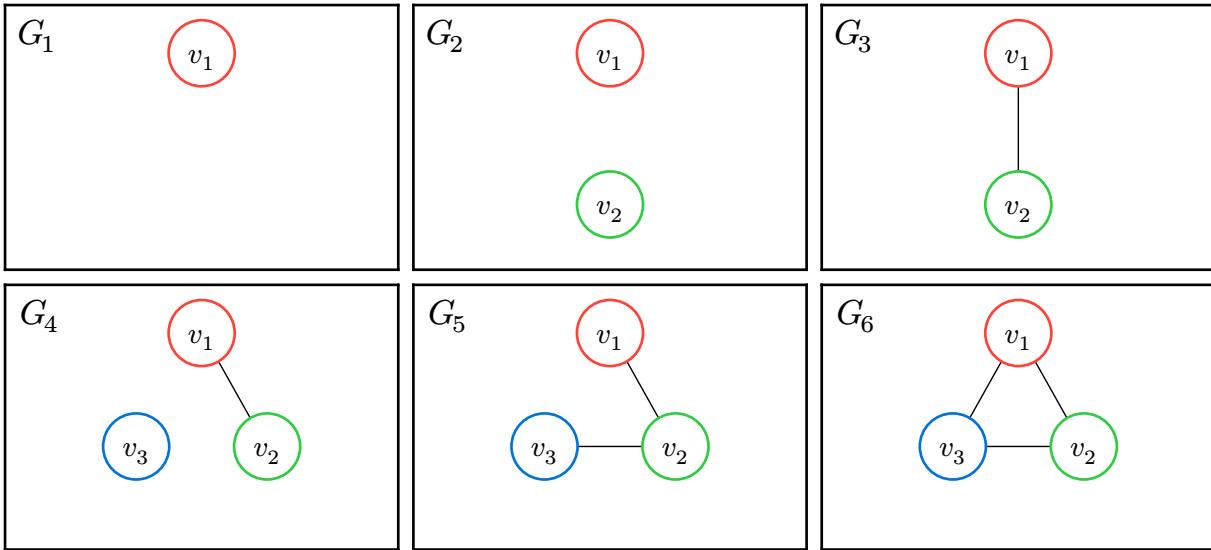
Por ejemplo, veamos el siguiente enunciado formal:

#### Ejercicio 4.3.1 (Fórmula de suma de grados)

Sea  $G = (V, E)$  un grafo,  $m = |E|$ , y  $d_G : V \rightarrow \mathbb{N}$  el grado de cada vértice. Entonces:

$$\sum_{v \in V} d_G(v) = 2m$$

Para ver «por qué» se tiene que cumplir esa ecuación, podemos comenzar viendo grafos pequeños.



Mirando qué pasa con  $G_1$  y  $G_2$ , vemos que en ambos casos, ambos lados de la ecuación son cero. El lado izquierdo es porque todos los vértices tienen grado cero, y el lado derecho porque no hay aristas. Mirando qué pasa en  $G_3$  y  $G_4$ , vemos que ambos lados de la ecuación son 2. El lado izquierdo vale 2 porque cuando sumamos el grado de  $v_1$  eso agrega 1, y cuando sumamos el grado de  $v_2$  eso agrega otro 1, quedando de resultado 2. El lado derecho vale porque hay una arista, luego  $m = 1$ , y luego  $2m = 2$ . En  $G_5$ , en la sumatoria sumamos 1 en  $v_1$  y  $v_3$  (porque hay una arista incidente a cada uno), y 2 en  $v_2$  (porque tiene dos aristas incidentes). Finalmente, en  $G_6$ , cada vértice es incidente a dos otros, y luego al sumar los grados estoy sumando 2 cada vez, obteniendo  $2 \times 3 = 6$ . Como hay dos tres aristas, el lado derecho es también  $2 \times 3 = 6$ .

Lo que está pasando, entonces, es que cada arista le suma 1 al grado de cada vértice que toca. Al sumar el grado de cada vértice, vamos a estar sumando 2 por cada arista, y luego obtendremos  $2m$  como total de la suma.

Quisieramos escribir algo así, entonces.

*Demostración.* Cada arista en  $E$  aumenta en 1 el grado de sus dos vértices incidentes. En total, las aristas suman  $2m$  a los grados de todos los vértices. Es decir,  $\sum_{v \in V} d_G(v) = 2m$ .  $\square$

Esa demostración es vaga, porque está haciendo referencia a un «procedimiento» (que implícitamente comienza todos los grados en cero). Si queremos hacerla formal, un «procedimiento» se traduce a una inducción. Si este procedimiento está modificando un grafo, en la inducción vamos a tener un grafo cada paso, y el grafo en un paso se construye a partir de grafos anteriores, en los que podemos usar la hipótesis inductiva.

*Demostración.* Sea  $G = (V, E)$  un grafo, y sean  $\{e_1, \dots, e_m\} = E$  sus aristas. Sea  $G_0 = (V, \emptyset)$  el grafo con los mismos vértices que  $G$ , pero sin aristas, y definimos para todo  $1 \leq i \leq m$ ,  $G_i = (V, \{e_1, \dots, e_i\})$ . Vemos que  $G_m = G$ .

Sea la proposición  $P(i)$ : Si  $i \leq m$ , entonces  $\sum_{v \in V} d_{G_i}(v) = 2i$ . Vamos a probar  $P$  por inducción.

Caso base:  $P(0)$ . Como  $i \leq m$ , pues  $m \geq 0$  es un natural, el antecedente vale. Todos los vértices de  $G_0$  tienen grado cero puesto que no hay aristas, y luego  $\sum_{v \in V} d_{G_0}(v) = 0 = 2 \times 0$ , que es lo que queríamos demostrar.

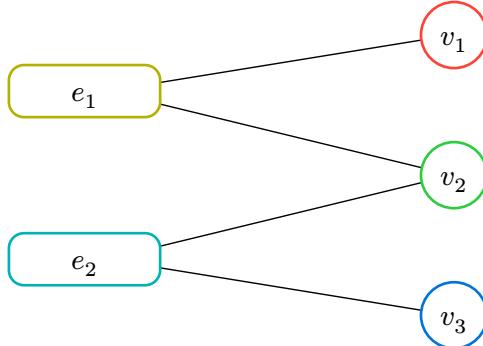
Paso inductivo. Asumimos  $P(i)$ , queremos probar  $P(i + 1)$ . Vemos que  $P(i + 1)$  es una implicación. Para probar una implicación, podemos asumir el antecedente. Luego, tenemos que  $i + 1 \leq m$ . Como vale eso, a fortiori vale  $i \leq m$ , que es el antecedente de  $P(i)$ , y luego podemos usar el consecuente de  $P(i)$ . Esto nos dice que  $\sum_{v \in V} d_{G_i}(v) = 2i$ . Sean  $\{v_j, v_k\} = e_{i+1}$  los dos vértices incidentes a  $e_{i+1}$ . Para todo otro  $v \in V$  que no es ni  $v_j$  ni  $v_k$ , tenemos  $d_{G_{i+1}}(v) = d_{G_i}(v)$ . Para esos dos, tenemos que  $d_{G_{i+1}}(v_j) = d_{G_i}(v_j) + 1$ , y  $d_{G_{i+1}}(v_k) = d_{G_i}(v_k) + 1$ . Luego,  $\sum_{v \in V} d_{G_{i+1}}(v) = 2i + 2 = 2(i + 1)$ , y queda probada la conclusión de  $P(i + 1)$ .

Como vale  $P(i)$  para todo  $i \in \mathbb{N}$ , en particular sabemos  $P(m)$ , y como  $m \leq m$ , sabemos que  $\sum_{v \in V} d_{G_m}(v) = 2m$ . Luego, como  $G_m = G$ , tenemos que  $\sum_{v \in V} d_G(v) = 2m$ , que es lo que queríamos demostrar.  $\square$

Otra forma de demostrar es darse cuenta que lo que queremos hacer es «contar lo mismo dos veces», hacer esto explícito creando un grafo con dos subconjuntos de vértices, y contando algo en cada «lado» del grafo:

*Demostración.* Consideremos el grafo bipartito  $G' = (V', E')$ , donde  $V' = V \sqcup E$ , la unión disjunta de  $V$  y  $E$ , y  $E' = \{\{e, v\} \mid e \in E, v \in e\}$ . Cada arista en  $E$  está unida en  $G'$  con sus dos vértices incidentes, en  $V'$ .

Para el grafo  $G_5$  que dibujamos arriba, esto se ve así:



Como  $|e| = 2$  para todo  $e \in E$ , tenemos que  $|E'| = 2|E| = 2m$ . También, como sabemos que  $d_G(v)$  es el número de veces que  $v$  aparece en una arista, tenemos que  $\sum_{v \in V} d_G(v) = \sum_{v \in V} |\{e \in E \mid v \in e\}| = |\{\{v, e\} \mid e \in E, v \in e\}| = |E'|$ . Luego, tenemos que  $\sum_{v \in V} d_G(v) = 2m$ .  $\square$

Vemos cómo el considerar ejemplos nos guió primero a una demostración informal, y luego la formalizamos. El hacer cálculos manuales también nos mostró otra estrategia posible para demostrar (contar lo mismo de dos formas distintas, una estrategia combinatórica clásica).

Veamos otro ejemplo de cómo jugar con ejemplos pequeños nos revela cómo lidiar con casos generales. Esta es la cadena de pensamientos que sigo al intentar resolver este ejercicio. No es una demostración formal, sólo una serie de ideas.

### Ejercicio 4.3.2

Demostrar que para todo  $n \in \mathbb{N}$ ,  $\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$ .



Veamos qué pasa con  $n = 1, 2, 3$ .

1. Hay un sólo término, y es  $1^2$ . Esto es  $\frac{1 \times 2 \times 3}{6}$ , pero no me es claro «por qué» todavía.
2.  $1^2 + 2^2 = 5 = \frac{2 \times 3 \times 5}{6}$ . Mmm todavía nada. No creo que el patrón de  $2 \times 3 \times \frac{X}{6}$  continúe.
3.  $1^2 + 2^2 + 3^2$ . ¿Puedo hacer algo con esto más que sumar cada cuadrado? Esto es  $1 + 2 \times 2 + 3 \times 3$ , o quizás,  $1 + 2 + 2 + 3 + 3 + 3 + 4 + 4 + 4 + 4$ . Me queda algo sospechoso,  $(1 + 2 + 3) + (2 + 3) + 3$ . Como sumas de no-cuadrados, hasta  $n$ , que se van haciendo más cortas. Esto es  $\frac{n(n-1)}{2} + \left(\frac{n(n-1)}{2} - 1\right) + \left(\frac{n(n-1)}{2} - 3\right)$ . El patrón queda algo como  $\sum_{i=0}^{n-1} N - \frac{i(i+1)}{2}$ , con  $N = \frac{n(n+1)}{2}$ .

A ver, pensémoslo con cuidado. Con  $n = 4$ , tenemos

$$\begin{aligned}\sum_{i=0}^n i^2 &= 0^2 + 1^2 + 2^2 + 3^2 + 4^2 \\&= 0 + 1 \times 1 + 2 \times 2 + 3 \times 3 + 4 \times 4 \\&= 1 + 2 + 2 + 3 + 3 + 3 + 4 + 4 + 4 + 4 \\&= (1 + 2 + 3 + 4) + (2 + 3 + 4) + (3 + 4) + (4) \\&= N + (N - 1) + (N - (1 + 2)) + (N - (1 + 2 + 3)) \\&= N + \left(N - 1 \times \frac{2}{2}\right) + \left(N - 2 \times \frac{3}{2}\right) + \left(N - 3 \times \frac{4}{2}\right)\end{aligned}$$

OK, esto es entonces  $\sum_{i=1}^n N - \frac{(i-1)i}{2} = Nn - \sum_{i=1}^n \frac{(i-1)i}{2}$ . Ese  $\frac{(i-1)i}{2}$  se va a convertir en  $\frac{i^2 - i}{2}$ , y si separo eso, ¡me va a quedar otra vez la suma de cuadrados!

Juguemos, entonces. Llamemos  $X = \sum_{i=0}^n i^2 = \sum_{i=1}^n i^2$ .

$$\begin{aligned}X &= Nn - \sum_{i=1}^n \frac{(i-1)i}{2} \\&= \frac{n^2(n+1)}{2} - \left(\frac{1}{2}\right) \left( \sum_{i=1}^n i^2 - i \right) \\&= \frac{n^2(n+1)}{2} - \left(\frac{1}{2}\right) \left( \sum_{i=1}^n i^2 \right) + \left(\frac{1}{2}\right) \sum_{i=1}^n i \\&= \frac{n^2(n+1)}{2} - \frac{X}{2} + \frac{n(n+1)}{4}\end{aligned}$$

Multiplicamos todo por 4, y pasamos todos los  $X$  a la izquierda.

$$\begin{aligned}6X &= 2n^2(n+1) + n(n+1) \\&= n(n+1)(2n+1)\end{aligned}$$

Con lo cual  $X = \frac{n(n+1)(2n+1)}{6}$ , que ¡es lo que quería!

Haciendo jugado con ejemplos pequeños, ahora sabemos «por qué» la proposición es cierta. Resta hacer un argumento convincente de que el patrón que encontramos se va a repetir para todo  $n \in \mathbb{N}$ . Usamos la herramienta formal de inducción, para hacer riguroso el argumento.

*Demostración.* Vamos a probar la proposición  $P(n) : \sum_{i=0}^n i^2 = \sum_{i=1}^n \frac{n(n+1)-(i-1)i}{2}$ , para todo  $n \in \mathbb{N}$ .

- $P(0)$ .  $\sum_{i=0}^1 i^2 = 0^2 = 0$ , mientras que  $\sum_{i=1}^0$  (lo que sea)  $= 0$ , porque hay cero términos en la suma. Luego vale  $P(0)$ .
- $P(n) \Rightarrow P(n + 1)$ . Entonces:

$$\begin{aligned}
\sum_{i=0}^{n+1} i^2 &= \left( \sum_{i=0}^n i^2 \right) + (n+1)^2 \\
&= \left( \sum_{i=1}^n \frac{n(n+1)-(i-1)i}{2} \right) + (n+1)(n+1) \\
&= \left( \sum_{i=1}^{n+1} \frac{n(n+1)-(i-1)i}{2} \right) + (n+1)(n+1), \text{ pues el último término es } 0 \\
&= \sum_{i=1}^{n+1} \left( \frac{n(n+1)-(i-1)i}{2} + n+1 \right), \text{ sumo } n+1 \text{ a cada uno de los } n+1 \text{ términos} \\
&= \sum_{i=1}^{n+1} \frac{2n+2+n(n+1)-(i-1)i}{2} \\
&= \sum_{i=1}^{n+1} \frac{(n+1)(n+2)-(i-1)i}{2}
\end{aligned}$$

que es lo que queríamos demostrar.

Habiendo probado  $P(n)$  para todo  $n \in \mathbb{N}$ , veamos cómo probar que  $\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$  para todo  $n \in \mathbb{N}$ . Sea  $n \in \mathbb{N}$ , y llamemos  $X = \sum_{i=0}^n i^2$ .

$$\begin{aligned}
X &= \sum_{i=0}^n i^2 = \sum_{i=1}^n \frac{n(n+1)-(i-1)i}{2} \\
&= n^2 \frac{n+1}{2} - \sum_{i=1}^n \frac{i(i-1)}{2} \\
&= n^2 \frac{n+1}{2} - \left( \frac{1}{2} \right) \sum_{i=1}^n i^2 + \left( \frac{1}{2} \right) \sum_{i=1}^n i \\
&= n^2 \frac{n+1}{2} - \left( \frac{1}{2} \right) X + \frac{n(n+1)}{4}
\end{aligned}$$

Multiplicando por 4 y pasando las  $X$  a la izquierda...

$$\begin{aligned}
6X &= 2n^2(n+1) + n(n+1) \\
&= n(n+1)(2n+1)
\end{aligned}$$

Y por lo tanto  $\sum_{i=0}^n i^2 = X = \frac{n(n+1)(2n+1)}{6}$ , que es lo que queríamos demostrar.  $\square$

Nuestro cerebro es muy eficiente para ser perezoso, y frecuentemente al hacer cuentas repetitivas, vamos a encontrar patrones que nos ahorren esfuerzo, y al mismo tiempo revelen propiedades sobre los objetos que estamos mencionando. Jamás se me hubiera ocurrido distribuir un  $n + 1$  a cada uno de los  $n + 1$  términos de la sumatoria, si no fuera porque fue exactamente lo que hice para  $n = 3$  y  $n = 4$ , notando ese patrón en un caso chico.

## 4.4 Estrategias de demostración

Una vez que entendemos qué es lo que se no pide probar y tenemos una idea intuitiva de por qué funciona, podemos planear estructuras que va a tener nuestra demostración. En general vamos a usar varias de ellas en la misma demostración.

### 4.4.1 Inducción

Si los objetos con los que estamos trabajando tienen una estructura recursiva, como los números naturales, los árboles, los grafos, o las cadenas de texto, entonces podemos considerar inducción como técnica de demostración.

#### ⚠️ Advertencia

Sugiero ser formal cuando hacen inducción. Es muy común que se confundan y prueben algo como «Si  $G = (V, E)$  es tal que  $|V| = n$ , entonces me armo  $G'$  de  $n + 1$  agregándole un vértice conectado con una arista a  $G$ . Como asumo que vale mi propiedad para  $G$ , y pruebo que vale mi propiedad para  $G'$ , entonces queda probada la propiedad para todos los grafos». Eso es incorrecto. Vamos a ver luego más ejemplos de errores. En general, se usa la formalidad para no cometer errores, cuando están aprendiendo a demostrar.

El siguiente es un ejemplo de una demostración incorrecta de una proposición falsa, por ser informal y no definir explícitamente una proposición sobre los números naturales, con cuantificadores.

#### Proposición 4.4.1

Sea  $G$  un grafo. Si todos los vértices de  $G$  tienen grado mayor o igual a 1, entonces  $G$  es conexo.



#### Demostración.

1. Caso base. Si  $G$  tiene un sólo vértice, entonces no vale el antecedente, y luego la implicación es cierta (recordar que «falso implica todo»).
2. Paso inductivo. Sea  $G = (V, E)$  un grafo con todos los grados mayores o iguales a 1. Construimos  $G'$  tomando  $G$  y agregándole un vértice nuevo  $v$ , conectado con algún número positivo de vértices en  $G$ . Sea  $V' = V + \{v\}$  el conjunto de vértices de  $G'$ . Sea  $w$  alguno de los vértices en  $V$  a los que conectamos a  $v$ . Entonces  $\{w, v\}$  es una arista en  $G'$ .

Sea  $z \in V$ . Por hipótesis inductiva,  $G$  es conexo, entonces existe un camino  $z \rightsquigarrow w$  en  $G$ , y luego también en  $G'$ . Como  $\{w, v\}$  es una arista en  $G'$ , tenemos un camino  $z \rightsquigarrow v$  en  $G'$ .

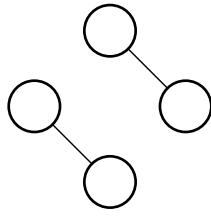
Veamos que  $G'$  es conexo. Sean  $a, b \in V'$  cualquier par de vértices distintos en  $G'$ . Si  $a \neq v$  y  $b \neq v$ , entonces ya había un camino entre  $a$  y  $b$  en  $G$  por ser  $a, b \in V(G)$ , y  $G$  conexo. Como  $G$  es subgrafo de  $G'$ , este camino sigue existiendo en  $G'$ . Si alguno de los dos es  $v$ , vimos arriba que hay un camino desde cualquier vértice de  $V$  hasta  $v$  en  $G'$ .

Luego hay un camino entre todo par de vértices en  $G'$ , y luego  $G'$  es conexo, que es lo que queríamos demostrar.



Esto es falso, y tenemos este contraejemplo:

$$G = K_2 \cup \\ K_2:$$



Todos los vértices de este grafo tienen grado mayor o igual a 1, y sin embargo no es conexo.



Esa demostración está mal, fundamentalmente, porque está siendo informal (y luego errónea) en cómo funciona la inducción. La inducción es sobre naturales. No es sobre grafos.

Lo que sí está probando es que si empezamos con un grafo conexo, y varias veces agregamos vértices conectados con al menos un vértice preexistente, seguimos teniendo un grafo conexo. Esto es cierto, pero no es lo que se nos pidió probar.

La diferencia está en que no todo grafo cuyos vértices tienen grados mayores o iguales a 1, viene de hacer ese procedimiento iterativo. Luego, la recomendación es que sean formales, que escriban cuál exactamente es la propiedad sobre números naturales que quieren demostrar.

### Advertencia

Si nuestra demostración de  $P(n)$  requiere que  $P(n-1), P(n-2), \dots, P(n-k)$  sea cierto para algún  $k \geq 1$  fijo, entonces vamos a necesitar  $k$  casos base. Esto es porque nuestra demostración no va a tener sentido cuando  $n < k$ , porque estaríamos diciendo que  $P(n-k)$  vale, con  $n-k < 0$ , que no tiene sentido pues  $P$  es una proposición sobre los naturales. Veamos un ejemplo.

Veamos primero una demostración correcta que toma esto en cuenta, y luego dos que son incorrectas por no hacerlo.

#### Ejercicio 4.4.2 (Fórmula cerrada para la sucesión de Fibonacci)

Sea  $a_0 = 0, a_1 = 1$ , y para todo  $n > 1$ , definamos  $a_n = a_{n-1} + a_{n-2}$ . Entonces para todo  $n \in \mathbb{N}$ , tenemos  $a_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}$ , con  $\varphi = \frac{1+\sqrt{5}}{2}$ , y  $\psi = \frac{1-\sqrt{5}}{2}$ .

*Demostración.* La propiedad que vamos a probar por inducción es  $P(n) : a_n = \frac{\varphi^n - \psi^n}{\sqrt{5}}$ . Como  $a_n$  está definida en términos de  $a_{n-1}$  y  $a_{n-2}$ , vamos a querer decir algo sobre  $a_{n-1}$  y  $a_{n-2}$ , lo cual significa que vamos a usar  $P(n-1)$  y  $P(n-2)$ . Luego, tenemos dos casos base.

- Caso base  $n = 0$ . Si  $n = 0$ , entonces  $a_n = a_0 = 0$ , por definición.  $\varphi^n = \varphi^0 = 1$ , y también  $\psi^n = \psi^0 = 1$ . Luego  $\varphi^n - \psi^n = 0$ , y por lo tanto  $a_n = a_0 = 0 = \frac{\varphi^0 - \psi^0}{\sqrt{5}} = \frac{\varphi^n - \psi^n}{\sqrt{5}}$ , que es lo que queríamos demostrar.
- Caso base  $n = 1$ . Si  $n = 1$ , entonces  $a_n = a_1 = 1$ , por definición.  $\varphi^1 = \frac{1+\sqrt{5}}{2}$ , y  $\psi^1 = \frac{1-\sqrt{5}}{2}$ , luego  $\varphi^n - \psi^n = \varphi^1 - \psi^1 = \frac{1+\sqrt{5}}{2} - \frac{1-\sqrt{5}}{2} = 2\frac{\sqrt{5}}{2} = \sqrt{5}$ , y luego  $\frac{\varphi^n - \psi^n}{\sqrt{5}} = 1 = a_0 = a_n$ , que es lo que queríamos demostrar.
- Paso inductivo. Podemos asumir  $P(n-1)$  y  $P(n-2)$ , y queremos probar  $P(n)$ . Vemos que  $\varphi^{-1} = \frac{\sqrt{5}-1}{2}$ , y de ahí vemos que  $1 + \varphi^{-1} = \varphi$ , y que  $1 - \varphi = -\frac{1}{\varphi}$ . Realmente ambos

hechos son consecuencia de que  $\varphi$  es una de las raíces de  $\varphi + 1 = \varphi^2$ . Notemos que  $\psi = 1 - \varphi = -\frac{1}{\varphi}$ . Entonces:

$$\begin{aligned}
a_n &= a_{n-1} + a_{n-2} \\
&= \frac{\varphi^{n-1} - \psi^{n-1}}{\sqrt{5}} + \frac{\varphi^{n-2} - \psi^{n-2}}{\sqrt{5}} \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} - \left(-\frac{1}{\varphi}\right)^{n-1} \right] + \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-2} - \left(-\frac{1}{\varphi}\right)^{n-2} \right] \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} - \left(-\frac{1}{\varphi}\right)^{n-1} \right] + \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} \frac{1}{\varphi} - \left(-\frac{1}{\varphi}\right)^{n-1} (-\varphi) \right] \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} - \left(-\frac{1}{\varphi}\right)^{n-1} + \varphi^{n-1} \frac{1}{\varphi} - \left(-\frac{1}{\varphi}\right)^{n-1} (-\varphi) \right] \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} \left(1 + \frac{1}{\varphi}\right) - \left(-\frac{1}{\varphi}\right)^{n-1} (1 - \varphi) \right] \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^{n-1} \varphi - \left(-\frac{1}{\varphi}\right)^{n-1} \left(-\frac{1}{\varphi}\right) \right] \\
&= \left(\frac{1}{\sqrt{5}}\right) \left[ \varphi^n - \left(-\frac{1}{\varphi}\right)^n \right] \\
&= \frac{\varphi^n - \psi^n}{\sqrt{5}}
\end{aligned}$$

□

Ahora veamos qué pasa si no somos cuidadosos al usar la hipótesis inductiva.

#### Proposición 4.4.3

Para todo  $n \in \mathbb{N}$ ,  $2^n = 1$ .



*Demostración.* Vamos a probar que vale  $P(n)$  para todo  $n \in \mathbb{N}$ , con  $P(n) : 2^n = 1$ .

1. Caso base,  $P(0)$ . Es cierto que  $2^0 = 1$ , y luego  $P(0)$  es cierta.
2. Paso inductivo. Asumimos que vale  $P(j)$  para todo  $j < n + 1$ , y probemos  $P(n + 1)$ .

Manipulamos algebraicamente:

$$\begin{aligned}
2^{n+1} &= 2^{2n-(n-1)} \\
&= \frac{2^{2n}}{2^{n-1}} \\
&= \frac{2^n 2^n}{2^{n-1}} \\
&= 1 \times \frac{1}{1} \text{ por hipótesis inductiva, tres veces} \\
&= 1
\end{aligned}$$

que es lo que queríamos demostrar. □

El error vino de usar  $P(n - 1)$ . Esto no tiene sentido cuando estamos probando  $P(1)$ , porque entonces  $1 = n + 1$  y entonces  $n = 0$ , y no tiene sentido decir  $n - 1$ . No fuimos cuidadosos al asumir que  $n > 1$ , lo cual nos hubiera marcado que debemos probar  $P(1)$  por separado, no sólo  $P(0)$ .

¿Pueden detectar dónde está el error en la siguiente demostración?

#### Proposición 4.4.4

Probar que para todo  $n \in \mathbb{N}$ ,  $5n = 0$ .



*Demostración.* Sea la proposición  $P(n) : 5n = 0$ . Vamos a probar  $P(n) \forall n \in \mathbb{N}$  por inducción.

- $P(0)$ . Queremos probar  $P(0)$ , que significa  $5 \times 0 = 0$ , y esto es cierto. Luego vale  $P(0)$ .
- $n > 0 \Rightarrow P(n)$ . Sean  $i, j \in \mathbb{N}$ , con  $i < n, j < n$ , tales que  $i + j = n$ . Entonces por hipótesis inductiva vale  $P(i)$  y  $P(j)$ , entonces  $5i = 0$  y  $5j = 0$ . Entonces,  $5n = 5(i + j) = 5i + 5j = 0 + 0 = 0$ , lo cual prueba  $P(n)$ .



El error está en asumir que existen naturales  $i < n, j < n$ , con  $i + j = n$ . Esto sólo es cierto si  $n \geq 2$ , y entonces nuestra demostración falla para  $n = 1$ , y se cae la inducción. No fuimos cuidadosos, y nos faltó el caso base  $n = 1$ .

#### 4.4.2 Correctitud de ciclos en algoritmos

Frecuentemente vamos a probar propiedades sobre algoritmos que usan ciclos. La herramienta principal que tenemos para esto es el teorema del invariante. Esto no es nada más que inducción en el número de iteraciones, con un formalismo al rededor para evitar que cometan errores (como, por ejemplo, olvidarse de probar que el ciclo efectivamente termina).

Para usar el teorema del invariante, necesitamos definir cinco cosas:

1. Una precondición  $P$ . Esto es algo que asumimos que vale antes de comenzar el ciclo.
2. Una postcondición  $Q$ . Esto es algo que queremos probar que vale al terminar el ciclo.
3. Un invariante  $I$ . Esto es algo que es válido antes de cada iteración, y después de cada iteración (pero no necesariamente durante una iteración).
4. Una guarda  $B$ , que nos dice si ejecutaremos la próxima iteración del ciclo, o no.
5. Una función variante  $v$ , que nos obliga a terminar el ciclo cuando llega a cero.

Y tenemos que demostrar las siguientes seis proposiciones:

1. La precondición vale antes de comenzar el ciclo.
2. La precondición implica el invariante.
3. La función variante decrece con cada iteración.
4. Si la función variante es cero, la guarda es falsa.
5. El invariante y la negación de la guarda implican la postcondición.
6. La guarda y el invariante al comenzar la iteración implican el invariante al terminar la iteración.



En Algoritmos 1, aprenden a especificar estas proposiciones, y a demostrar estas proposiciones. En la sección de ejemplos muestro varios, pero para mostrarles uno acá, vamos a probar la correctitud del algoritmo de exponenciación en tiempo logarítmico.

```

1: procedure EXP( $a \in \mathbb{N}, n \in \mathbb{N}$ )
2:   if  $n = 0$  then
3:     return 1
4:   end
5:    $k \leftarrow n$ 
6:    $y \leftarrow 1$ 
7:   while  $k > 1$  do
8:     if  $k \bmod 2 = 1$  then
9:        $y \leftarrow x \times y$ 
10:       $k \leftarrow k - 1$ 
11:    end
12:     $x \leftarrow x^2$ 
13:     $k \leftarrow k/2$ 
14:  end
15:  return  $x \times y$ 
16: end

```

*Demostración.* Llamemos  $x_0$  al valor de  $x$  al comenzar el algoritmo. Queremos probar que el programa devuelve  $x_0^n$ .

1. La precondición del algoritmo es  $n > 0, k = n, y = 1, x = x_0$ .
2. La postcondición del ciclo es  $x \times y = x_0^n$ .
3. El invariante es  $1 \leq k \leq n \wedge x^k y = x_0^n$ .
4. La guarda es  $k > 1$ .
5. La función variante es  $k - 1$ .

Probemos el teorema del invariante para este ciclo.

1. Estamos asignando  $k \leftarrow n, y \leftarrow 1$ , luego valen  $k = n, y = 1$ . Lo primero que hace nuestro programa es salir si  $n = 0$ , en cuyo caso devuelve la respuesta correcta y no ejecuta nada más. Luego, al comenzar el ciclo, sabemos que  $n \in \mathbb{N}, n \neq 0$ , y luego  $n > 0$ . Finalmente, antes de comenzar el ciclo no modificamos  $x$ , con lo cual sigue valiendo  $x = x_0$ .
2. Como vale la precondición, entonces  $k = n$ , y por ende  $k \leq n$ . Asimismo, la precondición nos dice que  $n > 0$ , y por ende  $k > 0$ , o lo que es equivalente,  $k \geq 1$ . Juntando las dos oraciones anteriores, tenemos que  $1 \leq k \leq n$ . Finalmente, como la precondición nos dice que  $y = 1$  y  $x_0 = x$ , tenemos  $x^k y = x^n y = x^n = x_0^n$ , que prueba el invariante.

3. En cada iteración, estamos dividiendo a  $k$  por 2, y quizás restándole 1, con lo cual siempre va a decrecer, porque  $k > 1$  (sólo podría no-decrecer si  $k = 0$ , porque entonces  $k/2 = k$ ). Luego, como  $k$  decrece, también  $k - 1$ , la función variante, decrece.
4. Si la función variante es cero, entonces  $k - 1 = 0$ , y luego  $k = 1$ . Como la guarda es  $k > 1$ , el que la función variante sea cero obliga a que la guarda no se cumpla.
5. La negación de la guarda nos dice que  $k \leq 1$ . El invariante nos dice que  $1 \leq k$ . Luego, sabemos que  $k = 1$ . El invariante también nos dice que  $x^k \times y = x_0^n$ , entonces sabemos que  $x^1 \times y = x \times y = x_0^n$ , que es la postcondición.
6. Sabemos que  $k > 1$  porque vale la guarda. Partimos en casos, dependiendo de si  $k \bmod 2 = 0$  o  $k \bmod 2 = 1$ .
  - Si  $k = 2k'$ , entonces entramos al condicional. Lo que hacemos es cuadrar  $x$  obteniendo  $x'$ , y dividir  $k$  por dos, obteniendo  $k'$ . Como vale la guarda, sabemos que  $k > 1$ , y como  $k \in \mathbb{N}$ , tenemos  $k \geq 2$ . Como  $k = 2k'$ , vemos que  $k' \geq 1$ , que es la primer parte del invariante. Empezamos la iteración con  $x^k y = x_0^n$ , es decir  $x^{2k'} y = x_0^n$ . Usando el álgebra de exponentiación, vemos que  $x^{2k'} = (x^2)^{k'}$ , y  $k' = k/2$ . Luego, vemos que vale  $(x^2)^{k'} y = x_0^n$ , o  $x'^{k'} y = x_0^n$ , que es el invariante al terminar la iteración.
  - Si  $k = 2k' + 1$ , entonces entramos al condicional. El efecto de la iteración en este caso es transformar  $k$  en  $k'$ ,  $x$  en  $x' = x^2$ , y  $y' = x \times y$ . Como  $k > 1$ , y  $k \in \mathbb{N}$ , sabemos que  $k \geq 2$ , pero como  $k$  es impar,  $k \geq 3$ . Como  $k = 2k' + 1$ , entonces, tenemos  $2k' \geq 2$ , y luego  $k' \geq 1$ , que es la primer parte del invariante. Como vale el invariante al comenzar la iteración, sabemos que  $x^{2k'+1} y = x_0^n$ . Nuevamente usando álgebra de exponentiación, obtenemos  $x_0^n = x^{2k'+1} y = x^{2k'} xy = (x^2)^{k'} xy = (x^2)^{k'} y' = x'^{k'} y'$ , y por tanto vale el invariante al terminar la iteración.

Concluimos la postcondición, que junto con la salida temprana en el caso de  $n = 0$  nos deja concluir que el valor de retorno de nuestro programa es efectivamente  $x_0^n$ , con lo cual es un algoritmo de exponentiación correcto.  $\square$

### Nota

Noten cómo los algoritmos con estado son más engorrosos de demostrar correctos, pues necesitamos más formalidad sobre las transiciones de estado para evitar cometer errores. Un argumento poco riguroso sería decir «En cada iteración,  $x^k y = x_0^n$ », pero esto abre la puerta a errores (¿en qué momento de la iteración? ¿por qué es cierto eso?). El teorema del invariante es una herramienta formal para obtener rigor.

#### 4.4.3 Correctitud de algoritmos recursivos

Si nuestro algoritmo es recursivo, en general vamos a usar inducción para probar su correctitud. Veamos una versión recursiva del algoritmo `Exp`.

```

1: procedure Exp( $a \in \mathbb{N}$ ,  $n \in \mathbb{N}$ )
2:   if  $n = 0$  then
3:     return 1
4:   end
5:    $b \leftarrow \text{Exp}(a, \lfloor \frac{n}{2} \rfloor)$ 
6:    $c \leftarrow b^2$ 
7:   if  $n \bmod 2 = 1$  then
8:      $c \leftarrow c \times a$ 
9:   end

```

```

10:   return c
11: end

```

Para probar su correctitud, vamos a definir una noción de «tamaño» de entrada, y probar la correctitud de nuestro algoritmo para todas las entradas de tamaño menor o igual a  $n$ , para todo  $n \in \mathbb{N}$ .

#### Proposición 4.4.5

El algoritmo recursivo devuelve  $a^n$ .



*Demostración.* Definamos una propiedad  $P(n)$ : Para todo  $a \in \mathbb{N}$ ,  $\text{Exp}(a, n) = a^n$ .

- $P(0)$ . Queremos ver que  $\text{Exp}(a, 0) = a^0 = 1$  para todo  $a \in \mathbb{N}$ . Si  $n = 0$ , entonces entramos en el **if** de la línea 2, y devolvemos 1, que es la respuesta correcta.
- $\forall n \in \mathbb{N}. (n > 0 \wedge (\forall k \in \mathbb{N}. k < n \Rightarrow P(k))) \Rightarrow P(n)$ . Vamos a usar inducción global. Si  $n > 0$ , entonces no entramos en el **if** de la línea 2. Llamamos a  $\text{Exp}(a, \lfloor \frac{n}{2} \rfloor)$ . Como  $\lfloor \frac{n}{2} \rfloor < n$ , podemos usar la hipótesis inductiva para ver que  $b = a^{\lfloor \frac{n}{2} \rfloor}$ . Ahora partimos en casos:
  - Si  $n \equiv 0 \pmod{2}$ , entonces  $\lfloor \frac{n}{2} \rfloor = \frac{n}{2}$ . Luego,  $a^n = (a^{\frac{n}{2}})^2$ . Como asignamos  $b \leftarrow a^{\frac{n}{2}}$ , y luego  $c \leftarrow b^2$ , vemos que  $c = a^n$ . Como  $n \equiv 0 \pmod{2}$ , no entramos en el **if** de la línea 7. Luego, cuando devolvemos  $c$  en la línea 10, estamos devolviendo  $a^n$ , que es la respuesta correcta.
  - Si  $n \equiv 1 \pmod{2}$ , entonces  $\lfloor \frac{n}{2} \rfloor = \frac{n-1}{2}$ . Luego,  $a^n = a^{2\frac{n-1}{2}+1} = a^{2\frac{n-1}{2}} \times a = a^{(\lfloor \frac{n}{2} \rfloor)^2} \times a$ . Como asignamos  $b \leftarrow a^{\lfloor \frac{n}{2} \rfloor}$ , y luego  $c \leftarrow b^2$ , vemos que  $c = a^{2\frac{n-1}{2}}$ . Como  $n \equiv 1 \pmod{2}$ , entonces entramos al **if** de la línea 7, y multiplicamos a  $c$  por  $a$ , obteniendo  $c = a^{2\frac{n-1}{2}} \times a = a^{2\frac{n-1}{2}+1} = a^{n-1+1} = a^n$ . Luego, cuando devolvemos  $c$  en la línea 10, estamos devolviendo  $a^n$ , que es la respuesta correcta.

□

Vemos como es más fácil demostrar esto que un algoritmo iterativo, que cambia estados. Esto es cierto en general, y es parte del motivo por el cual la gente usa algoritmos y lenguajes de programación funcionales.

#### 4.4.4 Definiciones equivalentes

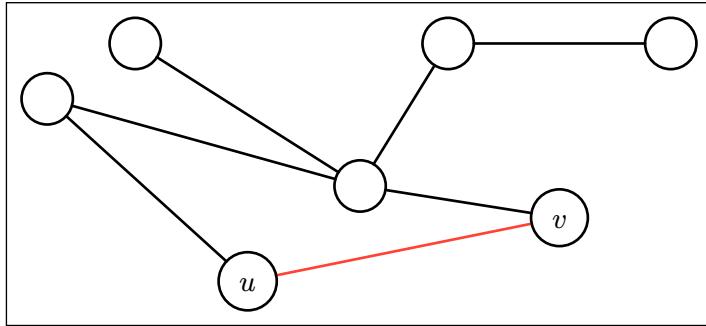
Si tenemos definiciones equivalentes para nuestro objeto, podemos hacer uso de cualquiera de ellas. Por ejemplo, si tenemos que probar que agregarle una arista a cualquier árbol crea un ciclo en el grafo resultante, podemos usar cualquiera de las definiciones equivalentes de árbol:

A un grafo  $G = (V, E)$  con  $n$  vértices y  $m$  aristas se le dice árbol cuando cumple cualquiera de las siguientes condiciones equivalentes:

- $G$  es conexo y acíclico.
- $G$  es conexo y  $m \leq n - 1$ .
- $G$  es acíclico y  $m \geq n - 1$ .
- $G$  es conexo y sacarle cualquier arista lo desconecta.
- $G$  es acíclico y agregarle cualquier arista produce un ciclo.
- Cualquier par de vértices en  $G$  está conectado por un único camino en  $G$ .

En este caso, probablemente nos sea útil que un árbol es un grafo  $G = (V, E)$  tal que para todo  $v, w$  en  $V$ , existe en  $G$  un único camino entre  $v$  y  $w$ . Al agregar una arista  $e = (v, w)$ , y sabiendo que ya

había un camino entre  $v$  y  $w$ , vamos a poder crear un ciclo. Para esto es importante saber *todas* las definiciones equivalentes de los objetos que usamos.



Si tenemos que probar condiciones equivalentes, frecuentemente nos va a ser útil ordenarlas de manera que cada una implique la próxima, y la última implique la primera, estableciendo así la equivalencia entre todas, usando sólo implicaciones.

#### Ejercicio 4.4.6

Sea  $X$  un conjunto, y  $\sim$  una relación de equivalencia en  $X$ . Sean  $a, b \in X$ . Entonces son equivalentes:

1.  $a \sim b$
2.  $[a] \cap [b] \neq \emptyset$
3.  $[a] = [b]$

*Demostración.*

- 1  $\Rightarrow$  2: Como  $a \sim b$ , entonces  $b$  es un elemento de  $[a]$ . También  $b$  es un elemento de  $[b]$ . Entonces  $b \in [a] \cap [b]$ , y luego  $[a] \cap [b] \neq \emptyset$ .
- 2  $\Rightarrow$  3: Sea  $y \in [a] \cap [b]$ . Luego  $y \sim a$  como también  $y \sim b$ . Entonces para todo  $x \in [a]$ , tenemos por transitividad que  $x \sim y \sim b$ , y luego  $x \in [b]$ , y luego  $[a] \subseteq [b]$ . De la misma manera, si  $x \in [b]$ , tenemos que  $x \sim b \sim y \sim a$ , y luego  $x \in [a]$ , y  $[b] \subseteq [a]$ . Entonces  $[a] = [b]$ .
- 3  $\Rightarrow$  1: Como  $[a] = [b]$  y  $a \in [a]$ , entonces  $a \in [b]$ , y luego  $a \sim b$ .

□

#### 4.4.5 Contrarecíproco y contradicción

Si tenemos que probar una implicación, es decir una proposición de la forma  $P \Rightarrow Q$ , podemos probar algo equivalente, que es el contrarecíproco de esa implicación. El contrarecíproco de  $P \Rightarrow Q$  es  $(\neg Q) \Rightarrow (\neg P)$ . Hacemos esto cuando nos es más cómodo tener de antecedente  $\neg Q$ , por ejemplo porque ya probamos  $\neg Q$  y queremos usar esto en un *modus ponens* para probar  $\neg P$ .

Esto también se puede usar para probar proposiciones en general, aún cuando no sean inmediatamente implicaciones. Si tenemos la proposición  $P$ , esta es equivalente a la proposición  $\top \Rightarrow P$ . Luego, es equivalente a su contrarecíproco, que es  $(\neg P) \Rightarrow \perp$ . Esto se llama probar  $P$  por contradicción, o a veces, «por absurdo».

Aún siendo equivalentes, a veces una va a ser más fácil de probar que la otra. Por ejemplo, si  $P$  es de la forma  $P : \neg(\exists y.Q(y))$ , parece difícil probarla, porque es difícil probar la no-existencia de algo,

uno parece carecer de herramientas. Al negarla obtenemos  $\exists y.Q(y)$ . Eso nos da otro objeto, otro sustantivo del cual hablar, y con el cual razonar. Cambiamos nuestra misión a ver si podemos obtener algo internamente contradictorio sobre  $y$ , y queremos llegar a  $\perp$ . Lo mismo sucede si tenemos una proposición de la forma  $P : \forall x.Q(x)$ . Probar esto por contradicción es probar  $(\neg(\forall x.Q(x))) \Rightarrow \perp$ , que es lo mismo que decir  $(\exists x.\neg Q(x)) \Rightarrow \perp$ . Eso otra vez nos da un objeto,  $x$ , del cual hablar.

Por ejemplo, el siguiente teorema era probablemente sabido por Aristóteles, y aparece en «Elementos» de Euclides:

### Teorema 7

$$\sqrt{2} \notin \mathbb{Q}$$



*Demostración.* Asumamos  $\sqrt{2} \in \mathbb{Q}$ . Luego, existen  $a, b \in \mathbb{Z}$  tal que  $\sqrt{2} = \frac{a}{b}$ , con  $b \neq 0$ . Podemos elegir a  $a, b$  coprimos, dado que si no lo fueran, tomamos  $\frac{a}{b} = \frac{\frac{a}{d}}{\frac{b}{d}}$ , con  $d = \gcd(a, b)$ , y ahora teniendo el numerador y denominador coprimos.

Como  $\sqrt{2} = \frac{a}{b}$ , tenemos que  $\frac{a^2}{b^2} = 2$ , y luego  $a^2 = 2b^2$ . Luego  $a$  es par, pues  $a^2$  lo es, y el cuadrado de un número impar sería impar.

Entonces existe  $k \in \mathbb{Z}$ , tal que  $a = 2k$ . Usando esto obtenemos que  $2b^2 = (2k)^2 = 4k^2$ , y luego  $b^2 = 2k^2$ . Luego,  $b$  también es par, porque su cuadrado es par.

Luego ambos  $a$  y  $b$  son pares. Esto no puede suceder, porque dijimos que eran coprimos. Luego, lo que asumimos es falso, y no es cierto que  $\sqrt{2} \in \mathbb{Q}$ . Esto prueba que  $\sqrt{2} \notin \mathbb{Q}$ , que es lo que queríamos demostrar.  $\square$

El motivo por el que la demostración es por contradicción es porque el que  $x$  no esté en  $\mathbb{Q}$  nos dice poco, sólo que  $x \in \mathbb{R} \setminus \mathbb{Q}$ , pero no es cómodo hablar sobre los irracionales. El demostrar por contradicción nos deja «imaginar» que  $x \in \mathbb{Q}$ , y el objetivo es llegar a  $\perp$ , algo falso. Usando esta hipótesis concluimos cosas útiles sobre  $x$ , como que  $x = \frac{n}{m}$  con  $n, m \in \mathbb{Z}, m \neq 0$ .

### ⚠️ Advertencia

Frecuentemente vemos alumnos empezar usando esta estrategia automáticamente, sin pensar en por qué se lo está haciendo. Imaginamos que es porque les da algo que escribir, y «se parece a progreso». Pero realmente no sugerimos hacer esto sin tener una razón específica, muchas veces se confunden con la cantidad (y paridad) de negaciones, ya que cada vez que hacen esto agregan una negación más. Eventualmente cometan algún error, llegan a un absurdo, y dicen «¡Listo, terminé!», pero el absurdo vino de hacer otra cosa mal en el medio.

Si van a usar esta estrategia, recuerden las reglas de negación.

$$\begin{aligned}
\neg(\forall x \in X.P(x)) &\Leftrightarrow \exists x \in X.\neg(P(x)) \\
\neg(\exists x \in X.P(x)) &\Leftrightarrow \forall x \in X.\neg(P(x)) \\
\neg(P \Rightarrow Q) &\Leftrightarrow (\neg Q) \Rightarrow (\neg P) \\
\neg(P \vee Q) &\Leftrightarrow (\neg P) \wedge (\neg Q) \\
\neg(P \wedge Q) &\Leftrightarrow (\neg P) \vee (\neg Q) \\
\neg\perp &\Leftrightarrow \top \\
\neg\top &\Leftrightarrow \perp
\end{aligned}$$

Vemos frecuentemente el error  $\neg(P \Rightarrow Q) \Leftrightarrow \neg P \Rightarrow \neg Q$ , y  $\neg(P \Rightarrow Q) \Leftrightarrow Q \Rightarrow P$ .

#### 4.4.6 Si y sólo si

Si tenemos que probar un si-y-sólo-si ( $\Leftrightarrow$ ), podemos probar por separado  $\Rightarrow$  y  $\Leftarrow$ . Es muy común que uno de los dos sea prácticamente trivial, y el otro sea el difícil. Por ejemplo:

##### Teorema 8 (Erdős-Gallai)

Sea  $d_1 \geq d_2 \geq \dots \geq d_n$  una secuencia de números naturales no-creciente. Entonces existe un grafo  $G = (V, E)$  con  $|V| = n$  y  $d_G(v_i) = d_i$  para todo  $1 \leq i \leq n$ , si y sólo si  $\sum_i d_i$  es par y para todo  $1 \leq k \leq n$ , tenemos

$$\sum_{i=1}^k d_i \leq k(k-1) + \sum_{i=k+1}^n \min(d_i, k)$$

Probar que una secuencia gráfica cumple eso es fácil. El probar que si eso se cumple para una secuencia, entonces existe un grafo con esa secuencia gráfica, es bastante difícil. Una demostración constructiva es dada por el algoritmo de Havel-Hakimi.

Esencialmente lo que está pasando en esa demostración es que uno de los lados del si-y-sólo-si es una condición global (el existir un grafo que cumple con lo pedido), mientras que el otro lado es un montón de condiciones locales (una por cada  $k$ ). Es fácil implicar cada condición local, pero probar que la unión de todas las condiciones locales implica la condición global es difícil.

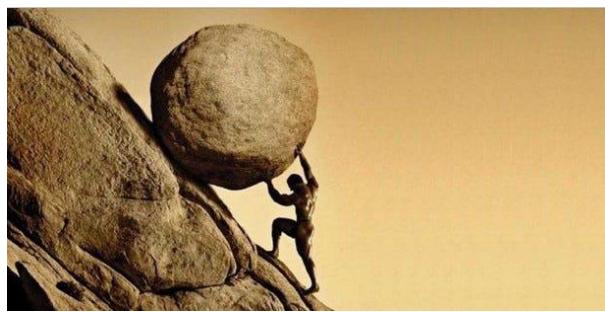


Figura 1: Sísifo probando la vuelta de Erdős-Gallai.



Figura 2: Sísifo probando la ida de Erdős-Gallai.

##### Advertencia

Uno puede probar un si-y-sólo-si mediante una cadena de  $\Leftrightarrow$ , pero tiene que tener cuidado que absolutamente todos los pasos que uno haga a las proposiciones que está manejando, sean equivalencias, y no sólo implicaciones. Por esto frecuentemente es más fácil hacer cada

implicación por separado, y luego si uno se da cuenta que se puede hacer ambas al mismo tiempo, reescribirlo de esa forma.

Por ejemplo, la siguiente demostración es incorrecta porque usa una implicación pero dice usar un si-y-sólo-sí.

**Proposición 4.4.7**

$$-2 = 2.$$



*Demostración.* Sea  $x = -2$ . Entonces:

$$\begin{aligned}x &= -2 \Leftrightarrow \\x^2 &= 4 \Leftrightarrow \\\sqrt{x^2} &= \sqrt{4} \Leftrightarrow \\x &= 2\end{aligned}$$



El error está en que  $x = 2$  implica  $\sqrt{x^2} = \sqrt{4}$ , pero la vuelta no vale. Para probar una cadena de si-y-sólo-sí, absolutamente todos los pasos deben ser si-y-sólo-sí. Con que haya una implicación sin vuelta cierta, todo está mal.

Sucede lo mismo dando vuelta el argumento:

*Demostración.* Sea  $x = 2$ . Entonces:

$$\begin{aligned}x &= 2 \Leftrightarrow \\\sqrt{x^2} &= \sqrt{4} \Leftrightarrow \\x^2 &= 4\end{aligned}$$

Por lo tanto cualquier solución a  $x^2 = 4$  es solución de nuestra ecuación, y en particular  $-2 = 2$ .



#### 4.4.7 Partir en casos

Si tenemos que probar  $\forall x \in X. P(x)$  y podemos dividir el dominio  $X$  de forma productiva, donde cada subconjunto de  $X$  tiene una demostración de  $P$  simple pero mayormente independiente, podemos partir en casos. Por ejemplo:

**Ejercicio 4.4.8**

Sea  $n \in \mathbb{Z}$ . Entonces  $n(n + 1)$  es par.



*Demostración.* Partimos en casos.

1. Si  $n$  es par, entonces existe  $k \in \mathbb{Z}$  tal que  $n = 2k$ . Luego,  $n(n + 1) = 2k(2k + 1)$ . Llamando  $m = k(2k + 1)$ , vemos que  $n(n + 1) = 2m$ , con lo cual  $n(n + 1)$  es par.

2. Si  $n$  es impar, entonces existe  $k \in \mathbb{Z}$  tal que  $n = 2k + 1$ . Luego,  $n(n+1) = (2k+1)(2k+1+1) = (2k+1)(2k+2) = (2k+1)2(k+1)$ . Llamando  $m = (k+1)(2k+1)$ , tenemos que  $n(n+1) = 2m$ , con lo cual  $n(n+1)$  es par.

□

Cuando hacemos esto, es importante tener en cuenta cuán difícil es la demostración en cada caso. Si partimos en, por ejemplo,  $n = 0$  y  $n \neq 0$ , pero uno de los casos es mucho más difícil que el otro, entonces el caso que es más difícil requiere mucho más esfuerzo y detalle. A veces vemos alumnos que hacen el caso simple con detalle, y el caso complejo lo dejan vago, lo cual es bastante inútil.

#### 4.4.8 Unicidad

Si tenemos que probar que «existe un único  $x$  en  $X$  tal que  $P(x)$ », una estrategia común es tomar dos objetos que cumplan  $P(x)$ , y concluir que son el mismo.

##### Ejercicio 4.4.9

Dado un grupo  $G$ , existe un elemento  $e \in G$  tal que para todo  $g \in G$ , tenemos  $eg = ge = g$ , llamado la identidad del grupo. Probar que existe un único tal elemento.

♦

*Demostración.* Sean  $e_1, e_2 \in G$  tal que para todo  $g \in G$ ,  $e_1g = ge_1 = g$  y  $e_2g = ge_2 = g$ . Como  $e_1$  es una identidad, multiplicamos a la izquierda para obtener  $e_1e_2 = e_2$ . Como  $e_2$  es una identidad, multiplicamos a la derecha para obtener  $e_1e_2 = e_1$ . Luego  $e_1 = e_2$ . □

♦

Otro ejemplo de unicidad, pero en grafos y árboles.

##### Ejercicio 4.4.10

Un ciclo es un circuito simple, es decir, que no repite vértices.

Sea  $G = (V, E)$  un grafo,  $T = (V, E_T)$  un árbol generador, y  $e \in E \setminus E_T$ . Entonces  $T' = (V, E_T \cup \{e\})$  tiene un único ciclo.

♦

*Demostración.* Sean  $\{u, v\} = e$  los extremos de  $e$ . Como  $T$  es un árbol, para todo par de vértices en  $T$ , hay un único camino entre ellos en  $T$ . Luego, sea  $P = [u, x_1, x_2, \dots, x_k, v]$  el camino en  $T$  entre  $u$  y  $v$ . Entonces  $C = [u, x_1, x_2, \dots, x_k, v, u]$  es un circuito en  $T'$  que usa  $e$ . Tenemos que probar que este circuito es simple, y que no hay otro ciclo en  $T'$ .

Si  $C$  no fuera simple, entonces existe un vértice  $w$  en  $C$  que  $C$  visita dos veces. Luego  $C = [u, x_1, x_2, \dots, x_{i-1}, x_i = w, x_{i+1}, \dots, x_{j-1}, x_j = w, x_{j+1}, \dots, v, u]$ , con  $j - i > 0$  (si no,  $C$  no está visitando  $w$  dos veces). Luego aún sin la arista  $\{v, u\}$ , tenemos un ciclo  $[x_i = w, x_{i+1}, \dots, x_{j-1}, x_j = w] \subseteq E_T$  con longitud  $j - i > 0$ , pero  $T$  era un árbol, luego acíclico, y esto no puede suceder. Luego  $C$  es simple.

Entonces,  $C$  es simple. Sea  $C'$  otro ciclo simple en  $T'$ . Si  $C'$  no contiene a  $e$ , entonces  $C'$  está también en  $T$ , lo cual no puede pasar por ser  $T$  acíclico. Entonces escribimos  $C' = [u, y_1, \dots, y_q, v, u]$ . Esto nos da dos caminos entre  $u$  y  $v$  en  $T$ :  $P = [u, x_1, \dots, x_k, v]$ , y  $P' =$

$[u, y_1, \dots, y_q, v]$ . Pero como  $T$  es un árbol, existe sólo un camino entre  $u$  y  $v$  en  $T$ , y luego  $[x_1, \dots, x_k] = [y_1, \dots, y_q]$ .

Luego,  $C' = C$ , que es lo que queríamos demostrar.  $\square$

A veces vamos a usar el contrarecíproco para probar unicidad, postulando que existen dos objetos distintos que cumplen una propiedad, y llegando a un absurdo. Les reitero que no usan el contrarecíproco mecánicamente, pero sí que lo conozcan como herramienta.

## 4.5 Pasar en limpio

Gran parte de una demostración es jugar con los objetos, e intentar ver qué sucede. Eventualmente, uno llega a un argumento formal sólido. Sin embargo, al comunicarle este argumento a alguien, no hace falta comunicar todas las cosas que pensamos, las ecuaciones que no llevaron a nada, los errores que cometimos, los ejemplos que intentamos, los dibujos que nos confundieron, etcétera.

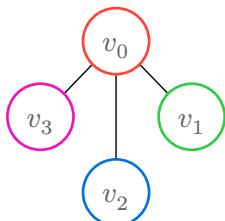
Es difícil comunicar el patrón incoherente de ideas que pasan por la cabeza mientras uno juega, pero el siguiente es un intento de mostrarlo, y luego la demostración final que uno pasa en limpio, obviando todos los caminos sin salida. No se supone que el texto a continuación sea totalmente comprensible, es sólo un camino vueltero que pueden transitar al jugar.

### Ejercicio 4.5.1

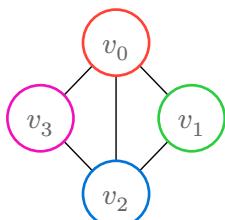
Sea  $G$  un grafo. Determinar si es cierto que si  $G$  tiene exactamente dos vértices de grado impar, entonces existe en  $G$  un camino entre ellos. Si es cierto, demostrarlo. Si es falso, dar un contraejemplo.

Hrm, OK. ¿Qué raro. ¿Cómo puede ser que algo totalmente local, como el grado de un vértice, concluya en algo global, como la existencia de un camino que puede ser larguísimo? A priori no le creo, vamos a intentar hacer un contraejemplo...

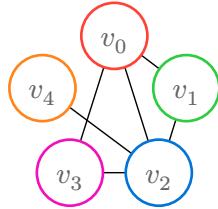
Empiezo con algo así, donde quiero que haya exactamente un vértice de grado impar. Si logro eso, después simplemente pongo dos copias de esto lado a lado, sin juntarlas con ninguna arista, y voy a tener dos vértices de grado impar que no se conectan.



Bueno pero eso tiene a las tres hojas con grado 1, entonces no hay exactamente 2 vértices de grado impar... a ver si juntándolas pasa algo...



Bueno eso no queda bien porque ahora  $v_2$  también tiene grado impar, y hay un camino... A ver, intento hacer que  $v_2$  tenga grado par otra vez. Como ya está conectado a los otros vértices, tengo que agregar uno. Si agrego uno solo, entonces no va a haber manera de lograr lo que quiero...



Eso arregla  $v_2$ ... pero  $v_4$  está roto... y sin importar cómo uno a  $v_4$ , no puedo hacer que haya exactamente *un* vértice de grado impar en el grafo resultante...

Parece que ser que hay algo invariante, que el número de vértices de grado impar, es par... ¿puede ser algo aritmético eso? A ver, le pongo nombre a las cosas, así puedo usar aritmética y ecuaciones...

Separaremos los vértices en los que tienen grado impar,  $V_1$ , y los que tienen grado par,  $V_2$ . La suma de los grados de todos los vértices ya sabíamos que es  $2m$ ... entonces  $\sum_{v \in V_1} d_G(v) + \sum_{v \in V_2} d_G(v) = \sum_{v \in V} d_G(v) = 2m$ . Como  $d_G(v)$  es impar para todos los vértices  $v$  en  $V_1$ , entonces quizás la suma esa da algo impar... pero la otra suma, sobre  $V_2$ , da algo par...

Entonces tendría algo como **impar + par = par**, y eso no puede pasar. El **impar** va a ser cuando el número de sumandos sea impar, porque un número impar de números impares, sumados, da resultado impar.

¿Cómo puedo usar esto para lo que me piden?

Si no hay un camino entre los dos vértices,  $u$  y  $v$ , entonces están en dos componentes conexas. Pero entonces puedo hacer este argumento en cada componente conexa, como si fuera cada una un grafo por separado. No puede ser que haya *un* vértice de grado impar en cada componente, porque estaríamos en la condición **impar + par = par** de arriba, en cada uno de estos subgrafos.

OK, creo que eso cierra. A ver cómo se puede escribir bien...

*Demostración.* Sea  $G = (V, E)$  un grafo. Queremos ver que si hay exactamente dos vértices con grado impar, entonces están en la misma componente conexa.

Asumo que no. Entonces existen dos componentes conexas en  $G$ ,  $G_1 = (V_1, E_1)$  y  $G_2 = (V_2, E_2)$ , tal que  $V_1 \cap V_2 = \emptyset$ , y hay dos vértices,  $v_1$  y  $v_2$ , tal que  $v_1 \in V_1$ ,  $v_2 \in V_2$ ,  $d_G(v_1) = d_{G_1}(v_1)$  es impar, y  $d_G(v_2) = d_{G_2}(v_2)$  es impar. Sabemos que  $d_G(v_1) = d_{G_1}(v_1)$  porque  $v_1$  sólo tiene aristas hacia  $V_1$  (porque definimos  $V_1$  como la componente conexa donde está  $v_1$ , no puede haber una arista desde  $v_1$  a alguien en  $V_2$ , pues entonces ese alguien estaría en la componente  $V_1$ , pero sabemos que  $V_1 \cap V_2 = \emptyset$ .)

Entonces  $G_1$  y  $G_2$  son subgrafos de  $G$ , y en cada uno existe un único vértice con grado par.

Probemos el siguiente lema.

### Lema 4.5.2

Sea  $H$  un grafo. Entonces el número de vértices en  $H$  de grado impar, es par.



*Demostración.* Sean  $W_1$  los vértices de  $H$  de grado impar, y  $W_2$  los vértices de  $H$  de grado par. Si los vértices de  $H$  son  $W$ , sabemos que  $W = W_1 \sqcup W_2$ . Entonces si  $H$  tiene  $m$  aristas, sabemos que  $\sum_{v \in W} d_H(v) = 2m$ . Pero como  $W = W_1 \sqcup W_2$ , esta suma es igual a  $\sum_{v \in W_1} d_H(v) + \sum_{v \in W_2} d_H(v)$ . Por claridad llamemos  $X = \sum_{v \in W_1} d_H(v)$ , y  $Y = \sum_{v \in W_2} d_H(v)$ .

Como llamamos a  $W_1$  el subconjunto de vértices de  $H$  de grado impar, sabemos que  $d_H(v) \equiv 1 \pmod{2}$  para todo  $v \in W_1$ . Luego la suma de todos ellos es congruente a  $|W_1|$ , módulo 2. Es decir,  $X \equiv |W_1| \pmod{2}$ .

Con un argumento análogo, vemos que  $Y \equiv 0 \pmod{2}$ .

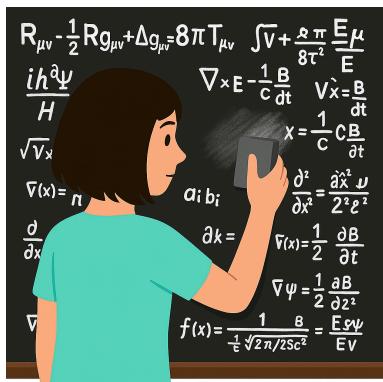
Entonces, tenemos que  $X + Y = 2m$ , con  $X \equiv |W_1| \pmod{2}$  y  $Y \equiv 0 \pmod{2}$ .

Tomando módulo dos a ambos lados de la primer ecuación, nos queda que  $|W_1| + 0 = 0 \pmod{2}$ , y luego  $|W_1|$  es par.

Luego, el número de vértices en  $H$  con grado impar, es par. □

Entonces llamando  $H = G_1$ , vemos que no puede ser que haya en  $G_1$  exactamente un vértice de grado impar. Luego, lo que asumimos tiene que estar mal, que había dos componentes conexas, una con un vértice de grado impar cada una.

Luego estos dos vértices tienen que estar en la misma componente conexa, y luego hay un camino de uno al otro. □



Hrm, quedó medio desordenado eso. Mejor lo emprolijo un poco. Puedo poner el lema primero así cuando lo uso ya lo tengo probado, no necesito usar  $G_2$ , no necesito aclarar que  $d_G(v) = d_{G_1}(v)$  para  $v$  en  $G_1$ , y quedó medio confusa la oración sobre  $V_1 \cap V_2 = \emptyset$ ...

*Demostración.* Primero, probemos un lema.

### Lema 4.5.3

Sea  $H$  un grafo. Entonces el número de vértices en  $H$  de grado impar, es par.



*Demostración.* Sean  $W_1$  los vértices de  $H$  de grado impar, y  $W_2$  los vértices de  $H$  de grado par. Si los vértices de  $H$  son  $W$ , sabemos que  $W = W_1 \sqcup W_2$ . Entonces si  $H$  tiene  $m$  aristas, sabemos que  $\sum_{v \in W} d_H(v) = 2m$ . Pero como  $W = W_1 \sqcup W_2$ , esta suma es igual a  $\sum_{v \in W_1} d_H(v) + \sum_{v \in W_2} d_H(v)$ . Por claridad llamemos  $X = \sum_{v \in W_1} d_H(v)$ , y  $Y = \sum_{v \in W_2} d_H(v)$ .

Como llamamos a  $W_1$  el subconjunto de vértices de  $H$  de grado impar, sabemos que  $d_H(v) \equiv 1 \pmod{2}$  para todo  $v \in W_1$ . Luego,  $\sum_{v \in W_1} 1 \pmod{2} \equiv |W_1| \pmod{2}$ . Es decir,  $X \equiv |W_1| \pmod{2}$ .

Con un argumento análogo, vemos que  $Y \equiv 0 \pmod{2}$ , por ser la suma de varios números pares.

Entonces, tenemos que  $X + Y = 2m$ , con  $X \equiv |W_1| \pmod{2}$  y  $Y \equiv 0 \pmod{2}$ .

Tomando módulo 2 a ambos lados de la primer ecuación, nos queda que  $|W_1| + 0 = 0 \pmod{2}$ , y luego  $|W_1|$  es par.

Luego, el número de vértices en  $H$  con grado impar, es par. □

Ahora sea  $G = (V, E)$  un grafo con exactamente dos vértices de grado impar,  $v_1$  y  $v_2$ . Asumamos, por contradicción, que  $v_1$  y  $v_2$  están en componentes conexas distintas. Entonces existen  $V_1, V_2 \subset V$ , tales que  $V_1 \cap V_2 = \emptyset$ ,  $v_1 \in V_1$ ,  $v_2 \in V_2$ , y  $d_G(v_1) \equiv d_G(v_2) \equiv 1 \pmod{2}$ . Tomemos la componente conexa inducida por  $V_1$ , llamémosla  $G_1$ . Sabemos que  $v_1 \in G_1$ , y  $v_2 \notin G_1$ . Como  $G$  tiene exactamente dos vértices con grado impar, son sólo  $v_1$  y  $v_2$ , entonces en  $G_1$  hay exactamente un vértice de grado impar, y es  $v_1$ .

Por el Lema 4.5.3, esto no puede pasar. Luego, lo que asumimos por contradicción era falso, y entonces esos dos vértices no están en componentes conexas distintas. Luego están en la misma componente conexa, y luego hay un camino entre ellos. □

Si sólo ven la demostración final, parece compacta, no comete errores, no intenta varias cosas, no nombra cosas que no usa, no deja cosas sin demostrar para después, tiene notación sensible, y hasta tiene estructura, probando un sub-lemma antes de usarlo. No piensen que la demostración nació así - como ven, uno pasa por jugar, probar cosas, planear, y emprolijar. No se frustren si sus demostraciones no se ven como esta última, en su primer pasada.

## 5 Errores comunes

### 5.1 Ser informal

Este es **de lejos** el error que más cometan los alumnos. En este momento de su educación, todavía no les recomiendo dejar de lado la formalidad, y hacer argumentos informales. Un argumento informal puede ser riguroso, pero requiere experiencia hacer esto sin cometer errores. Todavía no tienen esa experiencia. Por ende, a la hora de argumentar, sean formales. Algunas recomendaciones sobre formalidad:

1. Pónganle nombre a todo.

- Si un sustantivo no tiene nombre, no podemos hablar de él claramente. Muchas veces se quedan «sin saber cómo seguir», porque no tienen a mano suficientes sustantivos para ver relaciones entre ellos, o ver qué cosas cumple cada uno.

Ante el siguiente ejercicio:

**Ejercicio 5.1.1**

Demostrar, usando inducción en el número de vértices, que todo grafo de  $n$  vértices que tiene más de  $\frac{(n-1)(n-2)}{2}$  aristas es conexo.

Esta es una duda de un alumno:

«Ok, le saco  $v$  a  $G_{n+1}$ , el tema es que  $G_{n+1}$  tenía mas de  $\frac{n(n-1)}{2}$ , y le saco por ejemplo un  $v$  con  $n - 1$  aristas, me queda que  $G_n$  tiene mas de  $\frac{n(n-1)}{2} - (n - 1) = \frac{(n-1)(n-2)}{2}$  aristas y por tanto es conexo y ahí ya estoy. Si el  $v$  que sacara tuviera menos aristas, no sé si sigue andando.»

La duda acá sale de no haberle puesto nombre a todo. En particular, usar  $d_{G_{n+1}}(v)$ , que es el número de aristas que estamos sacando al sacar  $v$  de  $G_{n+1}$  para obtener  $G_n$ . El ponerle nombre nos deja usar álgebra. En particular, sabemos que  $0 \leq d_{G_{n+1}}(v) \leq n - 1$ , puesto que como máximo,  $v$  está conectado a los otros  $n$  vértices de  $G_{n+1}$ . Vamos a tener que argumentar cuidadosamente usando esto.

*Demostración.* Vamos a probar la proposición  $P(n)$ : Si  $G$  es un grafo con  $n$  vértices y más de  $\frac{n(n-1)}{2}$  aristas, entonces  $G$  es conexo.

$P(0)$  vale trivialmente porque no hay grafos sin vértices.  $P(1)$ : Un grafo con  $n = 1$  vértice siempre es conexo.

Ahora sea  $n \geq 2$ . Asumo que vale  $P(n - 1)$ , quiero ver que vale  $P(n)$ . Sea  $G = (V, E)$  un grafo de  $n$  vértices, con más de  $\frac{(n-1)(n-2)}{2}$  aristas, y  $v$  cualquier vértice en  $V$ . Sea  $m = |E|$ .

Consideremos  $G' = G - v = (V - \{v\}, E')$ , el grafo que resulta de sacarle  $v$  a  $G$ , junto con todas las aristas incidentes a  $v$  en  $G$ . Notar que sólo puedo hacer esto porque  $n \geq 2$ . No podría sacarle un vértice a un grafo que sólo tuviera un vértice<sup>6</sup>.

Sabemos que  $0 \leq d_G(v) \leq n - 1$ , porque como mínimo  $v$  tiene cero vecinos en  $G$ , y como máximo tiene a todos los otros  $n - 1$  vértices de  $G$  como vecinos.

Partimos en casos.

- Si  $d_G(v) = 0$ , entonces  $v$  es un vértice aislado. Esto no puede pasar, porque  $G$  tiene  $m > \frac{(n-1)(n-2)}{2}$  aristas, y aún poniendo una arista entre *todo* otro par de vértices en  $G$ , nos quedarían sólo  $\frac{(n-1)(n-2)}{2}$  aristas. Como  $m$  es mayor que  $\frac{(n-1)(n-2)}{2}$ , tiene que haber al menos una arista incidente a  $v$ .
- Si  $d_G(v) = n - 1$ , entonces  $v$  comparte una arista con cada uno de los otros vértices de  $G$ . Luego para cualquier par de vértices  $u, w \in V$ , tenemos un camino  $[\{u, v\}, \{v, w\}]$  en  $G$ , y luego  $G$  es conexo, que es lo que queríamos demostrar.
- Caso contrario,  $0 < d_G(v) \leq n - 2$ . Como sabemos que  $m > \frac{(n-1)(n-2)}{2}$ , al sacarle  $d_G(v)$  aristas a  $G$ , y sabiendo que  $d_G(v) \leq n - 2$ , obtenemos  $|E'| = m - d_G(v) \geq$

<sup>6</sup>En la carrera generalmente requerimos, en la definición de «grafo», que haya al menos un vértice.

$m - (n - 2) > \frac{(n-1)(n-2)}{2} - (n - 2) = \frac{(n-2)(n-3)}{2}$  aristas en  $G'$ . Por hipótesis inductiva, como  $G'$  es un grafo de  $n - 1$  vértices con más de  $\frac{(n-2)(n-3)}{2}$  aristas, sabemos que  $G'$  es conexo. Finalmente, como sabemos que  $d_G(v) > 0$ , al agregar  $v$  a  $G'$  con todas sus  $d_G(v)$  aristas que tenía en  $G$ , estamos conectando  $v$  con un grafo conexo ( $G'$ ) con al menos una arista, y luego  $G$  es conexo.

□

2. No usen el mismo nombre para dos cosas distintas. Si están modificando un objeto, no usen el mismo nombre para el objeto antes y después de modificarlo.
3. Si el objeto  $X$  depende de un objeto  $Y$ , nómbrélo  $X_Y$  o  $X(Y)$ , para recordar la dependencia.
4. Si terminan definiendo un sustantivo  $y$  y no lo usan para su conclusión, o no es necesario, pueden removerlo al terminar. Pero si no empezamos dándole nombre, seguro no lo podemos usar.
2. Cuantifiquen todo.
  1. Si usan una variable, cuantifíquenla. Una variable sin cuantificar es inútil. « $G$  es conexo.» ¿Quién es  $G$ ? ¿Vale para todo  $G$ ? ¿Existe algún  $G$ ? ¿Es un  $G$  particular que definimos nosotros?
  2. Presten atención al anidado de cuantificadores. En  $\forall x \in X. \exists y \in Y. P(x, y)$ ,  $y$  puede depender de  $x$ , pero  $x$  no puede depender de  $y$ . En  $\exists x \in X. \forall y \in Y. P(x, y)$  es una oración completamente distinta, no tienen nada que ver una con la otra. Recuerden la Sección 4.2, donde interpretamos demostraciones como una conversación entre nosotros y alguien que no está pidiendo demostrarles algo.
3. Usen ecuaciones y desigualdades. En vez de decir «El peor caso es que  $m = n$ », digan explícitamente que  $m \leq n$ , o  $m \geq n$ , sea cual fuere el caso. Muchas veces cometen el error de asumir que un objeto es «un peor caso» (y luego basta probar lo que tienen que probar sólo para ese objeto), pero están confundiéndose con la dirección de la desigualdad. Razonen formalmente, usen ecuaciones y desigualdades.
4. Usen lenguaje formal, cuando existe. La oración «La función seno se ve igual cada  $2\pi$ .» es vaga. ¿Qué significa «se ve igual»? ¿Quién la «ve», y cómo? Escribir esto con precisión resultaría en «Para todo  $x \in \mathbb{R}$ ,  $\sin(x) = \sin(x + 2\pi)$ », que es preciso, y nos da una ecuación con la cual trabajar y reemplazar en el futuro.
5. Sean claros en qué es lo que afirman, qué es lo que asumen, y qué es lo que quieren probar. Ver un montón de oraciones donde todas son afirmaciones dificulta la comprensión. Usen conectores lógicos, como «porque», «luego», «si», y «entonces». Pueden usar frases como «Vamos a probar que», «Acá usamos la hipótesis tal», «Asumimos por contradicción que tal cosa», «Vamos a usar tal estrategia (inducción, partir en casos, etcétera)».

## 5.2 No decir nada

Si la demostración les salió en una oración, está mal. Generalmente veo esto cuando sólo están reiterando el enunciado, o reiterando la conclusión, y no hay ningún argumento en el medio que los conecte. Saben que tienen que ir de  $P$  a  $Q$ , entonces saben que al menos van a tener que mencionar a  $P$  y a  $Q$ , se confunden, y sólo dicen « $P$ . Luego  $Q$ .», o sólo « $Q$ .». Si la demostración fuera una sola oración, no sería un ejercicio de una materia universitaria. Si un ejercicio les resultó totalmente

trivial, probablemente lo hicieron mal. Vuelvan a leer la Sección 4.2, sobre comprender qué nos están pidiendo.

### ⚠️ Advertencia

Esto es un ejemplo de un alumno, donde sólo se reitera lo que hay que probar.

#### Ejercicio 5.2.1

Dado un grafo  $G = (V, E)$  y un vértice  $v \in V$ , un árbol generador  $T$  de  $G$  es  $v$ -geodésico si  $\text{dist}_G(v, w) = \text{dist}_T(v, w)$  para todo  $w \in W$ .

Sea  $T$  un árbol que genera BFS al comenzar desde un vértice  $v$ . Probar que  $T$  es  $v$ -geodésico.

*Demostración.* El árbol que queda explícitamente definido después de correr BFS en  $G$  con el vértice  $v$  cumple que  $\text{dist}_T(v, w)$  es igual a  $\text{dist}_G(v, w)$  para todo  $w \in V$ . Por lo tanto si  $T$  es el árbol de BFS en  $G$  enraizado en  $v$ , entonces queda probado que  $T$  es  $v$ -geodésico. □

Esto no prueba nada. Sólo reitera la conclusión.

Esto sucede también cuando afirman proposiciones sin probarlas, a veces simplemente diciendo que «es obvio». La proposición «Todo árbol de al menos dos vértices tiene al menos dos hojas» es «obvia», e imagino que la mayoría no la puede probar fácilmente. Frecuentemente piensan en varios ejemplos, todos cumplen una propiedad, y concluyen que «es obvio». Lo que es obvio es *que esos ejemplos la cumplen*, lo que no es obvio es *cómo demostrar que todos los objetos la cumplen*. Después de todo, esta conjectura:

#### Conjetura 5.2.2 (Goldbach)

Para todo  $n \in \mathbb{N}$  par,  $n > 2$ , existen dos números primos  $p, q \in \mathbb{N}$ , tal que  $n = p + q$ .

jamás ha sido probada, aún después de cientos de años de intentos. Sin embargo, absolutamente todos los números naturales pares que ustedes piensen van a cumplirla, porque no se le conocen contraejemplos. Sólo afirmarla porque no se nos ocurren contraejemplos no dice nada.

En general, pueden usar sin probar (pero mencionando qué es lo que están usando!) lo que hayan visto en materias correlativas, y en el material demostrado en clase. Si quieren usar algo más, pregunten si lo pueden usar. La respuesta muchas veces va a ser «Lo podés usar sólo si lo podés probar», que es lo mismo que «No».

### 5.3 Empezar con la conclusión

No empiecen con la conclusión e intenten probar la premisa. Si logran hacer esto, de milagro, usando sólo implicaciones bilaterales ( $\Leftrightarrow$ ), en el mejor caso es una pobre y confusa exposición de la implicación pedida. En el peor caso, casi siempre van a cometer el error de usar una implicación que no tiene vuelta válida, y su demostración no va a decir nada.

### ⚠️ Advertencia

Asumir la conclusión y probar algo cierto no dice nada.

### Postulado 5.3.1

$$1 = -1$$

♥

#### Demostración.

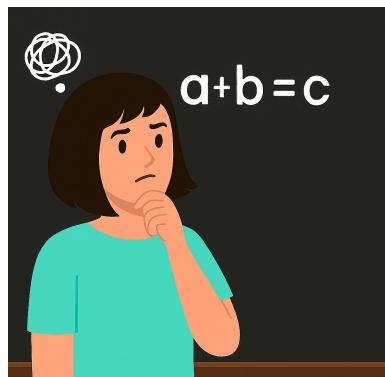
$$\begin{aligned} 1 &= -1 && \text{hago lo mismo en ambos lados} \\ 1^2 &= (-1)^2 && \text{simplifico} \\ 1 &= 1 \end{aligned}$$

Como llegamos a algo cierto, debemos haber empezado con algo cierto. □

Si llegaron a algo razonando así, este es el momento para emprolijar su idea, y empezar desde el principio. No se preocupen porque «parezca un galerazo» que su objeto justo cumpla lo pedido al final. Esa es una preocupación pedagógica del docente, no de ustedes.

## 5.4 No entender qué estamos asumiendo

Frecuentemente vemos que usan algo sin saber que lo están usando. Por ejemplo, «sean  $u, v \in V$  tales que hay un camino entre  $u$  y  $v$ ». Nada en esto nos deja concluir que  $u \neq v$ . Si luego asumimos que el camino tiene longitud mayor a cero, por ejemplo diciendo «sea  $P = [u, x, \dots, v]$  el camino entre  $u$  y  $v$ , y tomemos la arista  $(u, x)$ » (que no tiene por qué existir), estamos asumiendo algo sin siquiera entender que lo estamos haciendo.



Mismo cuando decimos «sea  $x$  el vértice tal que  $P(x)$ ». El decir *el* vértice implica que existe un único tal vértice que cumple  $P$ . Esto ambos requiere probar que existe, y que no hay otro que cumpla  $P$ . Si sólo queremos decir que existe (y lo demostramos anteriormente), podemos decir «sea  $x$  un vértice tal que  $P(x)$ ».

Finalmente, a veces lo que asumen no es explícito. Por ejemplo, si se les pide probar que  $a$  y  $b$  comutan (es decir, que  $ab = ba$ ), y usan  $(ab)^2 = a^2b^2$ , están precisamente asumiendo commutatividad para probar commutatividad.

## 6 Ejemplos

A continuación les voy a mostrar varios ejemplos de demostraciones matemáticas. Los temas están sacados del contenido de la carrera. Todas son rigurosas y formales. La idea de estos ejemplos es que vean varias estrategias de demostración puestas en práctica en una variedad de escenarios. Es posible que haya ejercicios que mencionen objetos que no conocen todavía, no hay ningún problema con eso, no hace falta que entiendan *todos* los ejemplos.

Sugiero *fuertemente* no leer la solución sin haber intentado resolver cada ejercicio, durante al menos una hora. Como dije en la Sección 3, no aprenden cuando terminan un ejercicio, y aprenden muy poco cuando ven cómo alguien más resuelve un ejercicio. El aprendizaje sucede principalmente cuando intentan, durante horas, resolver ejercicios. Intenten jugar con los objetos, pónganle nombre a todo, usen ecuaciones, vean qué pueden deducir, y no se rindan si no les sale en los primeros 20 minutos.

## 6.1 Conjuntos

### Ejercicio 6.1.1

Sean  $A, B$  conjuntos. Entonces

$$(A \cup B)^c = A^c \cap B^c$$



*Demostración.* Vamos a probar una igualdad de conjuntos probando que cada uno está incluido en el otro.

- $(A \cup B)^c \subseteq A^c \cap B^c$ : Sea  $x \in (A \cup B)^c$ . Entonces  $x \notin (A \cup B)$ . Como  $A \cup B = \{y \mid y \in A \text{ o } y \in B\}$ , entonces tenemos que es falso que  $(x \in A \text{ o } x \in B)$ , y luego tenemos que  $x \notin A$  y  $x \notin B$ . Pero  $A^c = \{y \mid y \notin A\}$ , y análogamente para  $B^c$ , luego  $x \in A^c$  y  $x \in B^c$ . Por definición de  $\cap$ , entonces,  $x \in A^c \cap B^c$ .
- $A^c \cap B^c \subseteq (A \cup B)^c$ . Sea  $x \in A^c \cap B^c$ . Luego por definición de  $\cap$ ,  $x \in A^c$  y  $x \in B^c$ . Luego por definición de  $X^c$  para conjuntos  $X$ ,  $x \notin A$ , y  $x \notin B$ . Luego  $x \notin A \cup B$ , dado que  $A \cup B = \{y \mid y \in A \text{ o } y \in B\}$ . Luego  $x \in (A \cup B)^c$ .



### Ejercicio 6.1.2

Sean  $A, B$  conjuntos. Probar que  $B \setminus (B \setminus A) = A \cap B$ .



*Demostración.* Hagamos esto enteramente por definición.

$$\begin{aligned} B \setminus (B \setminus A) &= \{x \mid x \in B \wedge x \notin (B \setminus A)\} \\ &= \{x \mid x \in B \wedge x \notin \{y \in B \mid y \notin A\}\} \\ &= \{x \mid x \in B \wedge (x \notin B \vee (x \in B \wedge x \in A))\} \\ &= \{x \mid x \in B \wedge x \in A\} \\ &= A \cap B \end{aligned}$$



### Ejercicio 6.1.3

Sean  $A, B, C$  conjuntos. Probar que  $(A \setminus B) \setminus C \subseteq A \setminus (B \setminus C)$ . Mostrar que no vale la igualdad, en general.



*Demostración.* Sea  $x \in (A \setminus B) \setminus C$ . Entonces  $x \in A$ ,  $x \notin B$ , y  $x \notin C$ . Como  $x \notin B$ , entonces con más razon  $x \notin (B \setminus C)$ , dado que  $B \setminus C \subseteq B$ . Luego  $x \in A \wedge x \notin (B \setminus C)$ , y luego  $x \in A \setminus (B \setminus C)$ .

La igualdad no vale en general. Por ejemplo, podemos tomar  $A = \{1, 2, 3\}$ ,  $B = \{2, 3\}$ , y  $C = \{3, 4, 5\}$ . Entonces  $(A \setminus B) \setminus C = \{1\}$ , pero  $A \setminus (B \setminus C) = \{1, 3\}$ . □

### Ejercicio 6.1.4 (Leyes de De Morgan)

Sean  $A, B, C$  conjuntos. Demostrar que:

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

◆

*Demostración.* Para la primer ecuación:

$$\begin{aligned} A \cup (B \cap C) &= \{x \mid x \in A \vee x \in (B \cap C)\} \\ &= \{x \mid x \in A \vee (x \in B \wedge x \in C)\} \\ &= \{x \mid (x \in A \vee x \in B) \wedge (x \in A \vee x \in C)\} \\ &= (A \cup B) \cap (A \cup C) \end{aligned}$$

y para la segunda:

$$\begin{aligned} A \cap (B \cup C) &= \{x \mid x \in A \wedge (x \in B \vee x \in C)\} \\ &= \{x \mid (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)\} \\ &= (A \cap B) \cup (A \cap C) \end{aligned}$$

□

### Ejercicio 6.1.5

Denotemos por  $A \triangle B$  la diferencia simétrica entre conjuntos, es decir,  $A \triangle B = (A \setminus B) \cup (B \setminus A)$ .

Sea  $f : X \rightarrow Y$ , y denotemos por  $f^{-1}(C) = \{x \in X \mid f(x) \in C\}$ , para cualquier subconjunto  $C \subseteq Y$ .

Probar que para todo  $A, B$  subconjuntos de  $Y$ ,  $f^{-1}(A \triangle B) = f^{-1}(A) \triangle f^{-1}(B)$ .

◆

*Demostración.*

- ⊆: Sea  $x \in f^{-1}(A \triangle B)$ . Entonces  $f(x) \in A \triangle B$ , es decir,  $f(x) \in (A \setminus B) \cup (B \setminus A)$ . Partimos en casos:
  1. Si  $f(x) \in A \setminus B$ , entonces  $f(x) \notin B$ , y  $f(x) \in A$ . Luego,  $x \notin f^{-1}(B)$ , y  $x \in f^{-1}(A)$ , y luego  $x \in (f^{-1}(A) \setminus f^{-1}(B)) \subseteq f^{-1}(A) \triangle f^{-1}(B)$ .
  2. Si  $f(x) \in B \setminus A$  sucede algo análogo, y por ende  $x \in f^{-1}(A) \triangle f^{-1}(B)$ .
- ⊇: Sea  $x \in f^{-1}(A) \triangle f^{-1}(B) = (f^{-1}(A) \setminus f^{-1}(B)) \cup (f^{-1}(B) \setminus f^{-1}(A))$ . Partimos en casos:
  1. Si  $x \in f^{-1}(A) \setminus f^{-1}(B)$ , entonces  $f(x) \in A$ , pero  $f(x) \notin B$ . Luego  $f(x) \in (A \setminus B) \subseteq A \triangle B$ , y por lo tanto  $x \in f^{-1}(A \triangle B)$ .
  2. Si  $x \in f^{-1}(B) \setminus f^{-1}(A)$ , pasa algo análogo, y tenemos que  $x \in f^{-1}(A \triangle B)$ .

□

## 6.2 Funciones y análisis asintótico

### Ejercicio 6.2.1

Sea  $f : A \rightarrow B$  una función. Si existe una función  $g : B \rightarrow A$  tal que  $f \circ g = \text{id}_B$ , y  $g \circ f = \text{id}_A$ , entonces  $f$  es biyectiva, y  $f^{-1} = g$ .

*Demostración.* Para probar que  $f$  es biyectiva, tenemos que mostrar que es inyectiva y sobreyectiva.

- Sean  $x, x'$  tal que  $f(x) = f(x')$ . Podemos aplicarle  $g$  a ambos lados de la ecuación, obteniendo  $g(f(x)) = g(f(x'))$ , luego  $(g \circ f)(x) = (g \circ f)(x')$ . El enunciado nos dice que  $g \circ f = \text{id}_A$ , luego esto es  $\text{id}_A(x) = \text{id}_A(x')$ , pero  $\text{id}_A(y) = y$  para todo  $y \in A$ , luego esto nos dice que  $x = x'$ , probando que  $f$  es inyectiva.
- Sea cualquier  $y \in B$ , y definamos  $x = g(y)$ . Entonces podemos aplicar  $f$  a ambos lados, obteniendo  $f(x) = f(g(y))$ , o equivalentemente,  $f(x) = (f \circ g)(y)$ . El enunciado nos dice que  $f \circ g = \text{id}_B$ , entonces sabemos que  $f(x) = \text{id}_B(y)$ . Pero  $\text{id}_B(y) = y$ , y luego  $f(x) = y$ . Luego, para todo  $y \in B$ , encontramos un  $x \in A$  tal que  $f(x) = y$ , probando que  $y$  es sobreyectiva.

Luego  $f$  es biyectiva. Para ver que  $f^{-1} = g$ , tenemos que probar que  $f^{-1}(y) = g(y)$  para todo  $y \in B$ . Sea  $y \in B$ . Como  $f$  es sobreyectiva, existe un  $x \in A$  tal que  $f(x) = y$ . Luego,  $f^{-1}(y) = f^{-1}(f(x)) = x$  por definición de función inversa. Asimismo,  $g(y) = g(f(x)) = (g \circ f)(x) = \text{id}_A(x) = x$ . Luego ambos  $f^{-1}(y)$  y  $g(y)$  son iguales a  $x$ , y luego  $f^{-1}(y) = g(y)$  para todo  $y \in B$ , que es precisamente la definición de  $f^{-1} = g$ .  $\square$

### Ejercicio 6.2.2

Sea la  $T : \mathbb{N} \rightarrow \mathbb{N}$  una función dada por  $T(0) = 1$ , y para todo  $n \in \mathbb{N}$  tal que  $n > 0$ , definimos  $T(n) = 2T(\frac{n}{2}) + \Theta(n \log n)$ . Es decir, existe una función  $h : \mathbb{N} \rightarrow \mathbb{N}$  tal que  $h \in \Theta(n \log n)$ , y  $T(n) = 2T(\frac{n}{2}) + h(n)$  para todo  $n \in \mathbb{N}, n > 0$ .

Probar que  $T \in \Theta(n \log^2 n)$ .

*Demostración.* Podemos usar el teorema maestro, que nos dice que si tenemos una función  $T$  de la forma  $T(n) = aT(\frac{n}{b}) + f(n)$  para todo  $n \in \mathbb{N}, n > 0$ , y  $f \in \Theta(n^{\log_b(a)} \log^k n)$  para algún  $k \in \mathbb{N}$ , entonces  $T \in \Theta(n^{\log_b(a)} \log^{k+1}(n))$ . Basta elegir  $k = 1$ ,  $a = 2$ ,  $b = 2$ , para ver que estamos dentro de las condiciones de este caso del teorema, y como  $\log_2(2) = 1$ , tenemos que  $T \in \Theta(n \log^2 n)$ .  $\square$

### Ejercicio 6.2.3

Sea  $T : \mathbb{N} \rightarrow \mathbb{N}$  una función dada por  $T(0) = 47$ , y para todo  $n \in \mathbb{N}, n > 0$ , definimos  $T(n) = 2T(\frac{n}{2} - 1) + \Theta(n)$ .

Probar que  $T \in \Theta(n \log n)$ .

*Demostración.* No caemos en ninguna de las hipótesis del teorema maestro, porque nuestra  $T$  no es de la forma  $T(n) = aT\left(\frac{n}{b}\right) + f(n)$ .

Recordemos entonces cómo está definida la expresión « $+ \Theta(n)$ ». Significa que existe alguna función  $h : \mathbb{N} \rightarrow \mathbb{N}$ , tal que  $h \in \Theta(n)$ , y  $T(n) = T\left(\frac{n}{2} - 1\right) + h(n)$  para todo  $n \in \mathbb{N}, n > 0$ .

Hay varias formas de resolver esto, pero una es intentando volver a caer en las hipótesis del teorema maestro.  $T$  no cumple lo que queremos por tener la forma incorrecta, pero  $\frac{n}{2} - 1 = \frac{n-2}{2}$ . Esto nos dice que la recurrencia que define  $T$  está meramente movida de  $n$  a  $n - 2$ , y nos sugiere ver qué pasa si definimos  $U(n) = T(n - 2)$ .

$$\begin{aligned} U(n) &= T(n - 2) \\ &= 2T\left(\frac{n-2}{2} - 1\right) + h(n - 2) \\ &= 2T\left(\frac{n}{2} - 1 - 1\right) + h(n - 2) \\ &= 2T\left(\frac{n}{2} - 2\right) + h(n - 2) \\ &= 2U\left(\frac{n}{2}\right) + h(n - 2) \\ &= 2U\left(\frac{n}{2}\right) + g(n) \text{ definiendo } g(n) = h(n - 2) \end{aligned}$$

Esto se parece a lo que pide el teorema maestro, y quizás nos ayude para acotar  $T(n)$ .

Notemos que no podemos simplemente decir que  $g \in \Theta(n)$ , porque no es cierto que  $\Theta(g) = \Theta(h)$  por el mero hecho de tener  $g(n) = h(n - 2)$  (pensar, por ejemplo, qué pasaría si  $h(n) = n!$ , son el mismo conjunto  $\Theta(n!)$  y  $\Theta((n - 2)!)$ ?).

Para analizar en qué caso del teorema maestro caemos, vamos a tener que analizar cuidadosamente a  $g$ . Como  $h \in \Theta(n)$ , entonces existen  $n_0 \in \mathbb{N}, \alpha, \beta > 0 \in \mathbb{R}$ , tal que para todo  $n \in \mathbb{N}, (n \geq n_0 \Rightarrow \alpha n \leq h(n) \leq \beta n)$ . Queremos ver que existen  $n_1 \in \mathbb{N}, \gamma, \delta > 0 \in \mathbb{R}$ , tal que para todo  $n \in \mathbb{N}, (n \geq n_1 \Rightarrow \gamma n \leq g(n) \leq \delta n)$ , que es la definición de  $g \in \Theta(n)$ .

Como sabemos que para todo  $n \in \mathbb{N}, n \geq n_0, \alpha n \leq h(n) \leq \beta n$ , definiendo  $n_1 = n_0 + 2$ , tenemos que para todo  $n \in \mathbb{N}, n \geq n_1$ , se cumple que  $\alpha(n - 2) \leq h(n - 2) \leq \beta n$ . Luego:

$$\begin{aligned} \alpha(n - 2) &\leq h(n - 2) \leq \beta(n - 2) \\ \alpha(n - 2) &\leq g(n) \leq \beta(n - 2) \\ \alpha n - 2\alpha &\leq g(n) \leq \beta n - 2\beta \end{aligned}$$

Recordemos, queremos encontrar  $\gamma, \delta > 0 \in \mathbb{R}$ , y  $n_2 \in \mathbb{N}$  tal que para todo  $n \geq n_2, \gamma n \leq g(n) \leq \delta n$ .

Intentemos tomando  $n_2 = n_1 = n_0 + 2$ . Entonces, por las ecuaciones de arriba vemos que  $g(n) \leq \beta n - 2\beta$ , pero  $\beta n - 2\beta < \beta n$ , porque  $\beta > 0$ . Luego, si tomamos  $\delta = \beta$ , nos aseguramos de que  $g(n) < \beta n - 2\beta < \beta n = \delta n$ , que es la primer parte de lo que tenemos que probar. Para encontrar  $\gamma$ , tenemos que encontrarlo tal que si  $\alpha n - 2\alpha \leq g(n)$ , entonces  $\gamma n \leq g(n)$ . Esto nos dice que  $\gamma n \leq \alpha n - 2\alpha$ . Por ende,  $\gamma \leq \alpha - 2\frac{\alpha}{n}$ . Pero esta cota tiene que valer para todo  $n \in \mathbb{N}, n \geq n_2$ . Si  $n \geq n_2$ , entonces como  $\alpha > 0$ , tenemos  $\frac{\alpha}{n} \leq \frac{\alpha}{n_2}$ , y luego  $\alpha - 2\frac{\alpha}{n} \geq \alpha - 2\frac{\alpha}{n_2} = \gamma$ .

$2\frac{\alpha}{n} \geq \alpha - 2\frac{\alpha}{n_2}$ . Luego, basta tomar  $\gamma = \alpha - 2\frac{\alpha}{n_2}$ . Verifiquemos si esto cumple lo que necesitamos.

Queremos ver que para todo  $n \in \mathbb{N}$ ,  $n \geq n_2$ , tenemos que  $\gamma n \leq g(n) \leq \delta n$ . Como  $n \geq n_2$ , y  $n_2 = n_0 + 2$ , entonces  $n - 2 \geq n_0$ . Luego, tenemos que  $\alpha(n - 2) \leq h(n - 2) \leq \beta(n - 2)$ . Esto no es nada menos que  $\alpha(n - 2) \leq g(n) \leq \beta(n - 2)$ . Como  $g(n) \leq \beta(n - 2)$ , y  $\beta(n - 2) < \beta n = \delta n$ , tenemos que  $g(n) < \delta n$ . Como  $g(n) \geq \alpha(n - 2)$ , si tuvieramos  $\alpha(n - 2) \geq \gamma n$ , podríamos concluir que  $g(n) \geq \gamma n$ . Calculemos, usando una cadena de si-y-sólo-si:

$$\begin{aligned}\alpha(n - 2) &\geq \gamma n \\ \alpha n - 2\alpha &\geq \left(\alpha - 2\frac{\alpha}{n_2}\right)n \\ \alpha n - 2\alpha &\geq \alpha n - 2\alpha \frac{n}{n_2} \\ -2\alpha &\geq -2\alpha \frac{n}{n_2} \\ \alpha &\leq \alpha \frac{n}{n_2} \\ 1 &\leq \frac{n}{n_2} \\ n_2 &\leq n\end{aligned}$$

Y esto es cierto, porque estamos diciendo que vale nuestra proposición para todo  $n \geq n_2$ .

Luego, como para todo  $n \geq n_0$ , existen  $\gamma = \alpha - 2\frac{\alpha}{n_2}$ , y  $\beta = \delta$ , tal que para todo  $n \geq n_2$ ,  $\gamma n \leq g(n) \leq \delta n$ , tenemos que  $g(n) \in \Theta(n)$ .

Ahora podemos usar el teorema maestro para concluir que  $U \in O(n \log n)$ , usando  $a = 2$ ,  $b = 2$ ,  $k = 0$ ,  $f = g$ . Esto todavía no nos dice que  $T \in O(n \log n)$ .

Como  $U \in O(n \log n)$ , sabemos que existe  $n_3 \in \mathbb{N}$ ,  $\varphi > 0 \in \mathbb{R}$  tal que para todo  $n \geq n_3$ , tenemos  $U(n) \leq \varphi n \log n$ . Como  $T(n) = U(n + 2)$ , y  $n + 2 > n \geq n_3$ , entonces para esos mismos  $n$  tenemos  $U(n + 2) \leq \varphi(n + 2) \log(n + 2)$ , y por lo tanto  $T(n) \leq \varphi(n + 2) \log(n + 2)$ . Si tomamos  $n_4 = \max(n_3, 2)$ , sabemos que para todo  $n \geq n_4$ ,  $n \geq 2$ , y  $T(n) \leq \varphi(n + 2) \log(n + 2)$ , y como  $n \geq 2$ , tenemos  $n + 2 \leq 2n$ . Por lo tanto,  $T(n) \leq 2\varphi n \log(2n)$ . Como  $n \geq 2$ ,  $n^2 \geq 2n$ . Por lo tanto, usando que  $\log$  es monotónicamente creciente,  $T(n) \leq 2\varphi n \log(2n) \leq 2\varphi n \log(n^2) = 4\varphi n \log n$ . Finalmente, si definimos  $\varepsilon = 4\varphi$ , tenemos que para todo  $n \geq n_4$ ,  $T(n) \leq \varepsilon n \log n$ , lo que nos deja concluir que  $T \in O(n \log n)$ .  $\square$

### Ejercicio 6.2.4

Se tiene el siguiente algoritmo recursivo:

```
def f(cuts: list[int], i: int, j: int) -> int:
    if j == i + 1:
        return 0
    r = inf
    for k in range(i, j):
        r = min(r, f(cuts, i, k) + f(cuts, k, j))
    return r + cuts[j] - cuts[i]
```

Dar una cota asintótica superior ajustada para el número de operaciones que realice este algoritmo, en función del tamaño de entrada  $j - i$ .



*Demostración.* Definimos el tamaño de una entrada  $(\text{cuts}, i, j)$  como  $j - i$ . La función que describe el tiempo que toma  $f$  en correr, en términos del tamaño de entrada, es  $T(n) = O(1) + \sum_{k=1}^{n-1} T(k) + T(n-k)$ .

Obviamente no podemos usar el teorema maestro para esto, porque  $T$  no tiene la forma correcta, así que vamos a tener que analizar cuidadosamente qué está pasando. Jugando un poco, vemos que:

$$\begin{aligned} T(n) &= O(1) + T(1) + T(n-1) + T(2) + T(n-2) + \dots + T(n-1) + T(1) \\ &= O(1) + 2 \sum_{i=1}^{n-1} T(i) \end{aligned}$$

Esto nos puede recordar a la fórmula para las potencias de un número. Por ejemplo,  $2^n = 1 + \sum_{i=0}^{n-1} 2^i$ , o  $3^n = 1 + 2 \sum_{i=0}^{n-1} 3^i$ . Como esto se parece bastante a la serie de potencias de tres, vamos a intentar adivinar que  $T(n) \leq \alpha 3^n$  para todo  $n \in \mathbb{N}$ , y algún  $\alpha > 0 \in \mathbb{R}$ . Esto nos diría que  $T \in O(3^n)$ .

Probemos esto por inducción. Por definición de « $O(1)$ », sabemos que existe una función  $h : \mathbb{N} \rightarrow \mathbb{N}$ , tal que  $T(n) = h(n) + 2 \sum_{i=1}^{n-1} T(i)$ , y existen  $n_0 \in \mathbb{N}$ ,  $\beta > 0$  tales que para todo  $n \geq n_0$ ,  $h(n) \leq \beta$ . Sea  $r = \max(\beta, \max_{i=0}^{n_0} h(i))$ . Entonces tenemos que para todo  $n$ ,  $h(n) \leq r$ . La cota vale para los primeros  $n_0$  valores de  $n$  por la segunda rama del max, y vale para todos los valores después de  $n_0$  por la primera rama, que a su vez vale por la definición de  $h \in O(1)$ .

Luego, para todo  $n$ ,  $T(n) \leq r + 2 \sum_{i=1}^{n-1} T(i)$ . Si esto tiene que ser menor o igual a  $\alpha 3^n$ , veamos quién tiene que ser  $\alpha$ .

$$\begin{aligned} T(n) &\leq r + 2 \sum_{i=1}^{n-1} T(i) \\ &\leq r + 2 \sum_{i=1}^{n-1} \alpha 3^i \\ &\leq r + 2\alpha \left(\frac{1}{2}\right)(3^n - 3) \\ &\leq r + \alpha 3^n - 3\alpha \end{aligned}$$

Vamos a querer concluir que  $T(n) \leq \alpha 3^n$ . Luego, queremos que  $r - 3\alpha = 0$ , y luego  $\alpha = \frac{r}{3}$ . Verifiquemos que con ese  $\alpha$  se cumple lo que queremos:

$$T(n) \leq r + \alpha 3^n - 3\left(\frac{r}{3}\right) = \alpha 3^n$$

Esto parece funcionar. Probémoslo por inducción, entonces. Sea  $P(n) : n \geq 1 \Rightarrow T(n) \leq \alpha 3^n$

- $P(1) = h(1) \leq r = 3\alpha = 3^1\alpha$ .
- Asumimos que  $P(j)$  vale para todo  $j < n$ , queremos probar  $P(n)$ .  $T(n) \leq r + 2\sum_{i=1}^{n-1} T(i) \leq \alpha 3^n$  por el argumento de arriba, donde usamos la hipótesis inductiva para acotar cada  $T(i)$  por  $\alpha 3^i$ .

Luego vemos que tomando  $n_0 = 1$ ,  $\alpha = \max\left(\beta, \frac{\max_{i=0}^{n_0} h(i)}{3}\right)$ , tenemos que para todo  $n \geq n_0$ ,  $T(n) \leq \alpha 3^n$ , y por lo tanto  $T \in O(3^n)$ .  $\square$

### Ejercicio 6.2.5

Determinar cuánto tiempo va a tomar el siguiente algoritmo, en el caso promedio:

```
def f(n: int) -> int:
    if n <= 0: return 0
    i = random(n)
    return f(i) + f(n - 1 - i)
```

asumiendo que `random` toma una unidad de tiempo, y todas las otras operaciones toman tiempo despreciable. `random(n)` devuelve un número uniformemente al azar en  $[0, n]$ .



*Demostración.* Si  $T(n)$  es la función que nos dice el tiempo que va a tomar ejecutar `f(n)`, en el tiempo promedio, entonces tenemos que  $T(0) = 0$ . Para ver qué es  $T(n)$  en general, vemos que

$$\begin{aligned} T(n) &= \mathbb{E}_{i \sim U[0,n]} [T(i) + T(n-1-i) + 1] \\ &= \sum_{i=0}^{n-1} \frac{T(i) + T(n-1-i) + 1}{n} \\ &= \sum_{i=0}^{n-1} \frac{2T(i) + 1}{n}, \text{ reordenando los términos} \end{aligned}$$

Luego  $nT(n) = n + \sum_{i=0}^{n-1} 2T(i)$ , para todo  $n \in \mathbb{N}$ . Queremos una fórmula más simple para  $T(n)$ . Jugando un poco con las ecuaciones, podemos usar que  $(n-1)T(n-1) = n-1 + \sum_{i=0}^{n-2} 2T(i)$ , y restar estos dos para hacer desaparecer la sumatoria. Obtenemos  $nT(n) - (n-1)T(n-1) = 2T(n-1) + 1$ . Esto nos da una relación simple entre  $T(n)$  y  $T(n-1)$ , que nos da la idea de seguir usando esta relación una y otra vez a ver si podemos hacer desaparecer las  $T$  de un lado, quedándonos sólo con  $T(n)$  en el otro.

Reescribiendo,  $nT(n) = (n+1)T(n-1) + 1$ , o también,  $\frac{T(n)}{n+1} = \frac{T(n-1)}{n} + \frac{1}{n(n+1)}$ . Esto finalmente es simétrico (los dos términos con  $T$  son similares, siendo de la forma  $\frac{T(x)}{x+1}$ ), y podemos repetir esto hasta llegar a  $n = 0$ :

$$\begin{aligned}
\frac{T(n)}{n+1} &= \frac{T(n-1)}{n} + \frac{1}{n(n+1)} \\
&= \frac{T(n-2)}{n-1} + \frac{1}{(n-1)n} + \frac{1}{n(n+1)} \\
&= \frac{T(n-3)}{n-2} + \frac{1}{(n-2)(n-1)} + \frac{1}{(n-1)n} + \frac{1}{n(n+1)} \\
&= \dots \\
&= \frac{T(0)}{1} + \sum_{i=1}^n \frac{1}{i(i+1)} \\
&= T(0) + \sum_{i=1}^n \frac{1}{i(i+1)}
\end{aligned}$$

que, como  $T(0) = 0$ , nos dice que  $\frac{T(n)}{n+1} = \sum_{i=1}^n \frac{1}{i(i+1)}$ .

Ahora queremos ver quién es  $T(n) = (n+1) \sum_{i=1}^n \frac{1}{i(i+1)}$ . Una estrategia que podemos usar es simplemente ver qué valores tiene  $T$  para los primeros números, a mano.

1.  $T(1) = 2(\frac{1}{2}) = 1$
2.  $T(2) = 3(\frac{1}{2} + \frac{1}{6}) = 3(\frac{3}{6} + \frac{1}{6}) = 3(\frac{4}{6}) = 2$
3.  $T(3) = 4(\frac{1}{2} + \frac{1}{6} + \frac{1}{12}) = 4(\frac{6}{12} + \frac{2}{12} + \frac{1}{12}) = 4(\frac{9}{12}) = 3$
4. Para el siguiente, intentemos hacernos la vida un poco más fácil:

$$\begin{aligned}
T(4) &= 5\left(\frac{1}{2} + \frac{1}{6} + \frac{1}{12} + \frac{1}{20}\right) \\
&= (4+1)\left(\left(\frac{1}{2} + \frac{1}{6} + \frac{1}{12}\right) + \frac{1}{20}\right) \\
&= 4\left(\frac{1}{2} + \frac{1}{6} + \frac{1}{6}\right) + \frac{4}{20} + \left(\frac{1}{2} + \frac{1}{6} + \frac{1}{12}\right) + \frac{1}{20} \\
&= T(3) + \left(\frac{4}{20} + \left(\frac{1}{2} + \frac{1}{6} + \frac{1}{12}\right) + \frac{1}{20}\right) \\
&= T(3) + \left(\frac{4}{20} + \frac{T(3)}{4} + \frac{1}{20}\right) \\
&= T(3) + \left(\frac{T(3)}{4} + \frac{1}{4}\right) \\
&= 4
\end{aligned}$$

El hacer esto a mano nos dio una pista sobre qué va a pasar en el caso general. No sólo vemos que  $T(n) = n$ , sino que para demostrarlo vamos a poder usar inducción, porque apareció  $T(3)$  al calcular  $T(4)$ .

Definimos, entonces,  $P(n) : T(n) = n$ .

1.  $P(0) : T(0) = 0$ . Esto es cierto porque  $T(0) = 1 \times 0 = 0$ , donde 0 es el valor de una suma de cero términos.
2.  $P(n) \Rightarrow P(n+1)$ . Hagamos lo mismo que hicimos para  $T(4)$ , pero en general:

$$\begin{aligned}
T(n+1) &= (n+2) \sum_{i=1}^{n+1} \frac{1}{i(i+1)} \\
&= ((n+1)+1) \left( \frac{1}{(n+1)(n+2)} + \sum_{i=1}^n \frac{1}{i(i+1)} \right) \\
&= \left( (n+1) \sum_{i=1}^n \frac{1}{i(i+1)} \right) + \frac{n+1}{(n+1)(n+2)} + \frac{1}{(n+1)(n+2)} + \sum_{i=1}^n \frac{1}{i(i+1)} \\
&= T(n) + \frac{1}{n+1} + \frac{T(n)}{n+1} \\
&= n + \frac{1}{n+1} + \frac{n}{n+1} = n+1
\end{aligned}$$

Luego vale  $P(n)$  para todo  $n \in \mathbb{N}$ , y luego  $T(n) = n$  para todo  $n \in \mathbb{N}$ . □

### Ejercicio 6.2.6

Se tienen tres algoritmos que resuelven el mismo problema. Todos realizan algún número positivo de operaciones cuando su entrada tiene tamaño cero.

1. El primero divide un problema de tamaño  $n$  en 5 subproblemas de tamaño  $\frac{n}{2}$  cada uno, y combina sus soluciones en  $O(n)$  operaciones.
2. El segundo divide un problema de tamaño  $n$  en 2 subproblemas de tamaño  $n - 1$  cada uno, y combina sus soluciones en  $O(1)$  operaciones.
3. El tercero divide un problema de tamaño  $n$  en 9 subproblemas de tamaño  $\frac{n}{3}$  cada uno, y combina sus soluciones en  $\Theta(n^2)$  operaciones.

Si nuestro  $n$  es enorme, y queremos minimizar el número de operaciones que requiere encontrar una solución, podemos determinar cuál algoritmo nos conviene usar sólo sabiendo esto?

Para esta demostración les voy a escribir el razonamiento que hago mientras escribo la demostración.

El primer caso es fácil,  $\log_2(5) > 1$ , y el costo de combinar soluciones es pequeño porque  $1 < \log_2(5)$ , entonces el costo está dominado por las hojas del árbol de recursión, y uso el teorema maestro para saber que esto es  $\Theta(n^{\log_2(5)})$ .

Para el segundo esto me recuerda a la función  $2^n = 2 \times 2^{n-1}$ , o también a la sucesión de Fibonacci,  $F_n = F_{n-1} + F_{n-2}$ . Esas tienen solución exponencial, entonces supongamos que  $T(n) \leq \gamma c^n$  para algún  $\gamma, c \in \mathbb{R}_{>0}$ .  $T(0)$  es alguna constante, pongámosle  $T(0) = \alpha$ . Entonces si quiero que  $T(0) \leq \gamma c^0 = \gamma$ , voy a tener  $\alpha \leq \gamma$ , y luego mi  $\gamma$  tiene que cumplir  $\gamma \geq \alpha$ . Como hay un  $O(1)$  en la definición de  $T$ , sea  $h : \mathbb{N} \rightarrow \mathbb{N}$  tal que  $T(n) = 2T(n-1) + h(n)$ , existen  $n_0 \in \mathbb{N}, \beta \in \mathbb{R}_{>0}$  tal que para todo  $n \geq n_0, h(n) \leq \beta$ . Veamos qué puedo deducir sobre  $\gamma$  usando  $\beta$ ...

$$\begin{aligned}
T(n) &= 2T(n-1) + h(n) \\
&\leq 2T(n-1) + \beta \\
&\leq 2\gamma c^{n-1} + \beta
\end{aligned}$$

Entonces, si quiero probar que  $T(n) \leq \gamma c^n$ , tengo que probar que  $2\gamma c^{n-1} + \beta \leq \gamma c^n$ , y por transitividad de  $\leq$  está. Voy a encontrar  $\gamma$  y  $c$  que hagan cierto esto. Moralmente espero que  $c \approx 2$  porque esa es la solución cuando  $\alpha = 1, \beta = 0$  (y  $T(n) = 2^n$  para todo  $n$  ahí).

$$2\gamma c^{n-1} + \beta \leq \gamma c^n$$

A ver qué pasa si pongo  $\gamma = \beta$ ...

$$\begin{aligned} 2\beta c^{n-1} + \beta &\leq \beta c^n \\ 2c^{n-1} + 1 &\leq c^n \\ 1 &\leq c^{n-1}(c - 2) \end{aligned}$$

pero eso no va a ser cierto si  $c = 2$ . Todavía no sé que  $c = 2$ , pero puede ser.

Expandamos un poco, asumiendo  $n$  grande (bastante más grande que  $n_0$  para poder usar  $\beta$ )...

$$T(n) \leq 2T(n-1) + \beta$$

$$\begin{aligned} &\leq 2(2T(n-2) + \beta) + \beta \\ &= 4T(n-2) + 2\beta + \beta \\ &\leq 4(2T(n-3) + \beta) + 2\beta + \beta \\ &= 8T(n-3) + 4\beta + 2\beta + \beta \\ &\dots, \end{aligned}$$

lo siguiente es falso porque la cota de  $h(n) \leq \beta$  sólo vale para  $n > n_0$ , pero estoy siendo informal...

$$\begin{aligned} &\leq 2^n T(0) + \sum_{i=0}^{n-1} 2^i \beta \\ &= 2^n \alpha + \beta(2^n - 1) \\ &= 2^n(\alpha + \beta) - \beta \end{aligned}$$

OK, eso me facilita las cosas, puedo usar  $\gamma = \alpha + \beta$ , y la forma que va a tener la cota es  $\gamma 2^n - \beta$ , no sólo  $\gamma 2^n$ . Después lo formalizo.

Eso igual sólo me da una cota superior. Tendría que también argumentar una cota inferior para saber bien quién es ese  $T$  y poder compararlo.

Para el tercero, tengo  $T(n) = 9T(\frac{n}{3}) + \Theta(n^2)$ . Como  $\log_3(9) = 2$ , no tenemos que  $2 < 2$ , entonces tenemos que usar el segundo caso del teorema maestro, que con  $a = 3, b = 3, k = 0$ , nos dice que  $T \in \Theta(n^2 \log n)$ .

*Demostración.* Calculemos el comportamiento asintótico de cada uno. Vamos a llamar  $T : \mathbb{N} \rightarrow \mathbb{N}$  a la función que, dado un  $n \in \mathbb{N}$ , nos dice cuántas operaciones toma cada algoritmo para resolver un problema de tamaño  $n$ . Como cada algoritmo está descrito usando  $O(\dots)$ , no vamos a poder encontrar quiénes son  $T$  exactamente, sino que vamos a conocer su comportamiento asintótico. Como nos dicen que  $n$  es enorme, esto es realmente lo único que importa, porque los factores constantes no van a importar si  $n$  es enorme.

1. El primer algoritmo cumple que  $T(n) = 5T(\frac{n}{2}) + O(n)$ . Usando el teorema maestro con  $a = 5, b = 2, f \in O(n)$ , tenemos que  $c = \log_2(5) > \log_2(4) = 2 > 1$ , donde 1 es el

exponente polinomial de  $f$ , por estar en  $O(n^1)$ . Como  $c > 1$ , entonces caemos en el primer caso del teorema maestro, el cual nos dice que  $T \in \Theta(n^{\log_2(5)})$ .

2. Como  $T(n)$  está definido en términos de  $T(n - 1)$ , vamos a usar inducción para ver quién es  $T$ . La recurrencia que cumple  $T$  es  $T(n) = 2T(n - 1) + O(1)$ . Sea  $\alpha = T(0)$ . Como  $T(n) = 2T(n - 1) + O(1)$ , ese  $O(1)$  nos dice que existe una función  $h : \mathbb{N} \rightarrow \mathbb{N}$ ,  $h \in O(1)$ , tal que  $T(n) = 2T(n - 1) + h(n)$  para todo  $n \geq 0$ . Como  $h \in O(1)$ , sean  $n_0 \in \mathbb{N}$ ,  $\beta > 0 \in \mathbb{R}$ , tales que para todo  $n \geq n_0$ ,  $h(n) \leq \beta$ . Sea  $\delta = \max\left(\beta, \max_{0 \leq i \leq n_0} h(i)\right)$ , y por lo tanto  $h(k) \leq \delta$  para todo  $k \in \mathbb{N}$ . Sea  $\gamma = \alpha + \delta$ .

$$P(n) : \alpha 2^n \leq T(n) \leq \gamma 2^n - \delta.$$

1. Caso base,  $P(0)$ . Por definición de  $\alpha$ ,  $T(0) = \alpha = \alpha + \delta - \delta = \gamma - \delta = \gamma 2^0 - \delta$ , que junto con  $\alpha 2^0 = \alpha$ , prueba  $P(0)$ .
2. Paso inductivo. Asumo  $P(n)$ , queremos probar  $P(n + 1)$ . Sabemos que  $T(n + 1) = 2T(n) + h(n)$ .

$$\begin{aligned} T(n + 1) &= 2T(n) + h(n) \\ &\leq 2T(n) + \delta \\ &\leq 2(\gamma 2^n - \delta) + \delta \\ &= \gamma 2^{n+1} - 2\delta + \delta \\ &= \gamma 2^{n+1} - \delta \end{aligned}$$

Asimismo,

$$\begin{aligned} T(n + 1) &= 2T(n) + h(n) \\ &\geq 2T(n) \\ &\geq 2(\alpha 2^n) \\ &\geq \alpha 2^{n+1} \end{aligned}$$

y concluimos que  $\alpha 2^{n+1} \leq T(n + 1) \leq \gamma 2^{n+1} - \delta$ , que es lo que queríamos demostrar,  $P(n + 1)$ .

Luego, como  $\alpha > 0$  porque nos lo dice el enunciado, y sabemos que  $T(n) \leq \gamma 2^n$  (descartando el término  $-\delta$  por transitividad), tenemos que  $T \in O(2^n)$  y  $T \in \Omega(2^n)$ , con lo cual  $T \in \Theta(2^n)$ .

3. Para el tercer caso, usamos el segundo caso del teorema maestro, con  $a = 3$ ,  $b = 3$ ,  $k = 0$ , y concluimos que  $T \in \Theta(n^2 \log n)$ .

Sabemos entonces que el número de operaciones que estos tres algoritmos hacen, ante una entrada de tamaño  $n$ , está respectivamente en  $\Theta(n^{\log_2(5)})$ ,  $\Theta(2^n)$ , y  $\Theta(n^2 \log n)$ .

Veamos cuál nos conviene usar. Llamemos a estas tres funciones  $T_1$ ,  $T_2$ , y  $T_3$ . Recordando que  $g \in O(f) \Leftrightarrow \lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \infty$ , y usando las reglas usuales de orden de conjuntos asintóticos:

- Como  $T_2 \in \Omega(2^n)$ , entonces llamando  $f(n) = 2^n$ , tenemos que  $f \in O(T_2)$ . Luego, como  $T_1 \in O(n^{\log_2(5)})$ , y  $O(n^{\log_2(5)}) \subset o(f)$ , tenemos que  $T_1 \in o(f)$ . Como  $T_1 \in o(f)$  y  $f \in O(T_2)$ , tenemos que  $T_1 \in o(T_2)$ .

- Como  $\lim_{n \rightarrow \infty} \frac{n^{\log_2(5)}}{n \log n} = \lim_{n \rightarrow \infty} \frac{n^{\log_2(5)-1}}{\log n} = \lim_{n \rightarrow \infty} \frac{n^\varepsilon}{\log n} = \lim_{n \rightarrow \infty} n\varepsilon n^{\varepsilon-1} = \varepsilon n^\varepsilon = \infty$ , porque  $\log_2(5) - 1 = \varepsilon > 0$ , y usamos la regla de L'Hopital, tenemos que  $T_3 \in o(T_1)$ .
- Como  $T_3 \in o(T_1)$  y  $T_1 \in o(T_2)$ , tenemos que  $T_3 \in o(T_2)$ .

Luego, como  $T_3$  está asintóticamente dominada por las otras dos, y  $n$  es enorme, nos conviene usar el tercer algoritmo.  $\square$

### 6.3 Sucesiones y series

#### Ejercicio 6.3.1

Sea  $n \in \mathbb{N}$ . Demostrar que  $\sum_{i=0}^n (2i + 1) = (n + 1)^2$ .



*Demostración.* Vamos a probar esto por inducción,  $P(n) : \sum_{i=0}^n (2i + 1) = (n + 1)^2$ .

- $P(0)$ . El lado izquierdo de la ecuación es  $\sum_{i=0}^0 (2i + 1) = 2 \times 0 + 1 = 1$ . El lado derecho es  $(0 + 1)^2 = 1$ . Luego, vale  $P(0)$ .
- Sea  $n \in \mathbb{N}$ . Queremos ver que  $P(n) \Rightarrow P(n + 1)$ . Luego, vamos a asumir  $P(n)$ , y vamos a ver  $P(n + 1)$ . Ver  $P(n + 1)$  es que  $\sum_{i=0}^{n+1} (2i + 1) = (n + 2)^2$ . Razonemos entonces:

$$\begin{aligned} \sum_{i=0}^{n+1} (2i + 1) &= 2(n + 1) + 1 + \sum_{i=0}^n (2i + 1) \\ &= 2(n + 1) + 1 + (n + 1)^2, \text{ usando } P(n) \\ &= 2n + 2 + 1 + (n + 1)^2 \\ &= 1^2 + 2 \times (n + 1) \times 1 + (n + 1)^2 \\ &= ((n + 1) + 1)^2 \\ &= (n + 2)^2 \end{aligned}$$

que es lo que queríamos demostrar.

Luego, vale  $P(n)$  para todo  $n \in \mathbb{N}$ .  $\square$

#### Ejercicio 6.3.2

Demostrar que para todo  $n \in \mathbb{N}, n \geq 5$ , tenemos  $2^n > n^2$ .



*Demostración.* Vamos a demostrar la propiedad  $P(n) : n \geq 5 \Rightarrow 2^n > n^2$ , por inducción. Para  $n < 5$ , la demostración es trivial, puesto que no hay nada que probar («falso implica todo»).

Luego vamos a empezar nuestra inducción con  $n = 5$  como caso base.

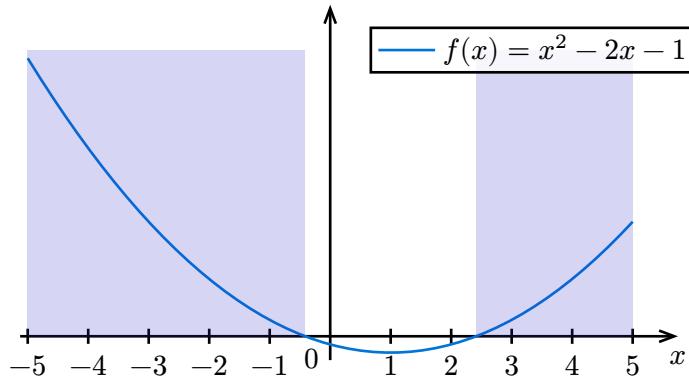
- $P(5)$ : Vemos que  $2^5 = 32$ , mientras que  $5^2 = 25$ , con lo cual efectivamente  $2^5 > 5^2$ , que es  $P(5)$ .
- Sea  $n \in \mathbb{N}$ .  $P(n) \Rightarrow P(n + 1)$ : Si  $n \leq 4$ , vimos que  $P(n + 1)$ , sea porque la premisa de  $P(n + 1)$  es falsa cuando  $n < 3$ , o porque  $P(5)$  es cierto porque lo probamos arriba, cuando  $n = 4$ . Luego asumimos  $n \geq 5$ . Por un lado, tenemos  $2^{n+1} = 2(2^n) > 2n^2$ , donde usamos  $P(n)$  y  $n \geq 5$  para obtener la desigualdad, mediante la conclusión de  $P(n)$ . Por el otro,

$(n+1)^2 = n^2 + 2n + 1$ . Si probamos que  $2n^2 > n^2 + 2n + 1$ , tendremos que  $2^{n+1} > (n+1)^2$ .

$$\begin{aligned} 2n^2 &> n^2 + 2n + 1 \\ n^2 - 2n - 1 &> 0 \end{aligned}$$

Consideremos ahora el polinomio  $f(x) = x^2 - 2x - 1$ . Lo podemos factorizar como  $f(x) = (x - (1 + \sqrt{2}))(x - (1 - \sqrt{2}))$ , es decir que sus raíces son  $a = 1 - \sqrt{2}$ , y  $b = 1 + \sqrt{2}$ .

Como el coeficiente principal de  $f(x)$  es 1,  $f$  es positiva en  $(-\infty, a)$ , es negativa en  $(a, b)$ , y es positiva en  $(b, \infty)$ .



Como  $n \geq 5 > b = 1 + \sqrt{5}$ , tendremos que  $f(n) > 0$ , y luego  $n^2 - 2n - 1 > 0$ , es decir,  $2n^2 > n^2 + 2n + 1$ . Como dijimos, esto implica que  $2^{n+1} > (n+1)^2$ , lo cual prueba  $P(n+1)$ .

□

### Ejercicio 6.3.3

Sea  $(a_n)_{n \in \mathbb{N}}$  la sucesión dada por  $a_0 = 1$ , y para todo  $n \in \mathbb{N}$ ,  $n > 0$ ,  $a_n = 2(n-1)a_{n-1} + 2^n(n-1)!$ .

Demostrar que para todo  $n \in \mathbb{N}$ ,  $n > 0$ ,  $a_n = 2^n n!$ .

◆

*Demostración.* Como toda sucesión definida recursivamente, tenemos una estructura inductiva. Luego, intentemos probar esto por inducción. La proposición que vamos a probar es  $P(n) : a_n = 2^n n!$ .

1.  $P(0)$ : Si  $n = 0$ , entonces queremos ver  $P(0)$ , es decir que  $a_0 = 2^0 1! = 1$ . Esto es cierto, porque por definición  $a_0 = 1$ .
2.  $P(n) \Rightarrow P(n+1)$ . Queremos probar que  $a_{n+1} = 2^{n+1}(n+1)!$ . Sabemos que  $a_{n+1} = 2na_n + 2^{n+1}n!$ . Como vale  $P(n)$ , podemos reemplazar  $a_n$  por  $2^n n!$ . Luego, sabemos que  $a_{n+1} = 2n(2^n n!) + 2^{n+1}n! = 2^{n+1}(nn! + n!) = 2^{n+1}n!(n+1) = 2^{n+1}(n+1)!$ , que es lo que queríamos demostrar.

□

### Ejercicio 6.3.4 (Series geométricas)

Sean  $n \in \mathbb{N}$ , y  $a \in \mathbb{C}, a \neq 1$ . Probar que

$$\sum_{i=0}^n a^i = \frac{a^{n+1} - 1}{a - 1}$$

*Demostración.* Vamos a probar esto mediante una serie de manipulaciones a esa ecuación, teniendo cuidado de que cada una sea un si-y-sólo-si.

$$\begin{aligned}\sum_{i=0}^n a^i &= \frac{a^{n+1} - 1}{a - 1} \\ (a - 1) \sum_{i=0}^n a^i &= a^{n+1} - 1 \\ a \left( \sum_{i=0}^n a^i \right) - \sum_{i=0}^n a^i &= a^{n+1} - 1 \\ 1 + \sum_{i=1}^{n+1} a^i - \sum_{i=0}^n a^i &= a^{n+1} \\ \sum_{i=0}^0 a^i + \sum_{i=1}^{n+1} a^i - \sum_{i=0}^n a^i &= a^{n+1} \\ \sum_{i=0}^{n+1} a^i - \sum_{i=0}^n a^i &= a^{n+1} \\ \sum_{i=n+1}^{n+1} a^i + \sum_{i=0}^n a^i - \sum_{i=0}^n a^i &= a^{n+1} \\ a^{n+1} &= a^{n+1}\end{aligned}$$

Como llegamos a algo cierto a través de manipulaciones reversibles, cada paso puede ser revertido para empezar con  $a^{n+1} = a^{n+1}$  y concluir con  $\sum_{i=0}^n a^i = \frac{a^{n+1} - 1}{a - 1}$ .

Notar que para revertir la multiplicación por  $a - 1$  que hacemos para ir de la primer ecuación a la segunda, estamos usando que  $a \neq 1$ .  $\square$

### Ejercicio 6.3.5

Sea la sucesión  $(a_n)_{n \in \mathbb{N}}$  definida como  $a_0 = 1$ , y para todo  $n \in \mathbb{N}, n > 0$ ,  $a_n = 4a_{n-1} - 2\frac{(2n-2)!}{n!(n-1)!}$ . Probar que para todo  $n \in \mathbb{N}$ ,  $a_n = \binom{2n}{n}$ .

*Demostración.* Nuevamente, vamos a usar inducción, porque  $(a_n)$  tiene una estructura inductiva (es decir, está definida recursivamente).

Vamos a probar la proposición  $P(n) : a_n = \binom{2n}{n}$ , para todo  $n \in \mathbb{N}$ . Recordemos que  $\binom{a}{b} = \frac{a!}{b!(a-b)!}$ , para todo  $a, b \in \mathbb{N}$ .

Sea  $n \in \mathbb{N}$ .

1. Si  $n = 0$ , entonces queremos ver  $P(0)$ , que es  $a_0 = \binom{0}{0} = \frac{0!}{0!0!} = \frac{1}{1} = 1$ , y esto es cierto pues definimos  $a_0 = 1$ .
2. Si  $n > 0$ , entonces podemos usar la hipótesis inductiva en  $n - 1$ , y sabemos  $P(n - 1)$ . Esto nos dice que  $a_{n-1} = \binom{2(n-1)}{n-1}$ . Queremos ver  $P(n + 1)$ , es decir, que  $a_n = \binom{2n}{n}$ . Como sabemos por definición que  $a_n = 4a_{n-1} - 2\frac{(2n-2)!}{n!(n-1)!}$ , podemos reemplazar lo que sabemos es  $a_{n-1}$  acá, y sabemos luego que  $a_n = 4\binom{2(n-1)}{n-1} - 2\frac{(2n-2)!}{n!(n-1)!}$ .

$$\begin{aligned}
 a_n &= 4\binom{2n-2}{n-1} - 2\frac{(2n-2)!}{n!(n-1)!} \\
 &= 4\frac{(2n-2)!}{(n-1)!(n-1)!} - \frac{(2n-2)!}{n!(n-1)!} \\
 &= \frac{(2n-2)!}{(n-1)!} \left( \frac{4}{(n-1)!} - \frac{2}{n!} \right) \\
 &= \frac{(2n-2)!}{(n-1)!} \left( \frac{4n!}{n!(n-1)!} - \frac{2(n-1)!}{n!(n-1)!} \right) \\
 &= \frac{(2n-2)!}{(n-1)!} \frac{4n(n-1)! - 2(n-1)!}{n!(n-1)!} \\
 &= \frac{(2n-2)!}{(n-1)!} \frac{4n-2}{n!} \\
 &= 2\frac{(2n-2)!}{(n-1)!} \frac{2n-1}{n!} \\
 &= 2\frac{(2n-1)!}{n!(n-1)!} \\
 &= 2\binom{2n-1}{n} \\
 &= \binom{2n-1}{n} + \binom{2n-1}{n} \\
 &= \binom{2n-1}{n} + \binom{2n-1}{2n-1-n}, \text{ pues } \binom{a}{b} = \binom{a}{a-b} \\
 &= \binom{2n-1}{n} + \binom{2n-1}{n-1} \\
 &= \binom{2n}{n}, \text{ pues } \binom{a}{b} + \binom{a}{b-1} = \binom{a+1}{b}
 \end{aligned}$$

que es lo que queríamos demostrar. □

### Ejercicio 6.3.6

Estamos pensando en una estrategia de inversión. Tenemos un fondo que crece, en valor esperado,  $1 - \alpha$  cada mes. Por ejemplo, si  $\alpha = 1.1$ , nuestro fondo crece en valor esperado un 10% por mes. Asumiendo que empezamos con  $c$  pesos en nuestra cuenta, y cada mes compramos  $b$  pesos más de este fondo, ¿cuál es el valor de los fondos que esperamos tener luego de  $k$  meses? ¿Qué necesitan asumir para encontrar este valor?

El objetivo de este ejercicio no es que hagan cuentas, sino que entiendan cómo descubrir qué están asumiendo cuando hacen cuentas.

*Demostración.* Sea  $a_i$  el retorno del  $i$ -ésimo mes. Sabemos que  $\mathbb{E}[a_i] = \alpha$ . Podemos definir la siguiente sucesión:

$$\begin{aligned} T(0) &= c \\ T(n) &= b + a_n T(n-1) \end{aligned}$$

Queremos encontrar  $\mathbb{E}[T(n)]$ . Veamos cómo se comporta esta función.  $\mathbb{E}[T(1)] = \mathbb{E}[b + a_1 c] = b + \alpha c$ , usando linearidad de esperanza, y que  $b$  y  $c$  son constantes. Para su segundo valor,  $\mathbb{E}[T(2)] = \mathbb{E}[b + a_2 T(1)] = b + \mathbb{E}[a_2 T(1)]$ . Nos gustaría decir que esto es  $b + \mathbb{E}[a_2] \mathbb{E}[T(1)]$ , pero acá debemos asumir que  $a_2$  y  $T(1)$  son independientes. **Para esto tenemos que asumir que los retornos  $\{a_i\}$  son independientes de a pares.** Es decir,  $\mathbb{E}[a_i a_j] = \mathbb{E}[a_i] \mathbb{E}[a_j] = \alpha^2$  para todo  $i, j$ . Luego, tenemos

$$\begin{aligned} \mathbb{E}[T(2)] &= b + \mathbb{E}[a_2(b + a_1 c)] \\ &= b + \mathbb{E}[a_2 b + a_2 a_1 c] \\ &= b + b \mathbb{E}[a_2] + c \mathbb{E}[a_2 a_1] \\ &= b + b\alpha + c\alpha^2 \end{aligned}$$

Veamos qué pasa para  $n = 3$ .

$$\begin{aligned} \mathbb{E}[T(3)] &= \mathbb{E}[b + a_3 T(2)] \\ &= b + \mathbb{E}[a_3 T(2)] \end{aligned}$$

Acá no podemos decir que  $\mathbb{E}[a_3 T(2)] = \mathbb{E}[a_3] \mathbb{E}[T(2)]$  porque no sabemos si  $a_3$  y  $T(2)$  están correlacionados. Veamos qué pasa si expandimos  $T(2)$  por definición:

$$\begin{aligned} \mathbb{E}[T(3)] &= b + \mathbb{E}[a_3 T(2)] \\ &= b + \mathbb{E}[a_3(b + a_2 T(1))] \\ &= b + \mathbb{E}[a_3 b + a_3 a_2 T(1)] \\ &= b + b\alpha + \mathbb{E}[a_3 a_2 T(1)] \\ &= b + b\alpha + \mathbb{E}[a_3 a_2(b + a_1 c)] \\ &= b + b\alpha + \mathbb{E}[a_3 a_2 b + a_3 a_2 a_1 c] \\ &= b + b\alpha + b\alpha^2 + c \mathbb{E}[a_3 a_2 a_1] \end{aligned}$$

El saber que los  $a_i$  no están correlacionados de a pares no nos deja decir que no estén correlacionados de a triples. Luego, para seguir acá vamos a necesitar asumir **que los  $\{a_i\}$  son independientes de a triples**, así como también lo que habíamos asumido antes, que son independientes de a pares.

Si asumimos eso, concluimos  $\mathbb{E}[T(3)] = b + b\alpha + b\alpha^2 + c\alpha^3$ .

Conjeturamos, entonces, que  $\mathbb{E}[T(n)] = \alpha^n c + \sum_{i=0}^{n-1} \alpha^i b$ , si asumimos que **los  $\{a_i\}$  son independientes tomados de a conjuntos de tamaño menor o igual a  $n$** .

Tenemos que tener cuidado al hacer inducción, entonces. Si quisieramos probar la proposición  $Q(n) : \mathbb{E}[T(n)] = \alpha^n c + \sum_{i=0}^{n-1} \alpha^i b$ , en algún momento vamos a tener  $\mathbb{E}[a_n T(n)]$ , y no vamos

a poder llegar a mucho si sólo tenemos eso, porque no sabemos que  $a_n$  y  $T(n)$  son independientes. Entonces, vamos a tener que ser precisos, y reescribir  $T$  de manera que podamos usar la independencia de los  $\{a_i\}$ .

Expresemos, entonces,  $T(n)$ , de una manera que nos deje usar esta independencia. Vimos que  $T(3) = b + ba_3 + ba_3a_2 + ca_3a_2a_1$ . Luego, definamos  $S_{i,n} = \prod_{i < j \leq n} a_j$ . Vemos que  $T(3) = bS_{3,3} + bS_{2,3} + bS_{1,3} + cS_{0,3}$ .

Probemos, entonces,  $P(n) : T(n) = cS_{0,n} + \sum_{i=1}^n bS_{i,n}$ .

1.  $P(0)$ .  $T(0) = c$  por definición.  $S_{0,0} = \prod_{0 < j \leq 0} a_j = 1$ , por definición de una productoria vacía. Asimismo,  $\sum_{i=1}^0 S_i b = 0$ , por definición de sumatoria vacía. Luego,  $T(0) = 1c + 0 = c$ , que prueba  $P(0)$ .
2.  $P(n) \Rightarrow P(n+1)$ .

$$\begin{aligned} T(n+1) &= b + a_n T(n) \\ &= b + a_n \left( cS_{0,n} + \sum_{i=1}^n bS_{i,n} \right), \text{ por hipótesis inductiva} \\ &= bS_{n+1,n+1} + ca_n S_{0,n} + a_n \sum_{i=1}^n bS_{i,n} \\ &= bS_{n+1,n+1} + cS_{0,n+1} + \sum_{i=1}^n bS_{i,n+1} \\ &= cS_{0,n+1} + \sum_{i=1}^{n+1} bS_{i,n+1} \end{aligned}$$

Luego vale  $P(n)$  para todo  $n \in \mathbb{N}$ . Veamos ahora qué nos dice esto sobre  $\mathbb{E}[T(n)]$ . Como los  $\{a_i\}$  son independientes en cualquier subconjunto de tamaño a lo sumo  $n$ , tenemos que  $\mathbb{E}[S_{i,n}] = \mathbb{E}\left[\prod_{i < j \leq n} a_j\right] = \prod_{i < j \leq n} \mathbb{E}[a_j] = \prod_{i < j \leq n} \alpha = \alpha^{n-i}$ .

Luego:

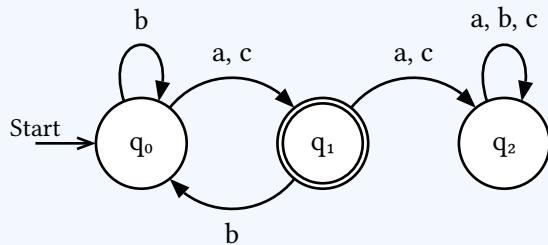
$$\begin{aligned} \mathbb{E}[T(n)] &= \mathbb{E}\left[cS_{0,n} + \sum_{i=1}^n bS_{i,n}\right], \text{ porque vale } P(n) \\ &= c\mathbb{E}[S_{0,n}] + \sum_{i=1}^n b\mathbb{E}[S_{i,n}] \\ &= c\alpha^n + \sum_{i=1}^n b\alpha^{n-i} \\ &= b\frac{\alpha^n - 1}{\alpha - 1} + c\alpha^n \end{aligned}$$

Para esto tuvimos que asumir que los  $\{a_i\}$  son independientes tomados de cualquier subconjunto de tamaño a lo sumo  $n$ .

Vemos con cuánto cuidado hay que deducir cosas, si no queremos cometer errores. □

### Ejercicio 6.3.7

Sea  $A$  el siguiente autómata finito:



Dado un  $n \in \mathbb{N}$ , cuántas cadenas de texto de longitud  $n$  acepta  $A$ ? ◆

El propósito de este ejercicio es que pasen un tiempo jugando con las ecuaciones, viendo qué pueden deducir. No imagino que hayan visto un ejercicio exactamente así. Sugiero que pasen un tiempo largo ( $>1$  hora) pensando y jugando antes de ver cómo lo hace esta demostración. Si no saben qué es un autómata finito, tomen  $T$  como definida en un enunciado.

#### Demostración.

Definiendo  $T(x, i)$  como el número de cadenas de texto que nos llevan al estado  $x$  luego de  $i$  caracteres, vemos que:

- $T(q, k) = 0$  para todo  $q$ , si  $k < 0$
- $T(q_0, 0) = 1$
- $T(q_0, i) = T(q_1, i - 1) + T(q_0, i - 1)$
- $T(q_1, i) = 2T(q_0, i - 1)$
- $T(q_2, i) = T(q_1, i - 1) + 3T(q_2, i - 1)$

Como el único estado que acepta es  $q_1$ , queremos encontrar  $T(q_1, n)$ . Para este  $A$  en particular, podemos notar que para todo  $i \geq 2$ ,  $T(q_0, i) = T(q_0, i - 1) + T(q_1, i - 1) = T(q_0, i - 1) + 2T(q_0, i - 2)$ , que se parece bastante a la sucesión de Fibonacci. Encontremos, entonces, una solución general para esta recurrencia. Llamemos por comodidad  $G(i) = T(q_0, i)$ .

Supongamos que existe una solución de la forma  $G(i) = c^i$  a la recurrencia  $G(i) = G(i - 1) + 2G(i - 2)$ . Entonces tendríamos que  $c^i = G(i) = G(i - 1) + 2G(i - 2) = c^{i-1} + 2c^{i-2}$ . Dividiendo ambos lados por  $c^{i-2}$ , nos queda  $c^2 = c + 2$ , que tiene como soluciones  $c = -1$  y  $c = 2$ .

Veamos entonces que ambas  $G(i) = 2^i$  y  $G(i) = (-1)^i$  son soluciones:

- $2^i = 2^{i-1} + 2 \cdot 2^{i-2} = 2^{i-1} + 2 \cdot 2^{i-1} = 2^i$
- $(-1)^i = (-1)^{i-1} + 2(-1)^{i-2} = (-1)^{i-1} + 2(-1)^i$ , y restando  $(-1)^i$  a ambos lados obtenemos la igualdad  $0 = (-1)^{i-1} + (-1)^i$ .

Luego ambas son soluciones. También, si  $A(n)$  y  $B(n)$  son soluciones a la recurrencia, pasa lo mismo con  $A(n) + B(n)$ , y si  $C(n)$  es una solución a la recurrencia, también  $\gamma C(n)$  lo es, para cualquier  $\gamma \in \mathbb{R}$ :

- Sea  $Z(n) = A(n) + B(n)$ . Entonces  $Z(n) = A(n - 1) + 2A(n - 2) + B(n - 1) + 2B(n - 2) = A(n - 1) + B(n - 1) + 2(A(n - 2) + B(n - 2)) = Z(n - 1) + 2Z(n - 2)$

- Si  $Z(n) = \gamma C(n)$ ,  $Z(n) = \gamma(C(n-1) + 2C(n-2)) = (\gamma C(n-1)) + 2(\gamma C(n-2)) = Z(n-1) + 2Z(n-2)$ .

Luego tenemos una familia de soluciones  $G(n) = \alpha 2^n + \beta (-1)^n$ , a la recurrencia  $G(n) = G(n-1) + 2G(n-2)$ .

Intentemos encontrar los coeficientes  $\alpha, \beta$ , tales que  $G(0) = T(q_0, 0) = 1$ , y  $G(1) = T(q_0, 1) = T(q_1, 0) + T(q_0, 0) = 2T(q_0, -1) + T(q_0, 0) = 2 \times 0 + 1 = 1$ .

- $\alpha 2^0 + \beta (-1)^0 = 1 \Leftrightarrow \alpha + \beta = 1$
- $\alpha 2^1 + \beta (-1)^1 = 1 \Leftrightarrow 2\alpha - \beta = 1$

Los únicos  $\alpha, \beta$  que cumplen esto son  $\alpha = \frac{2}{3}, \beta = \frac{1}{3}$ .

Entonces encontramos una fórmula general para *nuestra*  $G(n)$ :  $G(n) = \left(\frac{2}{3}2^n + \frac{1}{3}(-1)^n\right)$ .

Verificamos que  $G(0)$ , así definida, cumple que  $G(0) = 1$ , y  $G(1) = 1$ , y ya vimos arriba que esta definición cumple la recurrencia pedida. Notemos que hasta ahora, sólo habíamos supuesto la existencia de soluciones de la forma  $G(n) = c^n$  que nos sirvieran, pero la existencia de  $\alpha$  y  $\beta$  nos dice que efectivamente encontramos la solución que tiene valor exactamente iguales a nuestra función  $T(q_0, n)$ .

Luego, como queremos  $T(q_1, n)$ , y  $T(q_1, n) = 2G(n-1)$ , vemos que la respuesta al enunciado es  $T(q_1, n) = \frac{4}{3}2^{n-1} + \frac{2}{3}(-1)^{n-1}$ . □

## 6.4 Combinatoria

### Ejercicio 6.4.1

¿Cuántos números naturales hay menores o iguales que 1000 que no son ni múltiplos de 3 ni múltiplos de 5?



*Demostración.* Para esto vamos a usar el principio de inclusión-exclusión. Simbólicamente, este nos dice que para todo par de conjuntos  $A, B$ :

$$|A \cup B| = |A| + |B| - |A \cap B|$$

Para resolver este ejercicio, usamos  $A = \{n \in \mathbb{N} \mid n \leq 1000 \wedge n \not\equiv 0 \pmod{3}\}$ , y  $B = \{n \in \mathbb{N} \mid n \leq 1000 \wedge n \not\equiv 0 \pmod{5}\}$ . Lo que nos pide el enunciado es  $|A \cap B|$ . Luego, esto es  $|A| + |B| - |A \cup B|$ .

Por definición,  $|A \cup B| = \{n \in \mathbb{N} \mid n \leq 1000 \wedge (n \not\equiv 0 \pmod{3}) \vee (n \not\equiv 0 \pmod{5})\}$ . Pero si prestamos atención, el que algo sea o no múltiplo de tres o no múltiplo de cinco, es lo mismo que no ser múltiplo de quince. Podemos razonarlo o podemos hacer una tabla:

$n$	$n \pmod{3}$	$n \pmod{5}$	$n \pmod{15}$
0	0	0	0
1	1	1	1
2	2	2	2
3	0	3	3

4	1	4	4
5	2	0	5
6	0	1	6
7	1	2	7
8	2	3	8
9	0	4	9
10	1	0	10
11	2	1	11
12	0	2	12
13	1	3	13
14	2	4	14

Luego,  $A \cup B = \{n \in \mathbb{N} \mid n \leq 1000 \wedge n \not\equiv 0 \pmod{15}\}$ . Esto, a su vez, es  $A \cup B = \{n \in \mathbb{N} \mid n \leq 1000\} \setminus \{n \in \mathbb{N} \mid n \leq 1000 \wedge n \equiv 0 \pmod{15}\}$ . Es decir, sacarle los múltiplos de 15, al conjunto de todos los números menores o iguales a 1000. ¿Cuántos tales múltiplos hay? Uno de cada 15 números va a ser múltiplo de 15, luego hay  $\lfloor \frac{1000}{15} \rfloor$  tales números, y  $|A \cup B| = 1000 - \lfloor \frac{1000}{15} \rfloor = 1000 - 66 = 934$ .

Encontrar  $|A|$  y  $|B|$  es similar.  $|A| = 1000 - \lfloor \frac{1000}{3} \rfloor = 667$ , y  $|B| = 1000 - \lfloor \frac{1000}{5} \rfloor = 800$ .

Luego, lo que nos piden es  $|A \cap B| = 667 + 800 - 934 = 533$ . □

### Ejercicio 6.4.2

¿Cuántos palíndromes distintos de longitud  $n$  se pueden armar usando un conjunto de  $k$  símbolos?

*Demostración.* Si algo es un palíndrome, entonces se lee igual hacia adelante que hacia atrás. Sea  $S = \{x_1, x_2, \dots, x_{n-1}, x_n\}$  un tal palíndrome. Entonces  $x_n = x_1$ , y  $x_{n-1} = x_2$ , etcétera. Tenemos que tener cuidado, entonces, con qué pasa en el medio, cuando no hay una igualdad extra. Por ejemplo, si  $n = 3$ , entonces hay sólo una igualdad,  $x_1 = x_3$ . La igualdad  $x_2 = x_2$  no dice nada, entonces no restringe nuestras posibilidades.

Entonces, partamos en dos casos, dependiendo de si  $n$  es par o impar.

- Si  $n \equiv 0 \pmod{2}$ , entonces  $n = 2t$  para algún  $t \in \mathbb{N}$ . Los primeros  $t$  símbolos son totalmente arbitrarios, entonces tenemos  $k^t$  posibles cadenas. Como los siguientes  $t$  símbolos están totalmente determinados por los primeros, no hay posibilidades restantes, y el número de palíndromos de longitud  $n = 2t$  usando un conjunto de  $k$  símbolos es  $k^{\frac{n}{2}}$ .

$x_1$	$x_2$	$\dots$	$x_t$	$x_{t+1} = x_{n-(t+1)} = x_{t-1}$	$x_{t+2} = x_{n-(t+2)} = x_{t-2}$	$\dots$	$x_n = x_1$
-------	-------	---------	-------	-----------------------------------	-----------------------------------	---------	-------------

- Si  $n \equiv 1 \pmod{2}$ , entonces  $n = 2t + 1$ , para algún  $t \in \mathbb{N}$ . Los primeros  $t + 1$  símbolos son arbitrarios, y tenemos  $k^{t+1}$  cadenas. Los últimos  $t$  símbolos están totalmente determinados por los primeros  $t$ , entonces no hay más posibilidades, y tenemos  $k^{t+1}$  palíndromos posibles.

$x_1$	$x_2$	$\dots$	$x_t$	$x_{t+1}$	$x_{t+2} = x_{n-(t+1)} = x_t$	$\dots$	$x_n = x_1$
-------	-------	---------	-------	-----------	-------------------------------	---------	-------------

Vemos entonces que la fórmula general para el número de palíndromos de longitud  $n$  usando un conjunto de  $k$  símbolos es  $k^{\lceil \frac{n}{2} \rceil}$ .  $\square$

### Ejercicio 6.4.3

Sin calcular los valores explícitamente ni expandir a factoriales, probar que

$$\binom{10}{4} = \binom{9}{3} + \binom{9}{4}$$



*Demostración.* Esto es un caso particular de la fórmula  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ , con  $n = 10, k = 4$ .

Queremos dar una demostración de este hecho, sin expandir los coeficientes binomiales a sus factoriales. ¿Qué otra cosa sabemos sobre estos coeficientes? Que  $\binom{n}{k}$  es el número de subconjuntos de tamaño  $k$ , de un conjunto de tamaño  $n$ .

Vamos a usar la técnica de «contar lo mismo de dos formas distintas». Supongamos que tenemos un conjunto  $X$  de tamaño  $n$ . Sea  $Y$  el conjunto de subconjuntos de  $X$  de tamaño  $k$ .

- Por un lado,  $|Y| = \binom{n}{k}$ , porque esa es precisamente la semántica de  $\binom{n}{k}$ .
- Por otro lado, sea  $x \in X$ . Los elementos de  $Y$  se dividen en los que contienen a  $x$ , y los que no contienen a  $x$ . ¿Cuántos hay que no tienen a  $x$ ? Eso es lo mismo que elegir subconjuntos de tamaño  $k$  de  $X \setminus \{x\}$ , que tiene tamaño  $n - 1$ . Luego, hay  $\binom{n-1}{k}$  de esos. ¿Cuántos hay que sí tienen a  $x$ ? Si esos ya tienen a  $x$ , el número de cosas que pueden elegir es  $k - 1$  cosas más, pero no pueden volver a elegir a  $x$ , entonces tienen  $n - 1$  elementos de  $X$  para elegir. Luego, hay  $\binom{n-1}{k-1}$  de esos. Entonces, en total,  $|Y| = \binom{n-1}{k} + \binom{n-1}{k-1}$ .

Luego,  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$  para todo  $n, k \in \mathbb{N}$ , con  $k \leq n$ , en particular vale para  $n = 10, k = 4$ .  $\square$

### Ejercicio 6.4.4 (Teorema binomial)

Probar que para todo  $x, y \in \mathbb{R}$ , y para todo  $n \in \mathbb{N}$ , tenemos que

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$



Para esta demostración les voy a mostrar el proceso de deducción e ideas que hago antes de formalizarla.

Tengo una suma hasta  $n$ , quizás puedo descomponer la suma hasta  $n$  en una suma hasta  $n - 1$ ? A ver....

Es más, tengo algo «a la»  $n$ , entonces puedo descomponer eso como  $(x + y)(x + y)^{n-1}$ . Si ahí uso la hipótesis inductiva, veamos qué queda...

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n \\
&= (x+y) \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \\
&= \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1}
\end{aligned}$$

Quiero ver si puedo combinar estas dos ecuaciones, quizás emparejando coeficientes, porque así se sumarían los coeficientes binomiales. Pensando en cosas que valgan para sumas de coeficientes binomiales, recuerdo que  $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$ , puede ser que emparejando los coeficientes que tengan el mismo  $x^i y^j$ , me queden así los coeficientes? A ver...

Tengo  $i, j \in \mathbb{N}$ , quiero ver qué coeficiente binomial multiplica a  $x^i y^j$  en la primera sumatoria, y luego en la segunda. En la primera,  $i = k + 1, j = n - k$ . Entonces el coeficiente es  $\binom{n-j+(i-1)}{k-i-1}$ . En la segunda, tenemos  $i = k, j = n - k + 1$ , entonces el coeficiente es  $\binom{j+i-1}{i}$ .

Juntando estos dos, tenemos que la sumatoria es  $\sum_{i,j} x^i y^j (\binom{j+i-1}{i-1} + \binom{j+i-1}{i})$ , donde estoy sumando  $i, j$  sobre algún conjunto que no quiero pensar por ahora. Pero esto es bueno, los términos son precisamente de la forma que pensaba que eran, si los sumo me queda  $\binom{j+i}{i}$ .

Ahora tengo que pensar sobre dónde sumo los  $i, j$ . Todos los términos en las sumatorias tienen el mismo grado, son todos de grado  $n + 1$ . Los términos que estoy sumando ahora tienen grado  $i + j$ , entonces  $i + j = n + 1$ . Como quiero que me quede una sumatoria de sólo un índice, elijo  $i$ , y me queda  $j = n + 1 - i$ . ¿Cuáles son los bordes del índice  $i$ ? Hay un término de la sumatoria de la derecha donde tengo  $x^0$ , y acá estoy diciendo  $x^i$ , así que seguro tengo que tener un término con  $i = 0$ . Por otro lado, de la sumatoria izquierda tengo un término  $x^{n+1}$ , y nuevamente acá digo  $x^i$ , así que  $i$  tiene que llegar hasta  $n + 1$ .

Entonces, usando que  $j = n + 1 - i$  y el coeficiente del término  $x^i y^j$  es  $\binom{j+i}{i}$ , la suma de las sumatorias me queda  $\sum_{i=0}^{n+1} \binom{n+1-i+i}{i} x^i y^{n+1-i}$ , y esto es igual a  $\sum_{i=0}^{n+1} \binom{n+1}{i} x^i y^{n+1-i}$ , que es precisamente lo que quiero probar.

OK, perfecto. ¡A formalizarlo! Voy a tener que pensar un rato para hacer menos grotesco el reindexado de las sumas.

*Demostración.* Vamos a usar inducción. Sea  $P(n) : \forall x, y \in \mathbb{R}. (x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$ .

- Caso base,  $P(0)$ .  $P(0) : \forall x, y \in \mathbb{R}. (x+y)^0 = \sum_{i=0}^0 \binom{0}{k} x^k y^{0-k} = \binom{0}{0} x^0 y^0 = 1$ . Como  $(x+y)^0 = 1$  para todo  $x, y$  (sí,  $0^0 = 1$ ), esto es cierto.
- Paso inductivo. Sea  $n \in \mathbb{N}$ . Asumo  $P(n)$ , quiero ver  $P(n+1)$ . Sean  $x, y \in \mathbb{R}$ .

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)(x+y)^n \\
&= (x+y) \left( \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \right), \text{ por hipótesis inductiva} \\
&= x \left( \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \right) + y \left( \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \right) \\
&= \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1} \\
&= \sum_{i=1}^{n+1} \binom{n}{i-1} x^i y^{n-(i-1)} + \sum_{k=0}^n \binom{n}{k} x^k y^{n-k+1}, \text{ llamando } i = k+1, \\
&= \sum_{j=1}^{n+1} \binom{n}{j-1} x^j y^{n+1-j} + \sum_{j=0}^n \binom{n}{j} x^j y^{n+1-j}, \text{ llamando a ambos índices } j
\end{aligned}$$

Acá hay que tener cuidado. Lo que queremos hacer es que los índices sean iguales, que es sacarle el último término ( $j = n+1$ ) a la primer sumatoria, y el primer término ( $j = 0$ ) a la segunda. Lo único que sé es que  $n \in \mathbb{N}$ , entonces tengo al menos un término en cada sumatoria, porque  $n+1 \geq 1$  (para la primera) y  $n \geq 0$  (para la segunda). No podría, si quisiera, sacar dos términos de cada una, porque no sé si *hay* dos términos en cada una.

$$\begin{aligned}
&= \binom{n}{n} x^{n+1} y^{n+1-(n+1)} + \sum_{j=1}^n \binom{n}{j-1} x^j y^{n+1-j} \\
&\quad + \sum_{j=1}^n \binom{n}{j} x^j y^{n+1-j} + \binom{n}{0} x^0 y^{n+1-0} \\
&= \binom{n}{n} x^{n+1} y^0 + \binom{n}{0} x^0 y^{n+1} + \sum_{j=1}^n \left( \binom{n}{j-1} + \binom{n}{j} \right) x^j y^{n+1-j} \\
&= \binom{n+1}{n+1} x^{n+1} y^0 + \binom{n+1}{0} x^0 y^{n+1} + \sum_{j=1}^n \binom{n+1}{j} x^j y^{n+1-j} \\
&= \sum_{j=0}^{n+1} \binom{n+1}{j} x^j y^{n+1-j}
\end{aligned}$$

que es lo que queríamos demostrar.

□

#### Ejercicio 6.4.5

Sean  $f, g : \mathbb{N} \rightarrow \mathbb{N}$ . Se define la convolución de  $f$  y  $g$  como la función  $(f * g)(n) = \sum_{i=0}^n f(i)g(n-i)$ .

Probar que la convolución es asociativa. Es decir, que para cuales quiera  $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ , tenemos  $f * (g * h) = (f * g) * h$ .

Para este ejercicio también voy a escribirles en qué pienso a medida que lo hago.

Primero voy a expandir uno de los lados, a ver qué queda. La igualdad de funciones es igualdad en cada valor de entrada, así que sea  $n \in \mathbb{N}$ .

$$\begin{aligned}(f * (g * h))(n) &= \sum_{i=0}^n f(i)(g * h)(n - i) \\ &= \sum_{i=0}^n f(i) \left( \sum_{j=0}^{n-i} g(j)h(n - i - j) \right)\end{aligned}$$

Puedo que los índices tengan una relación simple, puedo en vez de ir desde  $j = 0$  hasta  $n - i$ , ir desde  $i$  hasta  $n$ , y lo único que estoy haciendo es sumándole  $i$  a  $j$ . Luego cuando en el término digo  $g(j)$ , se convierte en  $g(j - i)$ , y cuando digo  $h(n - i - j)$ , se convierte en  $h(n - i - (j - i)) = h(n - j)$ .

$$\begin{aligned}&= \sum_{i=0}^n f(i) \left( \sum_{j=i}^n g(j - i)h(n - j) \right) \\ &= \sum_{0 \leq i \leq j \leq n} f(i)g(j - i)h(n - j)\end{aligned}$$

Lo de  $f(i)g(j - i)$  se parece bastante a una convolución, específicamente  $(f * g)(j)$ . Tengo entonces que hacer que la suma de  $i$  esté afuera de la suma de  $j$ . Pero eso me quedaría  $\sum_{j=i}^m f(i)g(j - i)$  adentro, que no es lo que quiero tener, porque no es una convolución. Más aún, no podría sacar el  $h(n - j)$  afuera de esa convolución, porque es distinto para cada  $(j)$  término. OK, entonces la sumatoria interna tiene que ser sobre  $i$ .

$$\begin{aligned}&= \sum_{j=0}^n \sum_{i=0}^j f(i)g(j - i)h(n - j) \\ &= \sum_{j=0}^n (f * g)(j)h(n - j) \\ &= ((f * g) * h)(n)\end{aligned}$$

OK, pasémoslo en limpio.

*Demostración.* Sea  $n \in \mathbb{N}$ , y  $f, g, h : \mathbb{N} \rightarrow \mathbb{N}$ .

Entonces:

$$\begin{aligned}
(f * (g * h))(n) &= \sum_{i=0}^n f(i)(g * h)(n - i) \\
&= \sum_{i=0}^n f(i) \left( \sum_{j=0}^{n-i} g(j)h(n - i - j) \right) \\
&= \sum_{i=0}^n f(i) \left( \sum_{j=i}^n g(j - i)h(n - j) \right) \\
&= \sum_{0 \leq i \leq j \leq n} f(i)g(j - i)h(n - j) \\
&= \sum_{j=0}^n \sum_{i=0}^j f(i)g(j - i)h(n - j) \\
&= \sum_{j=0}^n (f * g)(j)h(n - j) \\
&= ((f * g) * h)(n)
\end{aligned}$$

□

## 6.5 Divide and conquer y programación dinámica

La computación está llena de algoritmos que se basan en dividir un problema en subproblemas más pequeños, resolver esos, y luego combinar los resultados. Estos en general van a tener demostraciones por inducción, donde hacemos inducción en el tamaño de los subproblemas que estamos resolviendo. Nuestra tarea es darle semántica al resultado del algoritmo, definir la noción de tamaño, y probar por inducción en el tamaño de una sub-solución, que el algoritmo es correcto con respecto a su semántica, para todos los subproblemas.

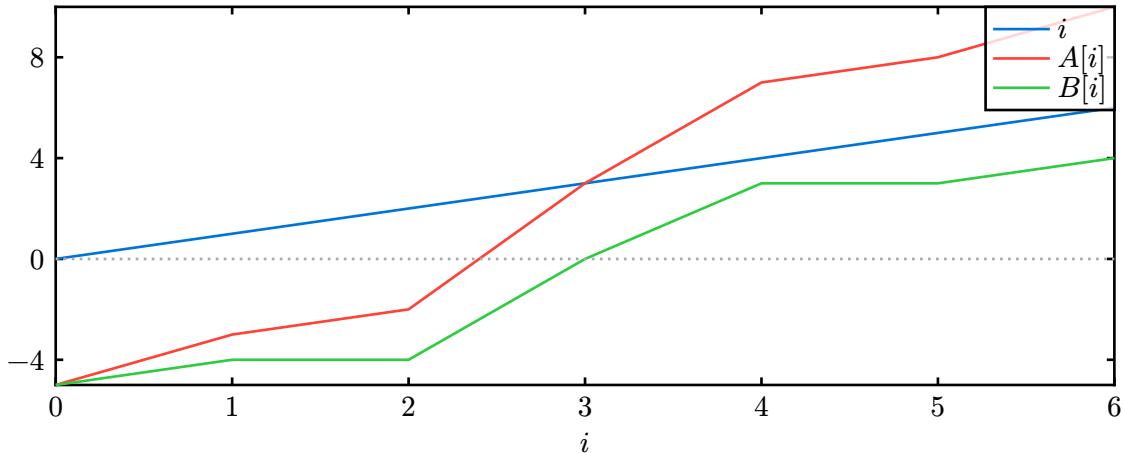


### Ejercicio 6.5.1

Se tiene un array  $A$  de  $n$  enteros, ordenado de manera estrictamente creciente. Dar un algoritmo basado en divide and conquer que determine si existe una posición  $i$  tal que  $A[i] = i$ . Probar su correctitud, y determinar su complejidad asintótica temporal y espacial en el peor caso.

♦

*Solución.* Podemos considerar la lista  $B[i] = A[i] - i$ , y vemos que  $A[i] = i$  exactamente cuando  $B[i] = 0$ . Por ejemplo, para  $A = (-5, -3, -2, 3, 7, 8, 10)$ , tenemos  $B = (-5, -4, -4, 0, 3, 3, 4)$ , que se ve así:



Como  $A$  es una lista creciente de enteros, entonces si  $i \in \mathbb{N}, i < n, A[i + 1] \geq A[i] + 1$ . Restando  $i$  a ambos lados, obtenemos  $A[i + 1] - i \geq A[i] + 1 - i$ , o equivalentemente,  $A[i + 1] - (i + 1) \geq A[i] - i$ , que es  $B[i + 1] \geq B[i]$ , luego  $B$  es creciente.

Luego, nuestro algoritmo tiene que encontrar un  $i$  tal que  $B[i] = 0$ , con  $B$  creciente. Podemos usar búsqueda binaria para esto. Notemos que no hace falta *crear B*, dado que preguntar si  $B[i] > 0$  es lo mismo que preguntar si  $A[i] - i > 0 \Leftrightarrow A[i] > i$ .

```
def f(A: list[int], i: int, j: int) -> bool:
    if j <= i: return False
    k = i + (j - i) // 2
    if A[k] > k:
        return f(A, i, k)
    elif A[k] < k:
        return f(A, k + 1, j)
    return True
```

La función se llama con  $f(A, 0, \text{len}(A))$ . El invariante, como en todas las búsquedas binarias, es que tenemos una propiedad  $P(i, j)$ : Si existe un  $k$  tal que  $A[k] = k$ , entonces  $i \leq k < j$ .

*Demostración.* La semántica que le vamos a dar a  $f(A, i, j)$  es que es `True` si y sólo si existe un  $k$  en  $[i, j)$  tal que  $A[k] = k$ . El algoritmo devuelve  $f(A, 0, \text{len}(A))$ . Si logramos probar que  $f$  es correcta, como todos los tales índices  $k$  están en  $[0, |A|)$ , el algoritmo va a ser correcto.

Definimos el tamaño de un argumento  $(A, i, j)$  como  $t(A, i, j) = j - i$ , y definimos  $P(n)$ : Nuestro programa es correcto para todas las entradas de tamaño menor o igual a  $n$ . Vamos a probar  $P$  por inducción.

1. Caso base,  $P(0)$ . Si  $t(A, i, j) = 0$ , entonces  $i = j$ , y el algoritmo devuelve `False`. La semántica que queríamos que  $f$  cumpliera es que devuelva `True` si y sólo si existe un  $k$  en  $[i, j)$  tal que  $A[k] = k$ , pero claramente no puede haber ningún tal  $k$  si  $i = j$ , pues  $[i, j) = []$ . Luego nuestra función es correcta para entradas con tamaño  $T(a, i, j) = 0$ .

2. Paso inductivo. Asumo que vale  $P(n)$ , pruebo  $P(n + 1)$ . Si  $t(A, i, j) = n + 1 > 0$ , entonces  $j > i$ . Definimos  $k$  como el promedio entre  $i$  y  $j$ ,  $k = i + \frac{j-i}{2} = 2\frac{i}{2} + \frac{j-i}{2} = \frac{i+j}{2}$ . Como vimos arriba,  $B$  es creciente, estricta. Luego:
  1. Si  $B[k] > 0$ , si existe un  $k'$  tal que  $B[k'] = 0$ , tenemos que tener  $k' < k$ . Como tenemos que devolver True exactamente si existe un tal  $k'$  en  $[i, j]$ , y sabemos que si tal  $k'$  existe está antes que  $k$ , entonces el  $k'$  no va a estar en  $[k, j]$ , y tiene que estar, si existe, en  $[i, k)$ . Por ende, nuestro algoritmo es correcto al devolver  $f(A, i, k)$ , que por hipótesis inductiva es True exactamente cuando hay un  $k'$  en  $[i, k)$  tal que  $B[k'] = 0$ .
  2. Si  $B[k] < 0$ , si existe un  $k'$  tal que  $B[k'] = 0$ , tiene que estar después de  $k$ . Como tenemos que devolver True exactamente si existe un  $k'$  en  $[i, j]$  tal que  $B[k'] = 0$ , y de existir tal  $k'$ , tiene que ser mayor a  $k$ , sabemos que debe estar en  $(k, j) = [k + 1, j)$ . Luego, nuestro algoritmo es correcto al devolver  $f(A, k + 1, j)$ , que por hipótesis inductiva es True exactamente cuando hay un  $k'$  en  $[k + 1, j)$  tal que  $B[k'] = 0$ .

Luego, nuestro algoritmo es correcto para todas las entradas.

Si denotamos por  $T(n)$  al número de operaciones que hace nuestro algoritmo al recibir una entrada de tamaño  $t(A, i, j) = n$ , vemos que  $T(0) = c$  para alguna constante  $c$  (no podría ser de otra forma,  $T(0)$  no puede depender de nada). Mientras tanto, que si  $n > 0$ ,  $T(n)$  llama a uno de dos problemas cuyo tamaño es  $k - i$ , y  $j - (k + 1)$  respectivamente, con  $k = \lfloor i + \frac{j-i}{2} \rfloor$ . Por definición de la función  $\lfloor \cdot \rfloor$ , sabemos que  $i + \frac{j-i}{2} - 1 < k \leq i + \frac{j-i}{2}$ .

El primer problema, entonces, tiene tamaño  $k - i \leq i + \frac{j-i}{2} - i = \frac{j-i}{2} = \frac{n}{2}$ . El segundo problema tiene tamaño  $j - (k + 1) \leq j - (i + \frac{j-i}{2} - 1 + 1) = 2\frac{j-i}{2} - \frac{j-i}{2} = \frac{j-i}{2} = \frac{n}{2}$ . Además de las llamadas recursivas, hacemos algún número constante de operaciones.

Lo que suelen hacer en la materia es decir que esto es  $T(n) = T(\frac{n}{2}) + O(1)$ , pero como vemos acá esto no es obviamente correcto (¿qué es  $T(\frac{5}{2})$ , si dijimos que  $T : \mathbb{N} \rightarrow \mathbb{N}$ ?). Nuestro algoritmo a veces va a llamar a un problema de tamaño  $\lfloor \frac{n}{2} \rfloor$ , y otras veces  $n - \lfloor \frac{n}{2} \rfloor - 1$ . Como nos piden hacer un análisis de peor caso, esto es  $\lfloor \frac{n}{2} \rfloor \geq n - \lfloor \frac{n}{2} \rfloor - 1$ , luego en el peor caso siempre caemos en la rama  $\lfloor \frac{n}{2} \rfloor$ , y para hacer esto lo más grande posible<sup>7</sup>, queremos que  $\lfloor \frac{n}{2} \rfloor = \frac{n}{2}$ , con lo cual para caer siempre en esta rama,  $n$  tiene que ser una potencia de 2. Esto nos dice que nuestro peor caso son entradas de tamaño potencia de 2. Usando esto, podemos concluir que la recurrencia que determina el número de operaciones necesarias *en el peor caso* es  $T(n) = T(\frac{n}{2}) + O(1)$ , pues en el peor caso,  $n$  es una potencia de 2.

Para un tratamiento formal sobre cómo resolver recurrencias que tengan  $\lfloor \cdot \rfloor$  y  $\lceil \cdot \rceil$ , pueden ver [3]. Noten el cuidado que hay que tener al argumentar, y cómo tuvimos que valernos de que estamos buscando el comportamiento asintótico *en el peor caso* para definir  $T$ . No existe una función  $T$  que nos da el número de pasos necesarios para *toda* entrada de un dado tamaño, porque el número de tales pasos va a variar. Si queremos

---

<sup>7</sup>Esto asume que  $T$  es creciente, lo cual es cierto en este algoritmo, pero no es cierto para todas las funciones.

definir el comportamiento asintótico en el peor caso, tenemos que primero encontrar una tal familia de casos, argumentar por qué son «el peor caso», y recién ahí podemos valernos de que nuestra entrada tiene alguna forma particular (en este caso, tener longitud potencia de 2).

Luego, usando el teorema maestro vemos que  $T \in O(\log n)$ . La complejidad asintótica espacial es también  $O(\log n)$ , porque usamos memoria para cada llamada recursiva, en particular para guardar  $k$ . Como hay a lo sumo  $\lceil \log_2 n \rceil$  llamadas recursivas, y usamos  $O(1)$  espacio en cada una, usamos  $O(\log n)$  espacio adicional a la entrada en total.  $\square$

### Ejercicio 6.5.2

Demostrar que el algoritmo «Mergesort» es correcto.

```
def merge(xs: [int], ys: [int]) -> [int]:
    i, j, n, m = 0, 0, len(xs), len(ys)
    ans = []
    while i < n or j < m:
        if i < n and (j >= m or xs[i] <= ys[j]):
            ans.append(xs[i])
            i += 1
        else:
            ans.append(ys[j])
            j += 1
    return ans

def mergesort(xs: [int]) -> [int]:
    n = len(xs)
    if n <= 1: return xs
    lefts = mergesort(xs[:n//2])
    rights = mergesort(xs[n//2:])
    return merge(lefts, rights)
```

*Solución.* Este algoritmo consiste de dos sub-algoritmos, `merge` y `mergesort`. Vamos a especificar cada uno, y demostrar que cumplen la especificación.

1. `merge` recibe dos listas de enteros, `xs` e `ys`, ambas ordenadas de forma no-decreciente. Devuelve una lista de enteros `ans`, donde `ans` contiene los mismos elementos que `xs + ys`, y `ans` está ordenada de forma no-decreciente.
2. `mergesort` recibe una lista de enteros `xs`, y devuelve una lista de enteros `zs`, donde `zs` contiene los mismos elementos que `xs`, y `zs` está ordenada de forma no-decreciente.

#### Lema 6.5.3

`merge` es correcto.

*Demostración.* Ya vieron cómo demostrar usando el teorema del invariante. También vieron que los algoritmos con estados son, en general, más difíciles de analizar que los que no mutan estado. En esta demostración quiero mostrarles cómo usar las herramientas que tienen para algoritmos recursivos, para algoritmos con estado.

Todo ciclo se puede transformar a un algoritmo con el mismo comportamiento, pero que usa recursión.

```

1: Estado ← EstadoInicial
2: while CONDICIÓN(Estado) do
3:   Estado ← Modificar(Estado)
4: end
5: return POSTPROCESAMIENTO(Estado)
```

Esto es equivalente al siguiente algoritmo recursivo:

```

1: procedure REC(Estado)
2:   if ¬CONDICIÓN(Estado) then
3:     return POSTPROCESAMIENTO(Estado)
4:   end
5:   return REC(MODIFICAR(Estado))
6: end
7: REC(EstadoInicial)
```

Estos algoritmos tienen la misma semántica, y costo computacional<sup>8</sup>. Para la función `merge`, la versión iterativa queda así:

```

def merge(xs: [int], ys: [int]) -> [int]:
    n, m = len(xs), len(ys)
    def g(i, j, ans):
        if not (i < n or j < m): return ans
        if i < n and (j >= m or xs[i] <= ys[j]):
            return g(i + 1, j, ans + [xs[i]])
        return g(i, j + 1, ans + [ys[j]])
    return g(0, 0, [])
```

La semántica que le vamos a dar a  $g$  es:

$$g : [0, \dots, n] \times [0, \dots, m] \times \text{List[int]} \rightarrow \text{List[int]}$$

$$g(i, j, a) = a + t, \text{ donde } t \text{ está ordenada de forma no-decreciente,}$$

$$\text{y tiene exactamente los elementos de } x[i, \dots, n-1] + y[j, \dots, m-1]$$

Queremos ver en qué vamos a hacer inducción para probar que  $g$  es correcta. Vemos que en cada llamada recursiva, aumenta  $i$  o aumenta  $j$ , y el otro queda igual. Entonces, lo que decrece es  $n - i$ , o  $m - j$ . Si decrece al menos uno, y el otro queda igual, dos opciones para algo en lo que hacer inducción son la suma,  $n - i + m - j$ , y el producto,  $(n - i)(m - j)$ . Si elegimos la suma, entonces vamos a tener problemas para probar el

---

<sup>8</sup>Para justificar que tienen el mismo costo computacional, hay que definir un modelo computacional. Por ejemplo, hay que definir si una llamada recursiva de este estilo usa espacio para el «stack». Es razonable asumir que las funciones de esta forma no requieren más espacio en memoria que el código iterativo. Alumnos interesados pueden leer sobre el concepto de «tail-call recursion» para entender por qué es razonable asumir esto. También los invito a leer cómo su lenguaje favorito implementa esta optimización.

caso base, porque  $n - i + m - j$  no nos dice que  $i = n, m = j$ , que es lo que usa  $g$  para no-hacer recursión. Luego miremos el producto,  $(n - i)(m - j)$ , que sí nos deja concluir eso. El problema acá va a ser que cuando  $i = n$  o  $j = m$ , las llamadas recursivas no bajan el valor de este producto (sigue siendo cero). Luego queremos modificar esto a  $(n - i + 1)(m - j + 1)$ , que no tiene ese problema. Entonces definimos la proposición

#### Definición 6.5.4

$P(k) : g(i, j, a)$  es correcta para todo  $i, j \in \mathbb{N}, 0 \leq i \leq n, 0 \leq j \leq m$  tal que  $(n - i + 1)(m - j + 1) = k$ .



Vamos a definir por comodidad la notación  $a \lesssim b$  como que  $a \leq t \forall t \in b$ . Por ejemplo,  $5 \lesssim [6, 5, 9]$ , pero  $5 \not\lesssim [8, 3, 10]$ .

Para ver que `merge` es correcta, tenemos que probar que vale  $g(0, 0, []])$  es correcta. Si probamos que  $P((n + 1)(m + 1))$ , entonces como  $(n - 0 + 1)(m - 0 + 1) = (n + 1)(m + 1)$ , vemos que  $g(0, 0, [])$  es correcta, y luego que `merge` devuelve una lista que contiene los mismos elementos que  $xs[0 : n] + ys[0 : m] = xs + ys$ , pero ordenada de forma no-decreciente, que es precisamente la semántica que queríamos darle a `merge`.

1. Caso base,  $P(1)$ . Tenemos que probar que para todo  $i, j \in \mathbb{N}, 0 \leq i \leq n, 0 \leq j \leq m$  tal que  $(n - i + 1)(m - j + 1) = 1$ ,  $g$  es correcta. Como  $i \leq n$  y  $j \leq m$ , entonces ambos factores de este producto son números naturales. Si tenemos un producto de naturales que es 1, entonces ambos naturales son 1. Luego,  $n - i + 1 = 1$ , y  $m - j + 1 = 1$ . Esto nos dice que  $i = n$ , y  $j = m$ . En este caso, tenemos que `not (i < n or j < m)`, y entonces  $g(n, m, a)$  devuelve  $a$ . Ahora bien,  $a$  es lo mismo que  $a + []$  y  $[]$  es precisamente la unión de todos los elementos de  $x[i, ..., n - 1] + y[j, ..., m - 1] = x[n, ..., n - 1] + y[m, ..., m - 1] = [] + [] = []$ . Luego,  $g$  es correcta para este caso.
2. Paso inductivo. Sabemos que vale  $P(r)$  para todo  $r < k$ , queremos ver que vale  $P(k)$ . Sean entonces  $i, j \in \mathbb{N}, 0 \leq i \leq n, 0 \leq j \leq m$ , tal que  $(n - i + 1)(m - j + 1) = k$ . Partimos en casos, si  $i = n$ , si  $j = m$ , o si ninguna es cierta.
  - a. Si  $i = n$  y  $j \neq m$ , entonces  $k = m - j + 1$ . Como  $j \leq m$ , entonces  $j < m$ , no salimos en la primer condición (`return ans`). Como  $i < n$  es falso,  $g(i, j, a)$  evalúa a  $g(i, j + 1, a + [y[j]])$ . Como  $m - (j + 1) + 1 < m - j + 1 = k$ , podemos usar la hipótesis inductiva  $P(m - (j + 1) + 1)$ , para concluir que si llamamos  $X = g(i, j + 1, a + [y[j]]) = a + [y[j]] + b$ , entonces  $b$  es una lista que contiene los elementos de  $x[i, ..., n - 1] + y[j + 1, ..., m - 1]$ , ordenados de forma no-decreciente. Como  $i = n$ , entonces  $x[i, ..., n - 1] + y[j + 1, ..., m - 1] = y[j + 1, ..., m - 1]$ . Como  $y$  está ordenada de forma no-decreciente, entonces  $y[j] \lesssim b$ . Luego  $t = [y[j]] + b$  está ordenada de forma no-decreciente, y tiene los mismos elementos que  $y[j, ..., m - 1] = x[i, ..., n - 1] + y[j, ..., m - 1]$ . Luego,  $X = a + t$ , con  $t$  teniendo los mismos elementos que  $x[i, ..., n - 1] + y[j, ..., m - 1]$ , ordenados de forma no-decreciente, que es lo que queríamos demostrar para  $P(k)$ .
  - b. Pasa algo análogo si  $j = m$  y  $i < n$ .

- c. Si  $i < n$  y  $j < m$ , entonces partimos en dos casos, dependiendo de si  $x[i] \leq y[j]$  o no.
- Si  $x[i] \leq y[j]$ ,  $g$  devuelve  $g(i+1, j, a + [x[i]])$ . Como  $n - (i+1) + 1 < n - i + 1$ , entonces  $(n - (i+1) + 1)(m - j + 1) < (n - i + 1)(m - j + 1) = k$ , y podemos usar la hipótesis inductiva  $P((n - (i+1) + 1)(m - j + 1))$  para ver que  $g(i+1, j, a + [x[i]]) = a + [x[i]] + b$ , con  $b$  una permutación no-decreciente de  $x[i+1, \dots, n-1] + y[j, \dots, m-1]$ . Como  $x$  e  $y$  son no-decrecientes,  $x[i] \leq x[i+1, \dots, n-1]$ , y  $x[i] \leq y[j] \leq y[j, \dots, m-1]$ . Luego,  $x[i] \leq b$ , y luego llamando  $t = [x[i]] + b$ , vemos que  $g(i, j, a)$  está devolviendo  $a + t$ , con  $t$  una lista no-decreciente, que contiene los mismos elementos que  $x[i, \dots, n-1] + y[j, \dots, m-1]$ . Esto es precisamente lo que hay que probar para  $P(k)$ .
  - Si  $x[i] > y[j]$ , pasa algo análogo con  $g(i, j+1, a + [y[j]])$ .

Luego, demostramos  $P(k)$  para todo  $k \geq 1$ . □

Habiendo probado que `merge` es correcta para toda entrada, probamos ahora fácilmente que `mergesort` es correcta.

#### Lema 6.5.5

`mergesort` es correcta. ♥

*Demostración.* Al ser `mergesort` una función recursiva, la primer herramienta que vamos a intentar es usar inducción.

Veamos primero, ¿qué es lo que decrece en cada llamada recursiva? Nos dan una lista,  $x$ , y la dividimos en dos partes, aproximadamente de la mitad del tamaño cada vez (lo de aproximado es porque no todas las entradas tienen un número par de elementos). Luego, lo que está decreciendo cada vez es el tamaño de la lista que nos pasan.

#### Definición 6.5.6

$P(k)$ :  $\text{merge}(x)$  tiene los mismos elementos que  $x$ , pero ordenados de forma no-decreciente, para toda lista  $x$  con a lo sumo  $k$  elementos. ♣

- Caso base,  $P(0)$ . Si  $x$  tiene 0 elementos, entonces  $x = []$ , y  $\text{merge}([]) = []$  por su primer `if`, que es la respuesta correcta. Luego vale  $P(0)$ .
- Caso base,  $P(1)$ . Si  $x$  tiene 1 elemento, entonces  $x = [\alpha]$  para algún  $\alpha$ , y  $\text{merge}([\alpha]) = [\alpha]$  por su primer `if`, que es la respuesta correcta. Luego vale  $P(1)$ .
- Paso inductivo. Sea  $k \in \mathbb{N}$ ,  $k > 1$ . Asumo que vale  $P(r)$  para todo  $r < k$ , quiero ver que vale  $P(k)$ . Sea  $a = \lfloor \frac{k}{2} \rfloor$ . Como  $k > 1$ , entonces  $a > 0$ . Sea  $b = k - a$ . Luego  $a < k$ , y  $b < k$ . Luego podemos usar las hipótesis inductivas  $P(a)$  y  $P(b)$ , para ver que `lefts` tiene los mismos elementos que  $x[0, \dots, a-1]$ , y `rights` tiene los mismos elementos que  $x[a, \dots, n-1]$ . Luego, su concatenación `lefts + rights` tiene los

mismos elementos que  $x$ . Vemos entonces que llamando a `merge(lefts, rights)`, tendremos una lista ordenada de forma no-decreciente, que tiene los mismos elementos que `lefts + rights`, que a su vez son los mismos elementos que  $x$ . Esto es precisamente la semántica que queríamos para `mergesort`, y luego vale  $P(k)$ .

□

### Ejercicio 6.5.7

Se tienen  $n$  objetos de pesos  $p_1, \dots, p_n$  no-negativos, y valores  $v_1, \dots, v_n$  no-negativos, y una mochila en la que caben varios objetos, pero aguanta como máximo un peso  $P$ .

1. Diseñar un algoritmo basado en programación dinámica que encuentre el máximo valor alcanzable poniendo objetos en la mochila.
2. Demostrar que el algoritmo es correcto.
3. Demostrar su complejidad temporal y espacial, en el peor caso. El mejor algoritmo que conocemos tiene complejidad espacial  $O(P)$  y complejidad temporal  $O(np)$ .

♣

Primero una explicación de cómo podemos pensar esto, y luego una solución como la que se espera que escriban en un parcial.

En principio, el espacio de búsqueda que tenemos es el conjunto de subconjuntos de los  $n$  objetos. Por ejemplo, si  $n = 3$ , y los objetos son  $X = \{a, b, c\}$ , el espacio de búsqueda que tenemos es  $\mathcal{P}(X) = \{\{\}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}$ . Una estrategia simple que usa esta estructura sería *backtracking*.

Si queremos ver si un problema puede ser resuelto con programación dinámica, tenemos que ver si podemos parametrizar al conjunto de subproblemas de tal manera que tengan un orden, y se cumplan las dos propiedades clásicas de subestructura óptima y subproblemas compartidos.

En este caso, podemos definir los subproblemas como pares  $G = \{(i, x) \mid 0 \leq i \leq n, 0 \leq x \leq P\}$ , donde un subproblema  $(i, x)$  significa «El mayor valor que podemos obtener usando peso a lo sumo  $x$ , y usando sólo los primeros  $i$  objetos.» Llamemos  $v^* : G \rightarrow \mathbb{R}$  la función que nos da tal valor. Vemos entonces que la respuesta al problema entero es la respuesta al subproblema  $(n, P)$ . Más aún, vemos que se cumplen las dos condiciones:

1.  $v^*(i, x)$  es o bien  $v^*(i - 1, x)$ , si una solución al subproblema  $(i, x)$  no usa el  $i$ -ésimo objeto, o  $v^*(i - 1, x - p_i) + v_i$ , si una solución al subproblema  $(i, x)$  usa el  $i$ -ésimo objeto. En el primer caso, tenemos el mismo peso disponible ( $x$ ) para usarlo de la mejor forma posible usando sólo los primeros  $i - 1$  objetos, y nuestro valor es el valor que nos den los objetos que elegimos de esos primeros  $i - 1$ . En el segundo caso, si usamos el  $i$ -ésimo objeto, tenemos  $x - p_i$  peso restante para los otros objetos que vamos a elegir dentro de los primeros  $i - 1$ , y el valor de esta solución es el valor de la solución a  $(i - 1, x - p_i)$ , más el valor que obtenemos por haber tomado el  $i$ -ésimo objeto,  $v_i$ .
2. La función  $v^*$  tiene dominio  $[0 \dots n] \times [0 \dots P]$ , luego hay sólo  $nP$  valores posibles. Sin embargo, en una implementación recursiva tradicional, en el peor caso vamos a tener un número exponencial de llamadas. Esto es fácil de ver si tomamos como familia de casos donde  $p_i = 0 \forall i$ ,

vemos que  $v^*(i, x) = \max(v^*(i - 1, x), v^*(i - 1, x) + v_i)$ , y luego una implementación recursiva tradicional tendría  $T(i) = 2T(i - 1)$  llamadas, lo cual termina siendo  $2^n$  llamadas.

Como tenemos ambos subestructura óptima, y subproblemas compartidos, podemos usar programación dinámica y esperar mejoras en el tiempo de cómputo.

Más aún, vemos que  $v^*(i, x)$  sólo depende de cosas en  $v^*(i - 1, \dots)$ , luego si usamos programación dinámica bottom-up, sólo necesitamos quedarnos con dos «filas» de la matriz de programación dinámica, porque al llenar una fila, sólo necesitamos ver la fila inmediatamente anterior, para responder el problema sólo necesitamos una entrada en la última fila.

*Solución.* Primero planteamos la función recursiva, junto con su semántica.

$$f : [0 \dots n] \times [0..P]$$

$f(i, x)$  : El máximo valor que puedo obtener usando los primeros  $i$  objetos y un peso de a lo sumo  $x$ .

La respuesta al enunciado es  $f(n, P)$ . Las ecuaciones recursivas para  $f$  son:

$$f(i, x) = \begin{cases} 0 & , \text{ si } i = 0 \\ f(i - 1, x) & , \text{ si } p_i > x \\ \max(f(i - 1, x), f(i - 1, x - p_i) + v_i) & , \text{ si no} \end{cases}$$

*Demostración.* Vamos a probar que  $f$  es correcta usando inducción. Para poder razonar formalmente, vamos a definir algunos conceptos.

- Definimos  $v(S) = \sum_{j \in S} v_j$ , la suma de los valores de los elementos de un subconjunto  $S \subseteq \{1, \dots, n\}$ . Asimismo  $p(S) = \sum_{j \in S} p_j$ , la suma de los pesos de los elementos de  $S$ .
- Definimos  $F(i, x) = \{S \subseteq \{1, \dots, i\} \mid p(S) \leq x\}$ . Estos son los conjuntos de objetos, de entre los primeros  $i$ , que podemos meter en la mochila, con suma de peso a lo sumo  $x$ . Algunos elementos de  $F(i, x)$  van a tener valor más alto que otros. Vemos que  $F(i, x) \subseteq F(i + 1, x)$  para todo  $1 \leq i < n$ .
- Definimos  $v^*(i, x) = \max_{S \in F(i, x)} \{v(S)\}$ , el máximo valor que podemos obtener, usando los primeros  $i$  objetos, con peso total a lo sumo  $x$ .

Sea  $P(i) : i \leq n \Rightarrow f(i, x) = v^*(i, x)$ . Vamos a probar  $P(i) \forall i \in \mathbb{N}$  por inducción.

- Caso base,  $P(0)$ . Tenemos que probar que  $f(0, x) = v^*(0, x)$ . Por definición,  $v^*(0, x) = \max_{S \in F(0, x)} \{v(S)\}$ , pero  $F(0, x) = \{S \subseteq \{1, \dots, 0\} \mid p(S) \leq x\} = \{\emptyset\}$ , pues el único subconjunto de los primeros 0 objetos es  $\emptyset$ . Luego,  $v^*(0, x) = v(\emptyset) = \sum_{j \in \emptyset} v_j = 0$ . Nuestra función  $f$  efectivamente devuelve 0, en su primer rama, donde  $i = 0$ . Luego  $f(0, x) = 0 = v^*(0, x)$ , lo cual demuestra  $P(0)$ .
- Paso inductivo. Sabemos que vale  $P(t)$ , queremos ver que vale  $P(t + 1)$ . Es decir, queremos probar que  $v^*(t + 1, x) = f(t + 1, x)$ . Llamemos  $i = t + 1$ . Si  $i > n$ , no hay nada que probar, pues «falso implica todo», y estamos probando una implicación ( $P(i)$ ) con antecedente falso. Luego, asumimos que  $i \leq n$ .

Por definición,  $v^*(i, x) = \max_{S \in F(i, x)} \{v(S)\}$ .

- Si  $p_i > x$ , sea  $S \in F(i, x)$ . Como  $S \in F(i, x)$ , sabemos que  $p(S) \leq x$ . Si  $i$  estuviera en  $S$ , tendríamos que  $p(S) = \sum_{j \in S} p_j \geq p_i > x$ , lo cual no sucede. Luego  $i \notin S$ . Como esto sucede para cualquier  $S \in F(i, x)$ , ningún  $S \in F(i, x)$  contiene a  $i$ , y entonces  $S \subseteq \{1, \dots, i\} \setminus \{i\} = \{1, \dots, i-1\}$ . Luego,  $F(i, x) \subseteq F(i-1, x)$ , y como sabíamos que  $F(i-1, x) \subseteq F(i, x)$ , tenemos que  $F(i, x) = F(i-1, x)$ . Luego,  $v^*(i, x) = \max_{S \in F(i, x)} \{v(S)\} = \max_{S \in F(i-1, x)} \{v(S)\} = v^*(i-1, x)$ . Como sabemos  $P(t)$ , es decir  $P(i-1)$ , sabemos que  $f(i-1, x) = v^*(i-1, x)$ . Como  $f(i, x)$  devuelve  $f(i-1, x) = v^*(i-1, x)$  cuando  $p_i > x$ , y en ese caso  $v^*(i-1, x) = v^*(i, x)$ , tenemos  $f(i, x) = v^*(i, x)$ , es decir  $P(i)$ .
- Si  $p_i \leq w$ , podemos particionar  $F(i, x)$  en dos subconjuntos,  $A$  y  $B$ , con  $A = \{S \in F(i, x) \mid i \notin S\}$ , y  $B = \{S \in F(i, x) \mid i \in S\}$ . Como  $A$  y  $B$  partitionan  $F(i, x)$ , entonces  $\max_{F(i, x)} \{v(S)\} = \max(\max_{S \in A} \{v(S)\}, \max_{S \in B} \{v(S)\})$ .
  - Tomemos un  $S \in A$ . Vemos que  $A = \{S \in F(i, x) \mid i \notin S\} = \{S \subseteq \{1, \dots, i\} \mid p(S) \leq x \wedge i \notin S\} = \{S \subseteq \{1, \dots, i-1\} \mid p(S) \leq w\} = F(i-1, x)$ . Luego,  $\max_{S \in A} \{v(S)\} = \max_{S \in F(i-1, x)} \{v(S)\} = v^*(i-1, x)$ .
  - Tomemos ahora un  $S \in B$ . Como  $S \in B$ , entonces  $i \in S$ . Llamemos  $S = S' \cup \{i\}$ , con  $S' \subseteq \{1, \dots, i-1\}$ . Luego,  $p(S) = p_i + p(S')$ . Como  $S \in B \subseteq F(i, x)$ ,  $p(S) \leq x$ , y luego  $p_i + p(S') \leq x$ . Por ende,  $p(S') \leq x - p_i$ . Luego,  $S' \in F(i-1, x - p_i)$ . Luego, cada  $S \in B$  se corresponde con un único  $S' \in F(i-1, x - p_i)$ , y la biyección es simplemente  $\varphi(X) = X \cup \{i\}$ . Para transformar los valores luego de esta biyección, vemos que  $v(S) = v_i + v(S')$ . Luego,
$$\begin{aligned}\max_{S \in B} \{v(S)\} &= \max_{S' \in F(i-1, x - p_i)} \{v(S') + v_i\} = \\ \max_{S' \in F(i-1, x - p_i)} \{v(S')\} + v_i &= v^*(i-1, x - p_i) + v_i.\end{aligned}$$

Juntando ambas ramas, vemos que  $v^*(i, x) = \max_{F(i, x)} \{v(S)\} = \max(v^*(i-1, x), v^*(i-1, x - p_i) + v_i)$ . Por  $P(t)$ , que es  $P(i-1)$ , sabemos que  $f(i-1, x) = v^*(i-1, x)$ , y  $f(i-1, x - p_i) = v^*(i-1, x - p_i)$ . Por lo tanto, como nuestra función devuelve  $\max(f(i-1, x), f(i-1, x - p_i))$ , está devolviendo  $v^*(i, x)$ , que prueba  $P(i)$ .

Esto prueba  $P(i)$  para todo  $i \in \mathbb{N}$ . En particular, para todo  $i \in \mathbb{N}$ ,  $0 \leq i \leq n$ , y para todo  $x \in \mathbb{N}$ ,  $0 \leq x \leq P$ , tenemos que  $f(i, x) = v^*(i, x)$ , y luego  $f(n, P) = v^*(n, P)$ , que muestra que nuestro algoritmo es correcto.  $\square$

Veamos el código ahora en Python:

```
def f(i, x):
    if i == 0: return 0
    if p[i] > x: return f(i - 1, x)
    return max(f(i - 1, x), f(i - 1, x - p[i]) + v[i])
```

Si queremos hacer esto un algoritmo de programación dinámica top-down, lo único que hay que hacer es mecánicamente agregar un cache.

```
def f(i, x, cache = {}):
    if (i, x) not in cache:
        if i == 0: r = 0
        elif p[i] > x: r = f(i - 1, x, cache)
        else: r = max(f(i - 1, x), f(i - 1, x - p[i]) + v[i])
    cache[(i, x)] = r
    return r
```

```

    else: r = max(f(i - 1, x, cache), f(i - 1, x - p[i], cache) + v[i])
    cache[(i, x)] = r
    return cache[(i, x)]

```

El número asintótico de operaciones en este algoritmo es más difícil de analizar, porque está mutando el estado (cache) a medida que hacemos llamadas recursivas. También va a depender del costo que tenga insertar y buscar en la estructura que usamos para el cache.

Para hacer este algoritmo bottom-up, tenemos que pensar en qué orden se llena la tabla, y llenarla nosotros mismos. Vemos que necesitamos leer un valor de  $f(i, x)$  sólo cuando estamos escribiendo el valor de  $f(i + 1, x')$  para algún  $x'$ . Luego, si llenamos la tabla en orden creciente de  $i$ , cada vez que queramos leer un valor, ya lo vamos a tener en la tabla.

```

def f(n, P):
    dp = [[0 for _ in range(P + 1)] for _ in range(n + 1)]
    for i in range(1, n + 1):
        for x in range(P + 1):
            if x < p[i]:
                dp[i][x] = dp[i - 1][x]
            else:
                dp[i][x] = max(dp[i - 1][x], dp[i - 1][x - p[i]] + v[i])
    return dp[n][P]

```

El comportamiento asintótico de este algoritmo es mucho más fácil de analizar, el número de operaciones está en  $\Theta(nP)$  en todos los casos, y el costo espacial es también  $\Theta(nP)$ , pues la tabla  $dp$  tiene  $n \times P$  entradas.

Por último, vemos que no hace falta mantener en memoria toda la tabla, si sólo queremos devolver  $dp[n][P]$ . Para llenar una fila,  $dp[i]$ , sólo hace falta la fila anterior,  $dp[i - 1]$ .

Luego podemos mantener sólo dos filas a la vez,  $dp1$  y  $dp2$ , donde  $dp2 = dp[i]$ , y  $dp1 = dp[i - 1]$ :

```

def f(n, P):
    dp1 = [0 for _ in range(P + 1)]
    dp2 = [0 for _ in range(P + 1)]
    for i in range(1, n + 1):
        for x in range(P + 1):
            if x < p[i]:
                dp2[x] = dp1[x]
            else:
                dp2[x] = max(dp1[x], dp1[x - p[i]] + v[i])
        dp1, dp2 = dp2, dp1
    return dp1[P]

```

El costo temporal es idéntico, pero bajamos el costo espacial a sólo  $\Theta(P)$  en todos los casos.

### Ejercicio 6.5.8

Sea  $X = [x_1, x_2, \dots, x_n]$  una secuencia de  $n$  booleanos (1 o 0) y sea  $k \in \mathbb{N}$  un número entre 1 y  $n$ . Supongamos que se pueden eliminar  $k$  ceros, queremos saber la longitud máxima que puede tener una cadena de 1s. Por ejemplo si  $k = 2$  y  $X = 11001010001$  la respuesta es 3, mientras que si  $k = 3$  la respuesta es 4.

1. Diseñar un algoritmo basado en programación dinámica que indique la longitud más larga de una subsecuencia de unos sacando a lo sumo  $k$  ceros de  $S$ . Debe tener complejidad a lo sumo  $O(nk)$ .
2. Demostrar que el algoritmo es correcto.
3. Demostrar su complejidad temporal y espacial en el peor caso.

Primero les voy a mostrar en qué pienso al resolver el ejercicio, y luego una resolución.

Hay una estructura de orden ahí, porque si hablo de «puedo eliminar  $k$  ceros», al eliminar un cero, caigo a un estado en el que «puedo eliminar  $k - 1$  ceros». Ese probablemente va a ser uno de los parámetros de mi función.

Si veo un 1, puedo armar una cadena de unos que termina en ese 1. Pero al hacer una llamada recursiva, no puedo preguntar algo como «la secuencia de 1s más larga borrando  $k$  ceros que termina antes de esta posición», porque esa cadena puede terminar mucho más atrás que la posición que estoy viendo, y entonces no podría tomar ese número y sumarle uno, extendiendo esa solución con el 1 que estoy mirando. Luego, tengo que restringir que la cadena de 1s termine exactamente donde estoy parado. Luego haré un pasada por la lista, diciendo que la cadena más larga de 1s borrando  $k$  ceros, es la cadena más larga de 1s borrando  $k$  ceros que termina exactamente en  $i = 0$ , o la que termina exactamente en  $i = 1$ , etc.

Veamos si puedo plantear esto.

- Si veo un 1 en  $x_i$ , entonces puedo tomar  $f(i - 1, k) + 1$  como la respuesta para esta posición.
- Si veo un 0 en  $x_i$ , entonces puedo tomar  $f(i - 1, k - 1) + 1$  si  $k > 0$ , o 0 si no. También puedo tomar una cadena de longitud 0 que terminar en este 0, si no quiero tomar la solución recursiva (que puede no existir, si por ejemplo estoy en un prefijo de más de  $k$  ceros).

$$f(i, k) = \begin{cases} -\infty & \text{si } k < 0 \\ 0 & \text{si } i < 0 \\ f(i - 1, k) & \text{si } x_i = 1 \\ \max(0, f(i - 1, k - 1)) & \text{si } x_i = 0 \wedge k > 0 \\ 0 & \text{si no} \end{cases}$$

algo como:

```
def F(x, k):  
    def f(i, kk):  
        if k < 0: return -99999  
        if i < 0: return 0  
        if x[i] == 1: return f(i - 1, kk) + 1  
        if x[i] == 0 and kk > 0: return f(i - 1, kk - 1)  
        return 0  
    return max(f(i, k) for i in range(len(x)))
```

Probablemente ni necesito lo de  $-\infty$ . OK, a escribir.

*Solución.* Definimos primero una función y su semántica.

$$f : [-1, \dots, n] \times [0, \dots, k]$$

$$f(i, r) = \begin{cases} f(i - 1, r) + 1 & \text{si } i \geq 0 \wedge x_i = 1 \\ f(i - 1, r - 1) & \text{si } i \geq 0 \wedge x_i = 0 \wedge r > 0 \\ 0 & \text{si no} \end{cases}$$

La semántica de  $f(i, r)$  es «La longitud de la cadena de unos más larga que termina en la posición  $i$ , borrando a lo sumo  $r$  ceros.»

El algoritmo devuelve

$$\max_{0 \leq i \leq n} f(i, k)$$

*Demuestra.* Está claro que toda cadena de unos borrando a lo sumo  $k$  ceros termina en alguna posición. Luego, si probamos que  $f$  es correcta (es decir, que cumple su semántica), estamos probando que nuestro algoritmo es correcto, pues estamos tomando el máximo sobre todas las posibles posiciones donde terminaría tal secuencia.

Vamos a probar que  $f$  es correcta usando inducción. Definimos  $v^*(i, t)$  como la longitud de cadena de unos más larga, borrando a lo sumo  $t$  ceros, que termina exactamente en  $i$ , 0 cero si no existen tales cadenas. Definimos  $P(i) : i \leq n \Rightarrow (f(i, t) = v^*(i, t) \forall t \in \mathbb{N})$ . Por comodidad, vamos a definir  $\theta(i, t)$  como el conjunto de cadenas de unos que termina en la posición  $i$ , y borra a lo sumo  $t$  ceros, y  $\theta^*(i, t)$  como el subconjunto de  $\theta(i, t)$  que tiene número máximo de unos, para cada  $i, t$ .

- 1) Caso base,  $P(0)$ . La longitud de una cadena de unos más larga que termina en la ( $i = 0$ )-ésima posición es o bien 1 o 0, dependiendo de si  $x_0 = 1$  o  $x_0 = 0$ . Luego, si  $v^*(0, t) = x_0$ . Si  $x_0 = 1$ , nuestra función  $f(0, t)$  devuelve  $f(-1, t) + 1$ , que evalúa a 1 inmediatamente. Si  $x_0 = 0 \wedge t > 0$ ,  $f(0, t)$  devuelve  $f(-1, t - 1)$ , que evalúa a 0 inmediatamente. Finalmente, si  $x_0 \wedge t = 0$ , entonces  $f(0, 0) = 0$ . Luego en todos los casos tenemos  $f(0, t) = v^{0,t}$ , lo que prueba  $P(0)$ .
- 2) Paso inductivo. Sabemos  $P(i)$ , queremos probar  $P(i + 1)$ . Sea  $t \in \mathbb{N}$ . Sea  $T \in \theta^*(i + 1, t)$ , y por ende  $|T| = v^*(i + 1, t)$ . Partimos en casos, dependiendo de quién es  $x_i$ :
  - a) Si  $x_{i+1} = 1$ . Como  $T$  termina en  $i + 1$ ,  $i + 1 \in T$ , puesto que sólo podemos borrar ceros. Sea  $T' = T \setminus \{i + 1\}$ . Como  $T'$  borra a lo sumo  $t$  ceros, y termina en  $i$ , está en  $\theta(i, t)$ . Si no estuviera en  $\theta^*(i, t)$ , podríamos tomar cualquier  $S \in \theta^*(i, t)$ , y por lo tanto  $|S| > |T'|$ , con lo cual  $|S + \{i + 1\}| > |T|$ , pero esto no puede pasar, porque  $T$  está en  $\theta^*(i + 1, t)$ , luego tiene el número máximo de unos. Luego,  $T' \in \theta^*(i, t)$ , y luego  $|T'| = v^*(i, t)$ . Como sabemos  $P(i)$ , esto es  $|T'| = f(i, t)$ , y por ende,  $|T| = f(i, t) + 1$ , que es precisamente lo que devuelve  $f(i + 1, t)$ , y por ende vale  $P(i + 1)$ .
  - b) Si  $x_{i+1} = 0 \wedge t > 0$ . Entonces  $i + 1 \notin T$ . Entonces, como  $t > 0$ , y  $T$  termina en  $i + 1$ , vemos que  $T \in \theta(i, t - 1)$ , y luego  $v^*(i + 1, t) = |T| \leq v^*(i, t - 1)$ . Tomemos ahora cualquier  $S \in \theta^*(i, t - 1)$ . Podemos expandir  $S$  a terminar en  $i + 1$ , borrando el elemento  $i + 1$ , con lo cual  $S$  también está en  $\theta(i + 1, t)$ . Luego  $v^*(i, t - 1) = |S| \leq v^*(i + 1, t)$ . Luego  $v^*(i, t - 1) = v^*(i + 1, t)$ . Por  $P(i)$ ,

$f(i, t - 1) = v^*(i, t - 1)$ , y entonces como  $f(i + 1, t)$  devuelve  $f(i, t - 1)$ , devuelve  $v^*(i, t - 1) = v^*(i + 1, t)$ , que muestra  $P(i + 1)$ .

- c) Si  $x_{i+1} = 0 \wedge t = 0$ , entonces  $T$  es una secuencia de unos que termina en un  $i + 1$ , pero no puede borrar ningún cero pues  $t = 0$ . Luego  $T$  no existe, y por definición,  $v^*(i + 1, 0) = 0$  en este caso, que es precisamente lo que devuelve nuestra función.

Luego tenemos  $P(i) \forall i \in \mathbb{N}$ . □

El código en Python para la función recursiva es:

```
def F(x, k):
    def f(i, t):
        if i < 0: return 0
        if x[i] == 1: return f(i - 1, t) + 1
        if x[i] == 0 and t > 0: return f(i - 1, t - 1)
        return 0
    return max(f(i, k) for i in range(len(x)))
```

La manera que formulamos el problema como un conjunto de subproblemas y un orden entre ellos cumple dos propiedades:

1. Subestructura óptima. Para resolver un problema  $(i, t)$ , necesitamos resolver algún número de subproblemas más pequeños  $((i - 1, t)$  o  $(i - 1, t - 1)$ ). Esto es esencialmente lo que nos deja usar recursión.
2. Subproblemas compartidos. Podemos ver que el peor caso sucede cuando  $k$  es muy grande o  $x$  está lleno de unos (y nunca nos quedamos «sin presupuesto» de ceros para borrar). En esos casos, llamar a  $f(i, t)$  resulta en  $i$  llamadas recursivas. Como luego tomamos un  $\max$  para cada  $i$  entre 1 y  $n$ , vamos a computar  $\frac{n(n-1)}{2}$  problemas en el peor caso, muchos de los cuales son idénticos.

Como siempre, para hacer esto un algoritmo de programación dinámica top-down, mecánicamente agregamos un cache:

```
def F(x, k):
    def f(i, t, cache = {}):
        if (i, t) not in cache:
            if i < 0: r = 0
            elif x[i] == 1: r = f(i - 1, t, cache) + 1
            elif x[i] == 0 and t > 0: r = f(i - 1, t - 1, cache)
            else: r = 0
            cache[(i, t)] = r
        return cache[(i, t)]
    return max(f(i, k) for i in range(len(x)))
```

Esto evita computar subproblemas dos veces, pero hace difícil el análisis de complejidad temporal, dado que estamos mutando estado (cache), y el tiempo que va a tomar una llamada va a depender del estado cuando es llamada. Asimismo, agregamos ahora el costo adicional de leer y escribir la estructura cache.

Para hacer más claro el análisis algoritmo, y bajar su complejidad en la práctica cuando vamos a llenar cache enteramente de todos modos, podemos usar programación dinámica bottom-up.

Esto implica ver en qué orden se llena cache en la versión top-down, y llenarlo nosotros mismos en ese orden. En este caso, vemos que llenamos una entrada  $\text{cache}[(i, t)]$  sólo luego de llamar a  $f(i - 1, \dots)$ , que va a escribir  $\text{cache}[(i - 1, \dots)]$ . Luego, si llenamos cache en orden creciente de  $i$ , vamos a estar llenando la estructura en un orden que garantiza siempre tener escritos los valores que queremos leer, al momento de querer leerlos.

```
def F(x, k):
    n = len(x)
    dp = [[0 for _ in range(k + 1)] for _ in range(n)]
    for i in range(n):
        for t in range(k + 1):
            if x[i] == 1:
                if i == 0: dp[i][t] = 1
                else: dp[i][t] = dp[i-1][t] + 1
            else:
                if t > 0:
                    if i == 0: dp[i][t] = 0
                    else: dp[i][t] = dp[i - 1][t - 1]
                else:
                    dp[i][t] = 0
    return max(dp[i][k] for i in range(n))
```

Esa es una traducción totalmente literal, donde quedaron varios condicionales porque no podemos leer  $\text{dp}[-1]$ . Podemos limpiar el código un poco:

```
def H(x, k):
    n = len(x)
    dp = [[0 for _ in range(k + 1)] for _ in range(n)]
    for t in range(k + 1): dp[0][t] = x[0]
    for i in range(1, n):
        for t in range(k + 1):
            if x[i] == 1: dp[i][t] = dp[i - 1][t] + 1
            elif x[i] == 0 and t > 0: dp[i][t] = dp[i - 1][t - 1]
            else: dp[i][t] = 0
    return max(dp[i][k] for i in range(n))
```

Vemos que el número de operaciones que hacemos en todos los casos es  $\Theta(nk)$ , pues eso cuesta construir el array  $\text{dp}$ . Los dos ciclos anidados hacen  $\Theta(nk)$  operaciones, y el ciclo de inicialización de  $\text{dp}[0]$  hace  $\Theta(k)$  operaciones. El ciclo final, que computa  $\max$ , hace  $\Theta(n)$  operaciones.

El costo espacial del algoritmo es, en todos los casos,  $\Theta(nk)$ .

### Ejercicio 6.5.9

Sea  $v = (v_0, v_1, \dots, v_{n-1})$  un vector de números enteros. Diseñar un algoritmo que indique la mínima cantidad de números que hay que eliminar del vector para que cada número que permanezca sea múltiplo del anterior (excepto el primero). Por ejemplo, para los vectores  $(-5, 5, 0)$ ,  $(0, 5, -5)$ , y  $(0, 5, -5, 2, 15, 15)$ , los resultados deberían ser respectivamente 0, 1, y 2. El algoritmo debe tener complejidad temporal  $O(n^2)$  y estar basado en programación dinámica.

1. Demostrar que el algoritmo es correcto.
2. Demostrar su complejidad temporal y espacial.

*Solución.* Definimos una función recursiva y su semántica.

$$f : [0, \dots, n) \rightarrow N$$

$f(i)$  = La longitud de la subsecuencia de  $v$  más larga,  
donde cada elemento es múltiplo del anterior,  
y que termina exactamente en  $v_i$ .

La respuesta que devuelve el algoritmo es

$$A = n - \max_{0 \leq i < n} \{f(i)\}$$

Y ahora las ecuaciones que definen  $f$ .

$$f(i) = \begin{cases} 1 & \text{si } i = 0 \\ 1 + \max_0 \{f(j) \mid j \in [0, i), v_j \mid v_i\} & \text{si no} \end{cases}$$

donde definimos  $\max_0$  como max, excepto que en el conjunto vacío devuelve 0. Tenemos que probar dos cosas:

1. La función  $f$  cumple la semántica que le dimos.
2. Si la función  $f$  cumple la semántica que le dimos, entonces  $A$  es la respuesta correcta al enunciado.

Probemos ambas, entonces.

#### Demostración.

1. Vamos a probar que  $f$  cumple la semántica que le dimos por (repitan conmigo) inducción. Definimos entonces  $S(i)$  como el conjunto de subsucesiones de  $v$  que terminan en  $v_i$  y cada elemento es múltiplo del anterior, y definimos  $g(i) = \max\{|s| \mid s \in S(i)\}$ . Queremos probar la proposición  $P(i) : i < n \Rightarrow f(i) = g(i)$ , para todo  $i \in \mathbb{N}$ .
  1. Caso base,  $P(0)$ . Queremos ver que  $f(0) = g(0)$ . Por definición,  $f(0) = 1$ .  $g(0)$  es la longitud de la subsecuencia de  $v$  más larga, donde cada elemento es múltiplo del anterior, y que termina en exactamente en  $v_0$ . Pero  $v_0$  es el primer elemento, luego hay una sola tal subsucesión, y es  $[v_0]$ , que tiene longitud 1. Por lo tanto,  $g(0) = 1$ , y tenemos  $f(0) = g(0)$ , probando  $P(0)$ .
  2. Paso inductivo. Sabemos  $P(j)$  para todo  $j < i$ , queremos ver  $P(i)$ . Si  $i \geq n$ , entonces vale  $P(i)$  trivialmente, pues  $P(i)$  es una implicación con antecedente  $i < n$ , y falso implica todo. Luego, sabemos que  $i < n$ , y tiene sentido hablar de  $v_k$ , con  $0 \leq k \leq i$ . Sea  $s \in S(i)$ . Como  $s \in S(i)$ , sabemos que  $s$  termina con  $v_i$ . Sea  $s'$  el prefijo de  $s$ , es decir,  $s = s' + [v_i]$ , y por ende  $|s| = |s'| + 1$ . Entonces  $s'$  es una subsucesión de  $v$ , donde cada elemento es múltiplo del anterior. No sabemos si  $s' = []$ , pero si no lo es, termina en algún  $v_j$ , con  $j < i$ ,  $v_j \mid v_i$ . Luego, si  $s' \neq []$ , entonces  $s'$  pertenece a la unión disjunta de  $S(j)$ , para algún  $0 \leq j < i$ . Si  $s' = []$ , entonces  $|s| = |s'| + 1 = 0 + 1 = 1$ . Ahora razonamos:

$$\begin{aligned}
g(i) &= \max\{|s| \mid s \in S(i)\} \\
&= \max\{\{1 + |s'| \mid s' \in S(j), 0 \leq j < i, v_j \mid v_i\} \cup \{1 + |s'| \mid s' \in \{\}\}\} \\
&= 1 + \max\{\{|s'| \mid s' \in S(j), 0 \leq j < i, v_j \mid v_i\} \cup \{|s'| \mid s' \in \{\}\}\} \\
&= 1 + \max\{\{|s'| \mid s' \in S(j), 0 \leq j < i, v_j \mid v_i\} \cup \{0\}\} \\
&= 1 + \max\{\{\max\{|s'| \mid s' \in S(j)\} \mid 0 \leq j < i, v_j \mid v_i\} \cup \{0\}\}, \text{ pues } S(j) \neq \emptyset \forall j \\
&= 1 + \max\{\{g(j) \mid 0 \leq j < i, v_j \mid v_i\} \cup \{0\}\} \\
&= 1 + \max\{\{f(j) \mid 0 \leq j < i, v_j \mid v_i\} \cup \{0\}\}, \text{ usando } P(j), \text{ pues } j < i \\
&= 1 + \max_0\{\{f(j) \mid 0 \leq j < i, v_j \mid v_i\}\} \\
&= f(i)
\end{aligned}$$

Luego vale  $P(i)$ .

3. Probemos ahora que, sabiendo que  $f$  tiene la semántica que dijimos,  $A$  es la respuesta al enunciado. Toda forma de borrar  $k$  elementos, dejando una subsecuencia de  $n - k$  elementos sin borrar, tal que cada uno es múltiplo del anterior, es lo mismo que encontrar esa subsecuencia de  $n - k$  elementos, y dejar sólo esos. Luego, la manera que borre *menos* elementos (que minimice  $k$ ), es la manera que encuentre la subsecuencia *más* larga (maximice  $n - k$ ).

Luego, podemos enfocarnos en encontrar la subsecuencia más larga donde cada elemento es múltiplo del anterior. El problema nos pide devolver cuántos elementos borramos, y eso es  $n - k$ , con  $k$  la longitud de tal secuencia.

Toda tal subsecuencia termina en alguna posición  $i$ . Luego, si tomamos  $k = \max_{0 \leq i < n} f(i)$ , sabiendo la semántica de  $f$ , habremos encontrado esa secuencia. Finalmente, al devolver  $A = n - k$ , estamos correctamente respondiendo el enunciado.

□

El código, en Python:

```

def F(v):
    n = len(v)
    def f(i):
        return 1 + max((f(j) for j in range(i)
                        if v[j] != 0 and v[i] % v[j] == 0),
                       default=0)
    return n - max(f(i) for i in range(n))

```

Este algoritmo va a computar muchas veces cada valor de  $f$ . En el peor caso, donde tenemos  $v = (1, 1, \dots, 1)$ , cada vez que llamemos a  $f(i)$  vamos a llamar a  $f(j)$  para todo  $0 \leq j < i$ . Si  $T$  es el número de operaciones que hace, entonces  $T(i) = O(n) + \sum_{j=0}^{i-1} T(j)$ . Esto termina siendo  $T \in \Theta(2^n)$ .

Vemos que se cumplen las dos condiciones para usar programación dinámica:

1. Subestructura óptima. Para resolver  $f(i)$ , basta con resolver los problemas más pequeños que  $i$ , en particular, con calcular  $f(j)$  para todo  $0 \leq j < i$ .
2. Subproblemas compartidos. A pesar de que hay sólo  $n$  posibles valores de  $f$ , la solución recursiva simple mira  $2^n$  subproblemas en el peor caso (donde  $v$  es todos unos, esto sale de resolver la recurrencia  $T(0) = 1; T(n) = 1 + \sum_{i=0}^{n-1} T(i)$ ). Luego hay muchos subproblemas compartidos, y programación dinámica probablemente haga más rápido nuestro algoritmo.

La manera mecánica de usar programación top-down es agregar un cache. Esto se convierte en:

```
def F(v):
    n = len(v)
    def f(i, cache={}):
        if i not in cache:
            cache[i] = 1 + max((f(j) for j in range(i)
                                if v[j] != 0 and v[i] % v[j] == 0),
                                default=0)
        return cache[i]
    return n - max(f(i) for i in range(n))
```

Es difícil analizar la complejidad temporal de este algoritmo, por estar mutando estado (cache), y depender su complejidad temporal del estado de cache en cada llamada. Podemos, entonces, usar programación dinámica bottom-up, llenando el cache en el mismo orden que se llenaría normalmente, pero a mano.

```
def F(v):
    n = len(v)
    dp = [1 for _ in range(n)]
    for i in range(1, n):
        dp[i] = 1 + max((dp[j] for j in range(i)
                           if v[j] != 0 and v[i] % v[j] == 0),
                           default=0)
    return n - max(dp[i] for i in range(n))
```

Esto nos deja entender el comportamiento asintótico muy fácilmente. El primer ciclo hace  $\Theta(n^2)$  operaciones en todos los casos, y el segundo  $\Theta(n)$  operaciones. Luego el algoritmo entero hace  $\Theta(n^2)$  operaciones en el peor caso. Finalmente, la complejidad temporal del algoritmo es  $\Theta(n)$ , que es el costo de guardar la tabla  $dp$ , más variables auxiliares, que cuestan  $O(1)$  cada una.

## 6.6 Greedy

Para muchos problemas, una solución greedy es fácil de imaginar. Por ejemplo, en el problema de la mochila, podemos pensar en «siempre tomo el objeto más valioso», o «siempre tomo el objeto más liviano». Lo difícil está en probar su correctitud.

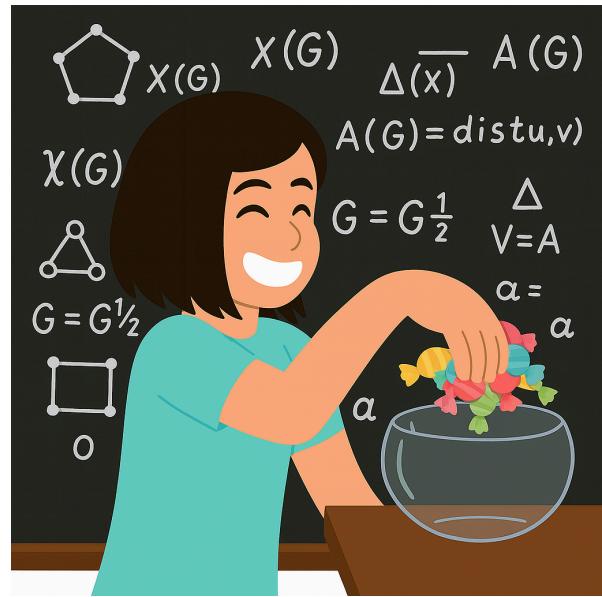
Hay varias maneras de probar la correctitud de los algoritmos greedy. En general, van a tener mucha flexibilidad para probar este tipo de algoritmos correctos. Tres formas clásicas son:

- Argumentos de intercambio, donde tomamos una solución óptima mejor a la nuestra lo más parecida posible a la nuestra, y hacemos un intercambio en esa solución que la hace o mejor que lo que era, o más parecida a la nuestra.
- Argumentos de liderazgo («greedy stays ahead»), donde definimos una métrica para sub-soluciones, y una noción de «longitud» para las mismas. Luego argumentamos que nuestra sub-solución siempre mejor, en esta métrica, a cualquier otra sub-solución de la misma «longitud».
- Argumentos de completación, donde argumentamos que en todo momento nuestra sub-solución puede ser completada a una solución óptima.

Vamos a mostrarles algunos ejemplos de cada una.

### Ejercicio 6.6.1

Queremos devolver el vuelto a un cliente, y tenemos monedas de 1, 5, 10, y 25 centavos. Diseñar un algoritmo greedy que resuelva el problema. Demostrar su correctitud, y su comportamiento asintótico.



*Demostración.* Un algoritmo como el que nos piden es el siguiente, donde  $n$  es el número de centavos que queremos devolver:

```

1: procedure GREEDYCHANGE( $n \in \mathbb{N}$ )
2:    $C \leftarrow [1, 5, 10, 25]$ 
3:    $v \leftarrow []$ 
4:   while  $n > 0$  do
5:      $c \leftarrow \max_{c \in C} \{c \mid c \leq n\}$ 
6:      $n \leftarrow n - c$ 
7:      $v \leftarrow v + [c]$ 
8:   end
9:   return  $v$ 
10: end
```

La semántica que le vamos a asignar a GREEDYCHANGE es que  $\text{GREEDYCHANGE}(n)$  devuelve una lista de denominaciones de monedas, tal que la suma de las denominaciones es  $n$ , y el número de monedas es mínimo entre todas las formas de sumar  $n$  con esas denominaciones.

Primero, veamos que GREEDYCHANGE termina. En cada iteración,  $n$  decrece en como mínimo 1, pues  $1 \leq n$  si entramos al ciclo, puesto que la guarda es  $n > 0$ , es decir  $n \geq 1$ .  $n$  nunca se hace negativo, porque siempre seleccionamos un  $c$  tal que  $c \leq n$ , y luego  $n - c \geq 0$ . Luego,

como en cada iteración decrece, vuelve al ciclo cada vez que  $n$  es positivo, y  $n$  nunca se hace negativo, sabemos que al final del ciclo,  $n = 0$ . Como cada vez que restamos  $c$  a  $n$ , agregamos  $c$  a  $v$ , vemos que  $\sum_{c \in v} c = n$ , y luego  $v$  es una forma de devolver el vuelto de  $n$  centavos. Asimismo, vemos que como  $n$  decrece en cada iteración, y  $c$  siempre es la máxima denominación menor a  $n$ , entonces  $c$  no puede crecer de una iteración a la otra, y luego  $v$  es llenado en orden de mayor a menor denominación.

Ahora veamos que GREEDYCHANGE devuelve una lista  $v$  de mínima longitud, tal que  $v$  contiene sólo elementos de  $C$ , y  $\sum_{c \in v} c = n$ . Vamos a hacer esto mediante un **argumento de intercambio**. Supongamos que existe una forma  $w$  de devolver el vuelto de  $n$  centavos, con menos monedas que  $v$ . De todas las posibles  $w$  de mínima longitud, tomemos cualquier que, al ser ordenada de forma no-creciente, tenga el máximo número de elementos en común con  $v$ . Es decir,  $w$  maximiza  $\max\{i \mid 0 \leq i \leq \min(|w|, |v|), \forall 0 \leq j < i, w_j = v_j\}$ , donde estamos ordenando  $w$  de forma no-creciente, entre todas las soluciones óptimas.

Sea  $i$  ese número. Entonces sabemos que antes de  $i$ ,  $w$  y  $v$  tienen los mismos elementos, mientras que  $v_i \neq w_i$ . Como GREEDYCHANGE produjo  $v$  de forma no-creciente, al elegir  $v_i$ , teníamos que  $v_i = \max_{c \in C} \{c \mid c \leq n'\}$ , con  $n' < n$  el cambio que hace falta hacer todavía. Entonces,  $n' = n - \sum_{j=1}^{i-1} v_j = \sum_{j=1}^{i-1} w_j$ , porque dijimos que para todos esos índices  $j$ ,  $v_j = w_j$ . Luego, como  $n = \sum_{j=1}^{|w|} w_j \geq \sum_{j=1}^i w_j = n - n' + w_i$ , o vemos que  $0 \geq -n' + w_i$ , o también,  $w_i \leq n'$ . Cómo elegimos  $v_i$  como el máximo  $c \in C$  tal que  $c \leq n'$ , y  $w_i \in C$ , sabemos que  $w_i < v_i$ .

Vamos a querer obtener, en cada caso, una forma de obtener una solución de menor largo que  $w$ , o de igual largo pero que tiene un prefijo más largo en común con  $v$ . Como  $w$  era una forma óptima que tiene el prefijo más largo con común con  $v$ , esto es una contradicción. Por lo tanto  $w$  no puede existir, y luego  $v$  es una solución óptima.

$v_1$	$v_2$	$\dots$	$v_{i-1}$	$v_i$	$\dots$	$v_{ v }$
		:		¶		
$w_1$	$w_2$	$\dots$	$w_{i-1}$	$w_i$	$\dots$	$w_{ w }$

$\overbrace{\quad \quad \quad \quad \quad \quad \quad}^{\sum_{j \geq i} v_j = n'}$

$\overbrace{\quad \quad \quad \quad \quad \quad \quad}^{\sum_{j \geq i} w_j = n'}$

Partimos en casos:

- Si  $v_i = 25$ , entonces  $n' \geq 25$ , y  $w_i < 25$ . Como ordenamos  $w$  de forma no-creciente, y el resto de  $w$  tiene que sumar  $n' \geq 25$ , consideremos qué monedas está usando  $w$ , a partir de  $w_i$ . Llamemos a estas monedas  $z = [w_j \mid |w| \geq j \geq i]$ .  $z$  va a ser de la forma  $z = \left[ \underbrace{10, \dots, 10}_{a \text{ dieces}}, \underbrace{5, \dots, 5}_{b \text{ cincos}}, \underbrace{1, \dots, 1}_{c \text{ unos}} \right]$ , con  $10a + 5b + c = n'$ .
  - Si  $a \geq 3$ , entonces podemos reemplazar tres 10s por [25, 5], obteniendo una solución más corta que  $w$ .
  - Si  $a = 2, b \geq 1$ , entonces podemos reemplazar dos 10s y un 5 por [25]. Si  $b = 0$ , podemos reemplazar dos 10s y cinco 1s por [25].

3. Si  $a = 1, b \geq 3$ , podemos reemplazar un 10 y tres 5s por [25]. Si  $b = 2$ , podemos reemplazar un 10, dos 5s y cinco 1s por [25]. Si  $b = 1$ , podemos reemplazar un 10, un 5 y diez 1s por [25]. Si  $b = 0$ , podemos reemplazar un 10 y quince 1s por [25].
4. Si  $a = 0$ , vemos que cualquier combinación no-creciente de 5s y 1s que sume  $n' \geq 25$  va a tener un prefijo que sume 25, y podemos ahí reemplazar ese prefijo monedas por [25]. En todos los casos, obteniendo una solución más corta que  $w$ .
2. Si  $v_i = 10$ , entonces  $n' \geq 10$ , y como  $w_i < v_i$ , entonces con la misma construcción que arriba, obtenemos  $z = \begin{bmatrix} 5, \dots, 5, 1, \dots, 1 \\ \underbrace{\phantom{5}}_{a \text{ cincos}} \quad \underbrace{\phantom{1}}_{b \text{ unos}} \end{bmatrix}$ , con  $5a + b = n' \geq 10$ .
  1. Si  $a \geq 2$ , entonces podemos reemplazar dos 5s por [10].
  2. Si  $a = 1$ , entonces podemos reemplazar un 5 y cinco 1s por [10].
  3. Si  $a = 0$ , entonces podemos reemplazar diez 1s por [10].
3. Si  $v_i = 5$ , entonces  $n' \geq 5$ , y como  $w_i < v_i$ , entonces con la misma construcción que arriba, obtenemos  $z = \begin{bmatrix} 1, \dots, 1 \\ \underbrace{\phantom{1}}_{a \text{ unos}} \end{bmatrix}$ , con  $a = n' \geq 5$ . Podemos entonces reemplazar cinco 1s por [5].
4. No podemos tener  $v_i = 1$ , porque no es posible tener  $w_i \in C, w_i < 1$ .

En todos los casos, obtenemos una solución más corta que  $w$ , lo cual contradice que  $w$  era una solución óptima. Luego, no puede existir tal  $w$ , y por lo tanto  $v$  es una solución óptima.  $\square$

### Ejercicio 6.6.2

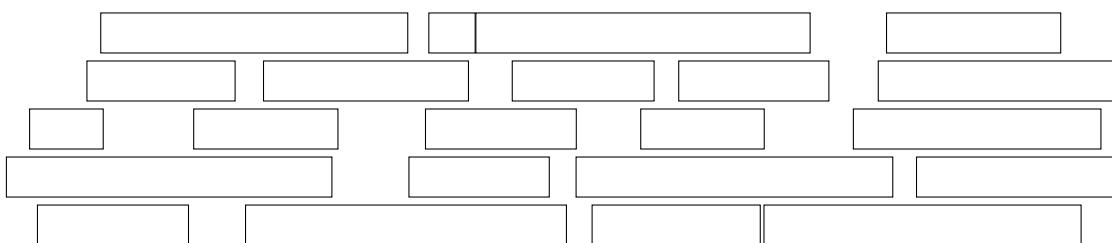
Tenemos un aula a nuestra disposición, y  $n$  clases que se pueden dar en ese aula. La  $i$ -ésima clase empieza a la hora  $s_i$ , y termina a la hora  $f_i$ . Dos clases no se pueden dar al mismo tiempo en ese aula. Queremos saber cual es el máximo número de clases que se pueden dictar en ese aula.

Diseñar un algoritmo greedy que resuelva este ejercicio. Probar que es correcto y probar su complejidad temporal y espacial asintótica.



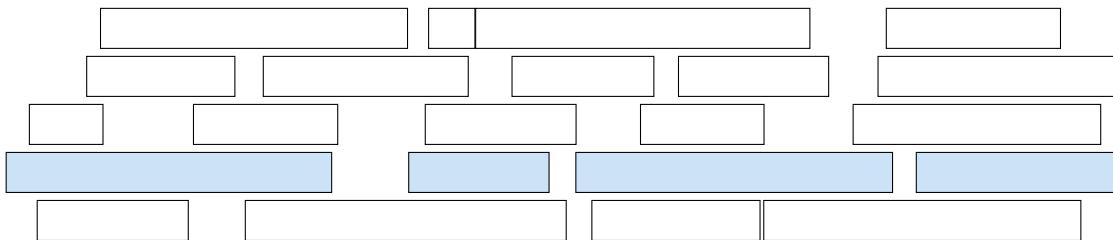
Cuando encontramos un problema nuevo y queremos ver si existe una estrategia greedy para resolverlo, lo primero que tenemos que hacer es intentar definir una estructura de «subproblemas», y adivinar cuál va a ser una estrategia de selección local (es decir, que toma decisiones en cada subproblema, sin replantearlas si «sale mal»).

Sirve entonces plantearse ejemplos, y ver cómo nuestras ideas funcionan o no.

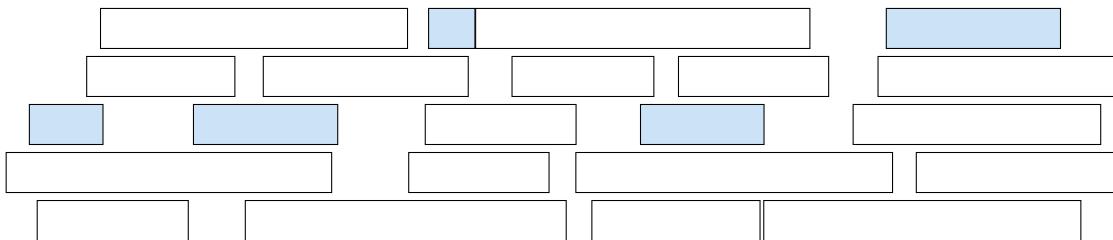


Pensemos en un par de estrategias para este ejemplo.

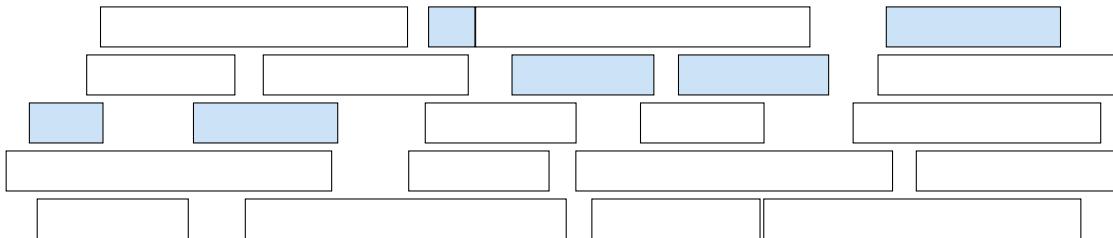
1. Podemos ordenar las materias por hora de comienzo, y agarrar la materia que empieze lo antes posible, que no tenga conflictos con materias ya seleccionadas. Esto nos da un subconjunto de tamaño 4.



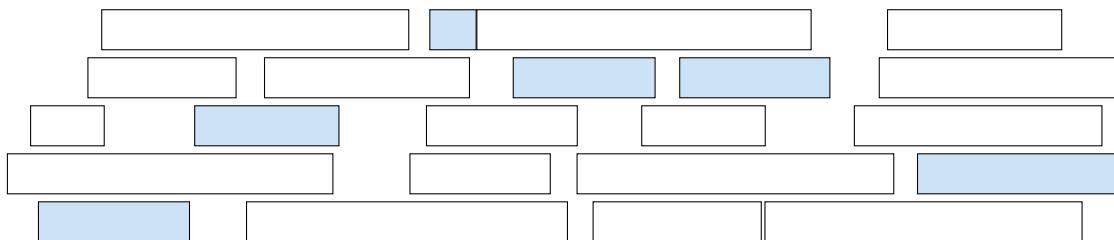
2. Podemos ordenar las materias por duración, eligiendo las más cortas que no causen conflictos con materias ya seleccionadas. Esto nos da un subconjunto de tamaño 5.



3. Podemos ordenar las materias por hora de finalización, y agarrar la materia que termine lo antes posible, que no tenga conflictos con materias ya seleccionadas. Esto nos da un subconjunto de tamaño 6.



Si buscamos a fuerza bruta, una solución óptima para este ejemplo también tiene tamaño 6.



Esto nos sugiere intentar probar que la estrategia de ordenar por momento de finalización, creciente, puede ser buena.

*Solución.* Proponemos el siguiente algoritmo.

```

1: procedure GREEDYCOURSES( $M = [(s_1, f_1), \dots, (s_n, f_n)] : \text{List}[\mathbb{N} \times \mathbb{N}]$ )
2:    $A \leftarrow []$ 
3:    $I \leftarrow \text{sort}([1, \dots, n], \text{key: } \lambda i. f_i)$ 
4:    $c \leftarrow 0$ 
```

```

5:   for  $i \in I$  do
6:     if  $s_i \geq c$  then
7:        $c \leftarrow f_i$ 
8:        $A \leftarrow A + [i]$ 
9:     end
10:   end
11:   return  $A$ 
12: end

```

*Demostración.* Para esta demostración les vamos a mostrar un argumento de liderazgo («greedy stays ahead»). Para esto tenemos que definir una estructura de «subproblemas» que nuestro algoritmo greedy resuelve en orden. Luego tenemos que dar una estructura de sub-solución a cualquier solución. Finalmente, vamos a dar una noción de «valor» para las sub-soluciones, y probamos que la  $i$ -ésima sub-solución de nuestro algoritmo tiene «mejor» valor que la  $i$ -ésima sub-solución de cualquier solución. «Mejor» puede lo vamos a definir a veces como «menor», y a veces como «mayor», dependiendo del problema.

Sean  $A = \{i_1, \dots, i_k\}$  los índices de las materias que devuelve nuestro algoritmo, en el orden en que fueron agregados. Sea  $O = \{j_1, \dots, j_m\}$  los índices de las materias en una solución óptima, ordenados por momento de finalización. Sea  $F(X) = \max_{x \in X} f_x$  el máximo tiempo de finalización entre todas las materias en un conjunto  $X$ . Notar que por cómo definimos  $A$  y  $O$ , tenemos que  $F(\{i_1, \dots, i_r\}) = F(\{i_r\}) = f_{i_r}$  y  $F(\{j_1, \dots, j_r\}) = F(\{j_r\}) = f_{j_r}$ , para todo  $r$ .

Las sub-soluciones de nuestro algoritmo son  $\{i_1, \dots, i_r\}$ , para cada  $r$ . Las sub-soluciones de una solución óptima son  $\{j_1, \dots, j_r\}$ , para cada  $r$ . La noción de valor que vamos a usar es  $F$ , el máximo tiempo de finalización entre las materias en la sub-solución.

Vamos a mostrar, entonces, que para todo  $1 \leq r \leq k$ , tenemos  $P(r) : F(\{i_1, \dots, i_r\}) \leq F(\{j_1, \dots, j_r\})$ . Probemos  $P$ .

1. Caso base,  $P(1)$ . Como  $A$  eligió inicialmente la materia que antes termina en  $M$ , sabemos que  $f_{i_1} \leq f_r \forall r$ . En particular, vemos que  $f_{i_1} \leq f_{j_1}$ .
2. Paso inductivo. Sea  $k \geq t \geq 2 \in \mathbb{N}$ , sabemos que  $P(t-1)$ , queremos probar  $P(t)$ .  $P(t-1)$  nos dice que  $f_{i_{t-1}} \leq f_{j_{t-1}}$ . Todas las materias que pueden seguirle a  $j_{t-1}$  en  $O$  empiezan después de  $f_{j_{t-1}}$ , luego como eso viene después que  $f_{i_{t-1}}$ , también son candidatas para agregar a  $A$ . Luego, si  $i_t$  es la siguiente materia que agrega  $A$ , entonces  $f_{i_t} \leq f_{j_t}$ , pues  $A$  toma la que antes termine, de todas las candidatas. Esto muestra que  $P(t)$  es cierto.

Eso demuestra que vale  $P(r)$  para todo  $1 \leq r \leq k$ . En particular, vale para  $r = k$ , y tenemos que  $F(A) = F(\{i_1, \dots, i_k\}) \leq F(\{j_1, \dots, j_k\})$ .

Si  $m > k$ , entonces existe  $j_{k+1}$ , una materia que la solución óptima  $O$  eligió, que no está en  $A$ . Como ordenamos  $O$ , debemos tener que  $s_{j_{k+1}} \geq f_{j_k}$ . Por  $P(k)$ , sabemos que  $f_{j_k} \geq f_{i_k}$ . Por lo tanto,  $j_{k+1}$  es una materia que empieza después de que termine la última materia que agregó  $A$ . Es más, como  $k$  fue la última materia que agregó  $A$ ,  $j_{k+1}$  empieza después que *todas* las materias en  $A$ . Entonces, al iterar por  $i = j_{k+1}$ ,  $A$  la

hubiera agregado, y no lo hizo. Esto no puede suceder, y por lo tanto  $m \leq k$ . Luego,  $A$  es una solución óptima.  $\square$

### Ejercicio 6.6.3

Tenemos  $n$  items, cada uno con valor  $v_i$  y peso  $p_i$ . Tenemos una mochila que puede aguantar un peso máximo  $P$ . Para este problema, podemos tomar pedazos fraccionarios de cada objeto. Por ejemplo, si tenemos una manzana de 300 gramos y valor 6, podemos tomar un tercio de la manzana, de peso 100 gramos y valor 2.

Queremos maximizar el valor de los objetos que llevamos en la mochila, sin exceder el peso máximo  $P$ . Podemos asumir que  $\frac{v_i}{p_i} \neq \frac{v_j}{p_j} \forall i, j$ , es decir, que no hay dos objetos con el mismo valor por unidad de peso.

Diseñar un algoritmo greedy que resuelva este ejercicio. Probar que es correcto y probar su complejidad temporal y espacial asintótica.



*Solución.* Primero formalicemos un poco la consigna. Lo que queremos hacer es encontrar un vector  $x = (x_1, \dots, x_n)$ , donde  $0 \leq x_i \leq 1$  para cada  $i$ , que maximice la suma

$$\sum_{i=1}^n v_i x_i$$

sujetos a

$$\sum_{i=1}^n p_i x_i \leq P$$

Una estrategia greedy para resolver esto va a ser repetidamente elegir algún elemento, mientras que quepa en la mochila. Si sólo cabe una fracción del objeto, metemos sólo esa fracción en la mochila. Veamos un ejemplo, con peso máximo  $P = 40$ .

Pesos	4	50	14	52	13	1	45	3
Valores	2	42	4	24	2	3	42	10

Algunas ideas para cómo elegir los objetos:

1. Podemos elegir el objeto con mayor valor, y meterlo en la mochila. Si no cabe entero, metemos la fracción que quepa. En este problema, elegiríamos ordenaríamos los objetos como  $[2, 7, 4, 8, 3, 6, 1, 5]$ . Luego intentaríamos meter el 2do objeto, que tiene peso 50 y valor 42. Como no cabe entero puesto que su peso es mayor a  $P = 40$ , meteríamos  $\frac{40}{50}$  de este objeto, que aporta un valor de  $\frac{40}{50} \times 42 = 33.6$ .
2. Podemos elegir el objeto con menor peso, y meterlo en la mochila. Si no cabe entero, metemos la fracción que quepa. En este problema, ordenaríamos los objetos como  $[6, 8, 1, 5, 3, 7, 2, 4]$ . Podemos meter los objetos 6, 8, 1, 5, y 3, enteros, obteniendo un valor de  $3 + 10 + 2 + 2 + 4 = 21$ . Para cuando vemos el 7mo objeto, nuestra mochila ya tiene un

peso de  $1 + 3 + 4 + 13 + 14 = 35$ , luego sólo podemos meter  $\frac{5}{45}$  de este objeto, que nos agrega un valor de  $\frac{5}{45} \times 42 = 4.67$ , obteniendo un valor final de  $21 + 4.67 = 25.67$ .

3. Podemos elegir el objeto con mayor valor por unidad de peso ( $\frac{v_i}{w_i}$ ), y meterlo en la mochila. Si no cabe entero, metemos la fracción que quepa. En este caso, ordenaríamos los objetos como  $[8, 6, 7, 2, 1, 4, 3, 5]$ . Agregamos los objetos 8 y 6, obteniendo un valor de 13, y un peso de 4. Luego vemos el 7mo objeto, y como tenemos 36 de peso para llenar, y el objeto pesa 45, sólo podemos meter  $\frac{36}{45}$  de este objeto, que nos agrega un valor de  $\frac{36}{45} \times 42 = 33.6$ , obteniendo un valor final de  $13 + 33.6 = 46.6$ .

De estas tres ideas, la última fue la que mejor valor nos dio. Intentemos, entonces, ese algoritmo.

```

1: procedure FRACTIONALKNAPSACK( $P \in \mathbb{N}$ ,  $p : \mathbb{N}^n$ ,  $v : \mathbb{N}^n$ )
2:    $A \leftarrow []$ 
3:    $I \leftarrow \text{sort}([1, \dots, n], \text{key: } \lambda i. \frac{v_i}{p_i})$ 
4:    $c \leftarrow 0$ 
5:    $v \leftarrow 0$ 
6:   for  $i \in I$  do
7:     if  $c + p_i \leq P$  then
8:        $c \leftarrow c + p_i$ 
9:        $v \leftarrow v + v_i$ 
10:       $A \leftarrow A + [(i, 1.0)]$ 
11:    else
12:       $x \leftarrow P - \frac{c}{p_i}$ 
13:       $v \leftarrow v + x * v_i$ 
14:       $A \leftarrow A + [(i, x)]$ 
15:      BREAK
16:    end
17:   end
18:   return  $(A, v)$ 
19: end
```

*Demostración.* Para este problema vamos a usar un argumento de intercambio. Vamos a considerar una solución óptima que es distinta y mejor a la nuestra, y vamos a ver que podemos mejorar esta solución óptima, lo cual nos diría que no puede existir.

### Nota

Notemos que en general *no* vamos a poder probar que nuestra solución es la *única* solución óptima. Puede haber muchas, y nuestro algoritmo va a elegir una sola. No vamos a poder probar, entonces, que no existe una solución óptima distinta que la nuestra. Lo que vamos a poder probar es que no hay una solución *mejor* que la nuestra.

Ordenemos los objetos por su valor por unidad de peso,  $\frac{v_i}{p_i}$ , de forma no-creciente. Luego, tenemos que  $\frac{v_i}{p_i} \geq \frac{v_j}{p_j}$  para todo  $i \leq j$ .

Sea  $A = [x_1, \dots, x_n]$  las fracciones de cada elemento que eligió nuestro algoritmo, y sea  $O = [y_1, \dots, y_n]$  una solución óptima. Por cómo ordenamos los objetos en esta demostración, nuestro algoritmo eligió primero un valor para  $x_1$ , luego un valor para

$x_2$ , etcétera. Notemos que  $0 \leq x_i \leq 1$  y  $0 \leq y_i \leq 1$  para todo  $i$ , porque estamos eligiendo fracciones de los objetos, y no podemos elegir más de un objeto entero.

Supongamos que  $A$  no es óptima. Entonces  $\sum_{i=1}^n v_i x_i < \sum_{i=1}^n v_i y_i$ . Sea  $i$  el primer índice donde  $x_i \neq y_i$ . Luego,  $\sum_{j=1}^{i-1} x_j p_j = \sum_{j=1}^{i-1} y_j p_j$ . Al momento de elegir  $x_i$ , nuestro algoritmo podía usar un peso máximo de  $P' = P - \sum_{j=1}^{i-1} x_j p_j$ . Nuestro algoritmo elige  $x_i = \frac{P'}{p_i}$ .

Asimismo, como  $O$  es una solución, tenemos que  $\sum_{j=1}^n y_j p_j \leq P$ . Restando  $\sum_{j=1}^{i-1} x_j p_j$  de cada lado, tenemos que  $\sum_{j=i}^n y_j p_j \leq P'$ . Como todos los pesos son positivos, así como también los  $y_j$ , esto implica que  $y_i p_i \leq P'$ . Esto es lo mismo que decir  $y_i \leq \frac{P'}{p_i} = x_i$ . Como  $y_i \leq x_i$  y  $y_i \neq x_i$ , sabemos que  $y_i < x_i$ . Como  $O$  es óptima, y tiene más peso disponible en la mochila, tiene que existir algún  $j > i$  tal que  $y_j > x_j$ . Si no fuera así, entonces  $O$  no es óptima, porque  $A$  tiene más valor que  $O$ .

Sea  $\varepsilon = \min\left(1 - y_i, y_j \frac{p_j}{p_i}\right)$ . Vemos que, como  $\varepsilon \leq 1 - y_i$ , entonces  $y_i + \varepsilon \leq 1$ . También, como  $\varepsilon \leq y_j \frac{p_j}{p_i}$ , entonces  $y_j - \varepsilon \frac{p_i}{p_j} \geq 0$ . Por lo tanto, podemos definir una nueva solución  $z$ , cuyos valores están siempre entre 0 y 1, como:

$$z_r = \begin{cases} y_i + \varepsilon, & \text{si } r = i \\ y_j - \varepsilon \frac{p_i}{p_j}, & \text{si } r = j \\ y_r, & \text{si } r \neq i \wedge r \neq j \end{cases}$$

Veamos que el peso total de  $z$  es el mismo que el de  $O$ .

$$\begin{aligned} \sum_{r=1}^n z_r p_r &= \sum_{r=1}^n y_r p_r - y_i p_i - y_j p_j + (y_i + \varepsilon) p_i + \left(y_j - \varepsilon \frac{p_i}{p_j}\right) p_j \\ &= \sum_{r=1}^n y_r p_r - y_i p_i - y_j p_j + y_i p_i + \varepsilon p_i + y_j p_j - \varepsilon p_i \\ &= \sum_{r=1}^n y_r p_r \end{aligned}$$

Y luego  $z$  es una solución válida. Sin embargo, veamos que el valor de  $z$  es mayor al de  $O$ :

$$\begin{aligned} \sum_{r=1}^n z_r v_r &= \sum_{r=1}^n y_r v_r - y_i v_i - y_j v_j + (y_i + \varepsilon) v_i + \left(y_j - \varepsilon \frac{p_i}{p_j}\right) v_j \\ &= \sum_{r=1}^n y_r v_r - y_i v_i - y_j v_j + y_i v_i + \varepsilon v_i + y_j v_j - \varepsilon \frac{p_i}{p_j} v_j \\ &= \sum_{r=1}^n y_r v_r + \varepsilon \left(v_i - \frac{p_i}{p_j} v_j\right) \end{aligned}$$

Ahora bien:

$$v_i - \frac{p_i}{p_j} v_j \geq 0$$

$$\frac{v_i}{p_i} \geq \frac{v_j}{p_j}$$

Pues así ordenamos los objetos. El problema nos dice que no hay dos objetos con el mismo valor por unidad de peso, entonces ese  $\geq$  es en realidad un  $>$ , pues  $j > i$ , y luego son objetos distintos.

Por otro lado, sabemos que  $y_i < x_i \leq 1$ , y luego  $y_i < 1$ , y entonces  $1 - y_i > 0$ .

Asimismo, como  $y_j > x_j \geq 0$ , entonces  $y_j > 0$ , y luego  $y_j \frac{p_j}{p_i} > 0$ . Luego, al ser  $\varepsilon = \min\left(1 - y_i, y_j \frac{p_j}{p_i}\right) > 0$ , tenemos que  $\varepsilon \left(v_i - \frac{p_i}{p_j} v_j\right) > 0$ .

Luego,  $\sum_{r=1}^n z_r v_r > \sum_{r=1}^n y_r v_r$ , lo cual contradice que  $O$  es óptima. Luego, no puede existir una tal solución óptima,  $O$ , que es estrictamente mejor que la nuestra,  $A$ .  $\square$

#### Ejercicio 6.6.4

Tenemos dos conjuntos de personas y para cada persona sabemos su habilidad de baile. Queremos armar la máxima cantidad de parejas de baile, sabiendo que para cada pareja debemos elegir exactamente una persona de cada conjunto de modo que la diferencia de habilidad sea menor o igual a 1 (en modulo). Ademas, cada persona puede pertenecer a lo sumo a una pareja de baile. Por ejemplo, si tenemos un multiconjunto con habilidades  $\{1, 2, 4, 6\}$  y otro con  $\{1, 5, 5, 7, 9\}$ , la maxima cantidad de parejas es 3. Si los multiconjuntos de habilidades son  $\{1, 1, 1, 1, 1\}$  y  $\{1, 2, 3\}$ , la máxima cantidad es 2.

Diseñar un algoritmo greedy que resuelva este ejercicio. Probar que es correcto y probar su complejidad temporal y espacial asintótica.

#### Demostración.

Veamos un algoritmo recursivo simple que resuelve el problema.

```
def F(G: [int], B: [int]) -> int:
    return f(sorted(G), sorted(B))

def f(G: [int], B: [int]) -> int:
    if not G or not B:
        return 0
    g = G[0]
    b = B[0]
    if b < g - 1:
        return f(G, B[1:])
    if g < b - 1:
        return f(G[1:], B)
    return 1 + f(G[1:], B[1:])
```

Queremos ver que  $F(G, B)$  devuelve el máximo número de parejas que se pueden armar con los multiconjuntos (en este programa, listas) de habilidades dados por  $G$  y  $B$ . Vamos a probar que  $f(G', B')$  devuelve la respuesta que  $F(G, B)$  tiene que dar, asumiendo que  $G'$  y  $B'$  tienen los mismos elementos que  $G$  y  $B$  respectivamente, sólo permutando su orden, lo cual no cambia la solución esperada, dado que estamos cambiando la representación del multiconjunto, pero no sus elementos.

Lo vamos a hacer por inducción. Como en  $f$  estamos sacando siempre un elemento de  $B$  o de  $G$  (y a veces de ambos), vamos a definir:

$$\text{tamaño}(G, B) = \text{len}(G) + \text{len}(B)$$

Esto nos va a ayudar porque nuestras llamadas recursivas siempre llaman a  $f$  con instancias de menor tamaño que la que recibe. Esto es lo que necesitamos para hacer inducción.

Definimos entonces:

$$P(k) : f(G, B) \text{ es correcta para todos los argumentos } (G, B) \\ \text{de tamaño}(G, B) \text{ a lo sumo } k.$$

### 1. Caso base

El caso base es  $P(0)$ . Tenemos que probar que  $f(G, B)$  es correcto para todos los argumentos de tamaño a lo sumo  $k = 0$ . Si  $\text{tamaño}(G, B) = 0$ , entonces  $\text{len}(G) + \text{len}(B) = 0$ , pero entonces como ambos  $\text{len}(G)$  y  $\text{len}(B)$  son enteros no-negativos, tenemos que  $\text{len}(G) = \text{len}(B) = 0$ . Luego,  $G = B = []$ . Entonces, no hay parejas posibles, porque una pareja tendría que tener un elemento de  $G$  y otro de  $B$ . Luego la respuesta correcta es  $0$ , y efectivamente nuestro programa devuelve  $0$ , en:

```
if not G or not B:
    return 0
```

### 2. Paso inductivo

Ahora probemos el paso inductivo. Asumo  $P(k)$ , y quiero probar  $P(k + 1)$ .

Me dan un par  $(G, B)$  con  $\text{tamaño}(G, B) = k + 1$ . Si uno de los dos está vacío (no pueden ambos estar vacíos porque  $\text{tamaño}(G, B) = k + 1 \geq 1$ ), no hay parejas posibles, luego la respuesta correcta es  $0$ , y el programa efectivamente devuelve eso

```
if not G or not B:
    return 0
```

Si ninguno está vacío, ambos tienen un primer elemento, llamémoslo  $g_0 \in G$  y  $b_0 \in B$ .

Como  $G$  y  $B$  están ordenados crecientes,  $g_0 \leq g \forall g \in G$ , y  $b_0 \leq b \forall b \in B$ .

Consideremos ahora  $b_0$  versus  $g_0$ .

- Si  $b_0 < g_0 - 1$ , entonces  $b_0$  no puede estar aparejado con  $g_0$ . Pero más aún, como  $g_0 \leq g \forall g \in G$ , todos los elementos de  $G$  son al menos tan grandes como  $g_0$ , y luego  $b_0$  no puede estar aparejado con nadie. Luego, toda solución a  $(G, B)$  es lo mismo que una solución a  $(G, B \setminus \{b_0\})$ . Como  $\text{tamaño}(G, B \setminus \{b_0\}) = k < k + 1$ , sabemos por inducción que  $f$  es correcta para ese caso, y luego nuestra  $f$  es también correcta para este caso de  $(G, B)$ , porque devolvemos exactamente  $f(G, B \setminus \{b_0\})$ :

```

if b < g - 1:
    return f(G, B[1:])

```

- Con un argumento similar, si  $g_0 < b_0 - 1$ ,  $g_0$  no puede estar aparejado con  $b_0$ , y  $b_0$  es menor o igual que todos los elementos en  $B$ , luego  $g_0$  no puede estar aparejado con nadie en  $B$ , y luego toda solución a  $(G, B)$  deja a  $g_0$  sin usar, y luego es idéntica a una solución a  $(G \setminus \{g_0\}, B)$ . Luego nuestra  $f$  es correcta para este caso de  $(G, B)$ , porque devolvemos exactamente  $f(G \setminus \{g_0\}, B)$ , y como tamaño( $G \setminus \{g_0\}, B$ ) =  $k < k + 1$ , por hipótesis inductiva  $f$  devuelve una solución óptima para ese caso:

```

if g < b - 1:
    return f(G[1:], B)

```

Si ninguna de las dos condiciones vale, entonces  $g_0 \geq b_0 - 1$ , y  $b_0 \geq g_0 - 1$ . Esto es lo mismo que  $b_0 - g_0 \leq 1$ , y  $g_0 - b_0 \leq 1$ , o lo que es lo mismo,  $|b_0 - g_0| \leq 1$ , es decir,  $b_0$  y  $g_0$  son compatibles. Notemos que en nuestro programa esta es la última rama, donde devolvemos  $1 + f(G[1:], B[1:])$ . Vamos a probar dos lemas:

- Si existe una solución óptima para  $(G, B)$  donde  $(g_0, b_0)$  están aparejados, entonces  $f$  es correcta para  $(G, B)$ .
- Existe una solución óptima para  $(G, B)$  donde  $(g_0, b_0)$  están aparejados.

Está claro que si probamos ambos lemas, probamos que  $f$  es correcta para  $(G, B)$ . Como  $(G, B)$  era cualquier argumento de tamaño  $k + 1$ , esto prueba que  $P(k) \Rightarrow P(k + 1)$ , que es lo que queríamos demostrar.

#### Lema 6.6.5

Queremos probar que «Si existe una solución óptima para  $(G, B)$  donde  $(g_0, b_0)$  están aparejados, entonces  $f$  es correcta para  $(G, B)$ .»



*Demostración.* Supongamos que existe una solución óptima para  $(G, B)$  donde  $(g_0, b_0)$  están aparejados. Sea  $S$  esa solución. Luego,  $S' = S \setminus \{(g_0, b_0)\}$  es una forma de aparejar a  $G \setminus \{g_0\}$  y  $B \setminus \{b_0\}$ .

Podemos decir algo más fuerte:  $S'$  es óptima para  $(G \setminus \{g_0\}, B \setminus \{b_0\})$ . Si no lo fuera, sea  $S^*$  una solución a  $(G \setminus \{g_0\}, B \setminus \{b_0\})$  que tiene más elementos que  $S'$ . Como  $S^*$  no menciona a  $g_0$  ni a  $b_0$ , podemos construir  $S^* \cup \{(g_0, b_0)\}$ , que tiene  $|S^*| + 1 > |S'| + 1 = |S|$  elementos, y es una solución a aparejar a  $G$  y  $B$ . Esto no puede pasar, porque  $S$  era óptima para  $(G, B)$ , no puedo tener a  $S^* \cup \{(g_0, b_0)\}$  más grande que  $S$ .

Luego,  $S'$  es óptima para  $(G \setminus \{g_0\}, B \setminus \{b_0\})$ . Por inducción,  $f$  es correcta para  $(G \setminus \{g_0\}, B \setminus \{b_0\})$ , porque tamaño( $G \setminus \{g_0\}, B \setminus \{b_0\}$ ) =  $k - 1 < k + 1$ . Luego  $|S'| = |f(G \setminus \{g_0\}, B \setminus \{b_0\})|$ , y luego  $|S| = 1 + |S'| = 1 + |f(G \setminus \{g_0\}, B \setminus \{b_0\})|$ . Luego  $f$  es correcta para  $(G, B)$ , dado que devolvemos exactamente eso:

```
return 1 + f(G[1:], B[1:])
```



### Lema 6.6.6

Queremos probar que «Existe una solución óptima para  $(G, B)$  donde  $(g_0, b_0)$  están aparejados.» Sabemos que  $g_0$  y  $b_0$  son compatibles.



*Demostración.* Sea  $S$  cualquier solución óptima a  $(G, B)$ . Sean  $S_G$  y  $S_B$  los elementos de  $G$  y  $B$  que *no* aparecen en  $S$ , respectivamente. Leér «single girls, single boys».

Partimos en casos:

- Si  $g_0 \in S_G$  y  $b_0 \in S_B$ , podríamos considerar  $S' = S \cup \{(g_0, b_0)\}$ , que tiene un elemento más que  $S$ , y es solución a  $(G, B)$ . Como  $S$  era solución óptima, esto no puede pasar.
- Si  $g_0 \in S_G$  pero  $b_0 \notin S_B$ , sabemos que existe una pareja  $(x, b_0) \in S$ . Luego, podemos considerar  $S' = (S \setminus \{(x, b_0)\}) \cup \{(g_0, b_0)\}$ . Esto deja sin aparejar a  $x$ , pero empareja a  $g_0$  y  $b_0$ , y como tiene el mismo número de elementos que  $S$ , vemos que  $S'$  es una solución óptima para  $(G, B)$  donde  $g_0$  y  $b_0$  están aparejados.
- Si  $b_0 \in S_B$  pero  $g_0 \notin S_G$ , sabemos que existe una pareja  $(g_0, y) \in S$ . Luego, podemos considerar  $S' = (S \setminus \{(g_0, y)\}) \cup \{(g_0, b_0)\}$ . Esto deja sin aparejar a  $y$ , pero empareja a  $g_0$  y  $b_0$ , y como tiene el mismo número de elementos que  $S$ , vemos que  $S'$  es una solución óptima para  $(G, B)$  donde  $g_0$  y  $b_0$  están aparejados.
- Si  $g_0 \notin S_G$  y  $b_0 \notin S_B$ , entonces ambos están aparejados. Hay dos opciones:
  - Si están aparejados el uno al otro, es decir,  $(g_0, b_0)$  está en  $S$ , ya está, conseguimos una solución óptima que contiene a  $(g_0, b_0)$ , y es  $S$ .
  - Si no, existen  $(x, b_0) \in S$  y  $(g_0, y) \in S$ . Quisieramos aparejar a  $b_0$  con  $g_0$ , y a  $x$  con  $y$ , y para eso tenemos que probar que  $x$  e  $y$  son compatibles. Como  $b_0$  era el elemento en  $B$  más chico, tenemos que  $y \geq b_0$ . De la misma manera sabemos que  $x \geq g_0$ . Veamos qué puede pasar:
    - Si  $b_0 = g_0$ , es decir, tienen la misma altura, entonces  $y \geq b_0 = g_0$ , y como en  $S$  están aparejados  $y$  con  $g_0$ ,  $y$  sólo puede ser  $g_0$  o  $g_0 + 1$ . Por el mismo motivo,  $x \geq g_0 = b_0$ , y en  $S$  están aparejados  $x$  con  $b_0$ , y entonces  $x$  sólo puede ser  $b_0$  o  $b_0 + 1$ . Como  $b_0$  y  $g_0$  son el mismo número, ambos  $x$  e  $y$  sólo pueden ser  $b_0$  o  $b_0 + 1$ , y luego  $x$  e  $y$  son compatibles.
    - Si  $b_0 = g_0 + 1$ , entonces  $y \geq b_0 = g_0 + 1$ , y como en  $S$  están aparejados  $y$  con  $g_0$ , tenemos que  $y$  es *exactamente*  $g_0 + 1$ . Luego tenemos que  $y = g_0 + 1$ .

$1 = b_0$ . Como  $x$  y  $b_0$  son compatibles, y  $b_0 = y$ , vemos que  $x$  e  $y$  son compatibles.

- Si  $g_0 = b_0 + 1$ , pasa lo mismo que en el caso anterior. Tenemos  $x \geq g_0 = b_0 + 1$ , y como  $x$  y  $b_0$  están aparejados en  $S$ , tenemos que  $x$  es *exáctamente*  $b_0 + 1$ . Luego tenemos que  $x = b_0 + 1 = g_0$ . Como  $y$  y  $g_0$  son compatibles, y  $g_0 = x$ , vemos que  $x$  e  $y$  son compatibles.

Como en todos los casos  $x$  e  $y$  son compatibles, podemos considerar  $S' = (S \setminus \{(x, b_0), (g_0, y)\}) \cup \{(g_0, b_0), (x, y)\}$ . Esta es una solución a  $(G, B)$  que apareja a  $g_0$  y  $b_0$ , y que tiene el mismo número de elementos que  $S$ , y por lo tanto también es óptima para  $(G, B)$ .

Vemos entonces que siempre existe una solución óptima para  $(G, B)$  donde  $g_0$  y  $b_0$  están aparejados.

□

□

### Ejercicio 6.6.7

Resolver el ejercicio anterior dando un algoritmo iterativo, en vez de recursivo.



*Demostración.* Veamos ahora una resolución iterativa del mismo ejercicio.

Nota: Se asume acá que los multiconjuntos están representados por listas no-decrescentes.

```
def f(G: [int], B: [int]) -> int:
    n = len(G)
    m = len(B)
    i = 0
    j = 0
    res = 0
    while i < n and j < m:
        g = G[i]
        b = B[j]
        if b < g - 1:
            j += 1
        elif g < b - 1:
            i += 1
        else:
            i += 1
            j += 1
            res += 1
    return res
```

Nuestro invariante va a ser que  $0 \leq i \leq n$ ,  $0 \leq j \leq m$ , y que existe un conjunto  $S$ , un conjunto  $T$ , y una solución óptima  $S^*$ , tal que:

1.  $S \subseteq G[0...i] \times B[0...j]$

2.  $T \subseteq G[i...n) \times B[j...m)$
3.  $S^* = S \sqcup T$ , con  $\sqcup$  siendo la unión disjunta
4.  $\text{res} = |S|$

Esto se puede entender como que  $S$  es «extensible» a una solución óptima  $S^*$ , usando sólo elementos que vienen no antes que  $i$  en  $G$  y  $j$  en  $B$ . Llamemos a esta noción « $(i, j)$ -extensible».

### El invariante vale inicialmente

Claramente vale el invariante antes de entrar al ciclo, dado que  $i = j = 0$ , y  $G[0 \dots i) = G[0 \dots 0) = []$ ,  $B[0 \dots j) = B[0 \dots 0) = []$ . Con ambos  $G$  y  $B$  vacíos, no se puede formar ninguna pareja, y luego definimos  $S = \emptyset$ , cuyo tamaño es exactamente  $\text{res} = 0$ . Asimismo, vemos que *toda* solución global  $S^*$ , es una extensión de  $\emptyset$ , agregándole parejas en  $G[i...n) \times B[j...m) = G[0...n) \times B[0...m) = G \times B$ . En particular, definimos  $T = S^*$ , y tenemos que  $S^* = \emptyset \sqcup T$ .

Los despiertos habrán notado que ese párrafo asume que *existe* una solución óptima. Un «para todo  $x$  en  $X$ , vale  $P(x)$ » sólo implica «existe  $x$  en  $X$  tal que  $P(x)$ » cuando  $X$  no es vacío. El conjunto de todos los emparejamientos posibles no es vacío (por ejemplo, podemos tomar el emparejamiento vacío), y es finito (está incluido en  $\mathcal{P}(G \times B)$ , el conjunto de partes de  $G \times B$ ). Luego tiene al menos un elemento de tamaño máximo, y luego *existe* al menos una solución óptima.

### El invariante es preservado por las iteraciones

Ahora veamos que si vale la guarda, y vale el invariante al entrar al cuerpo del ciclo, vale el invariante al finalizar el cuerpo del ciclo. Como vale la guarda, sabemos que  $0 \leq i < n$ , y  $0 \leq j < m$ . Vemos fácilmente que al final del cuerpo del loop sigue valiendo  $0 \leq i \leq n$ ,  $0 \leq j \leq m$ , porque sólo los aumentamos en uno en el cuerpo del loop, y antes eran  $\geq 0$  y  $< n$  y  $< m$ , respectivamente. Como los índices están en rango, tiene sentido referirse a  $g = G[i]$ , y  $b = B[j]$ . Partamos en los tres casos que parte el algoritmo:

1. Si  $b < g - 1$ , entonces  $b$  no es compatible con  $g$ , puesto que  $|g - b| = g - b > 1$ . Sabemos que  $G$  está ordenado crecientemente, luego para todo  $g' \in G[i...n)$ ,  $g' \geq g$ , y luego  $|g' - b| = g' - b \geq g - b > 1$ , y por lo tanto  $b$  es también incompatible con todos los  $g'$  que quedan. Luego, sea  $S^*$  la extensión de  $S$  que existe por el invariante. Sabemos que  $S^*$  se descompone como  $S^* = S \sqcup T$ , con  $S \subseteq G[0...i) \times B[0...j)$ , y  $T \subseteq G[i...n) \times B[j...m)$ . Como  $b = B[j]$ ,  $b \notin B[0...j)$ , y luego  $b$  no aparece emparejado en  $S$ . Vimos arriba que  $b$  no es compatible con nadie en  $G[i...n)$ , y luego  $b$  no puede estar en  $T$ , porque su pareja debería estar en  $G[i...n)$ . Luego, tenemos que

1.  $S \subseteq G[0...i) \times B[0...j + 1)$ , trivialmente, porque ya estaba incluído en  $G[0...i) \times B[0...j)$ .
2.  $T \subseteq G[i...n) \times B[j + 1...m)$ , porque sabíamos que estaba incluído en  $G[i...n) \times B[j...m)$ , y probamos que  $b$  no aparece emparejado con nadie en  $T$ , luego podemos restringir la segunda componente de  $T$  a  $B[j + 1...m)$ .
3.  $S^* = S \sqcup T$ , que ya valía antes. Luego, como nuestro código dice:

`j += 1`

Estamos manteniendo el invariante, porque cambiamos de  $j$  a  $j + 1$ .

2. Si  $g < b - 1$ , pasa algo exactamente análogo al caso anterior, y como decimos  $i += 1$ , mantenemos el invariante.
3. Si no, tenemos que  $g \geq b - 1$ , y  $b \geq g - 1$ . Esto implica que  $|g - b| \leq 1$ , y por lo tanto son compatibles. Definimos  $S' = S \cup \{(g, b)\}$ , con  $S' \subseteq G[0...i+1] \times B[0...j+1]$ , porque  $S \subseteq G[0...i] \times B[0...j]$ , y  $S'$  usa  $g = G[i]$  y  $b = B[j]$ . Queremos mostrar que  $S'$  es  $(i+1, j+1)$ -extensible, para probar que se mantiene el invariante. Definamos  $S_G^*$  como los elementos de  $G$  que **no** menciona  $S^*$ , y  $S_B^*$  análogamente para  $B$ . Leér «single girls, single boys». Partimos en casos.
  1. Si  $g \in S_G^* \wedge b \in S_B^*$ . Esto no puede suceder, porque  $(g, b)$  son compatibles, luego  $S^* \cup \{(g, b)\}$  sería una solución con tamaño mayor a  $S^*$ , con  $S^*$  siendo definida como de tamaño máximo.
  2. Si  $g \in S_G^*$ , pero  $b \notin S_B^*$ . Esto significa que existe un emparejamiento  $(x, b) \in S^*$ , con  $x \neq g$ , y que  $g$  no aparece emparejada en  $S^*$ . Como  $b$  no aparece emparejado en  $S$ , tiene que ser que  $(x, b) \in T$ . Consideremos entonces  $T' = T \setminus \{(x, b)\}$ . Definamos  $S^{\{*\}} = S' \sqcup T'$ . Tenemos que  $|S^{\{*\}}| = |S'| + |T'| = |S| + 1 + |T| - 1 = |S| + |T| = |S^*|$ , y luego  $S^{\{*\}}$  también es óptima, porque tiene el mismo tamaño que  $S^*$ . Notamos que agregar  $(g, b)$  es válido, porque  $g$  no aparecía emparejada en  $S^*$ , y rompimos la pareja de  $b$  cuando creamos  $T'$ . Notar también que  $T' \subseteq G[i+1...n] \times G[j+1...m]$ , porque  $g$  no aparece emparejado en  $S^*$  (a fortiori en  $T$ , y en  $T'$ ), y le sacamos a  $T$  la mención de  $b$ . Por lo tanto,  $S'$  es  $(i+1, j+1)$ -extensible.
  3. Si  $b \in S_B^*$ , pero  $g \notin S_G^*$ . Esto es totalmente análogo al caso anterior.
  4. Si  $b \notin S_B^*, g \notin S_G^*$ . Pueden pasar dos cosas, que estén emparejados entre sí, o que cada uno esté emparejado con alguien más.
    1. Si  $(g, b) \in S^*$ . Ya sabíamos que  $(g, b) \notin S$ , y luego como  $S^* = S \sqcup T$ , tenemos que  $(g, b) \in T$ . Definiendo  $T' = T \setminus \{(g, b)\}$ , vemos que  $S^* = S' \sqcup T'$ , y  $T'$  no usa ni  $g = G[i]$  ni  $b = B[j]$ , y luego  $T' \subseteq G[i+1...n] \times B[j+1...m]$ . Luego  $S'$  es  $(i+1, j+1)$ -extensible.
    2. Si no, entonces existen  $(g, x)$  y  $(y, b)$  en  $S^*$ . Como  $g = G[i]$  y  $b = B[j]$  no pueden estar en  $S$  porque  $S \subseteq G[0...i] \times B[0...j]$ , esas parejas tienen que estar en  $T$ . Notamos que  $x \geq b$  y que  $y \geq g$ , porque  $x$  e  $y$  vienen después que  $b$  y  $g$  en  $B$  y  $G$  respectivamente, y  $B$  y  $G$  están ordenados de forma no-decreciente. Vamos a probar que  $x$  e  $y$  son compatibles.
      1. Si  $g = b$ , entonces  $y \geq g = b$ . Como  $(y, g) \in T$ , son compatibles, y luego  $y = g$ , o  $y = g + 1$ . Por el mismo motivo,  $x \geq b = g$ , y como  $(g, x) \in T$ , son compatibles, luego  $x = g$ , o  $x = g + 1$ . Luego,  $\{x, y\} \subseteq \{g, g+1\}$ , luego están a distancia a lo sumo 1 del otro, y luego son compatibles.
      2. Si  $g > b$ , entonces  $g = b + 1$ , porque  $g$  y  $b$  son compatibles. Tenemos  $y \geq g > b$ , y luego  $y = b + 1$ , porque  $(y, b)$  son compatibles, estando emparejados en  $T$ . Como  $x$  es compatible con  $g = b + 1$ , estando emparejados en  $T$ , tenemos que  $x$  es compatible con  $y$ .
      3. Si  $b > g$ , tenemos algo análogo al caso anterior.

En todos los casos,  $x$  e  $y$  son compatibles. Luego, podemos considerar  $T' = T \cup \{(x, y)\} \setminus \{(g, x), (y, b)\}$ , y definimos  $S^{\{*\}} = S' \sqcup T'$ , con  $|S^{\{*\}}| = |S'| + |T'| = |S| + 1 + |T| + 1 - 2 = |S| + |T| = |S^*|$ , y luego  $S^{\{*\}}$  también es una solución óptima. Como  $T$  sólo usa elementos que vienen *después* de  $(g, b) = (G[i], B[j])$ , tenemos que  $T' \subseteq G[i+1\dots n] \times B[j+1\dots m]$ , y luego  $S'$  es  $(i+1, j+1)$ -extensible.

En todos los casos, tenemos que existe  $S'$ , un conjunto  $(i+1, j+1)$ -extensible, con  $|S'| = |S| + 1$ . Recordemos que `res`, antes de entrar al cuerpo de la iteración, era igual a  $|S|$ , por el teorema del invariante. Luego, como nuestro código dice:

```
i += 1
j += 1
res += 1
```

`res` es ahora  $|S'|$ , y el invariante es preservado.

### El ciclo termina

El cuerpo del ciclo siempre aumenta o  $i$  o  $j$ , empezando ambos en cero. Luego, si llegásemos a  $n+m-1$  iteraciones, tendríamos  $i+j = n+m-1$ . Esto haría que la guarda no se cumpla y el ciclo termine - veamos por qué.

1. Si  $i \geq n$ , la guarda no se cumple.
2. Si  $i < n$ , y sabemos que  $n+m-1 = i+j$ , entonces  $n+m+1 < n+j$ , y restando  $n$  de cada lado obtenemos  $m+1 < j$ , o lo que es lo mismo,  $j \geq m$ , con lo cual la guarda no se cumple.

### El invariante es suficiente para demostrar lo que queremos

Al final del ciclo, vale la negación de la guarda, y el invariante. La negación de la guarda es que o bien  $i \geq n$ , o bien  $j \geq m$ . Como vale el invariante, tenemos que  $i \leq n$  y  $j \leq m$ . Luego, sabemos que o bien  $i = n$ , o bien  $j = m$  (pueden pasar las dos juntas). Partimos en casos:

1. Si  $i = n$ , entonces  $S$  es  $(n, j)$ -extensible para algún  $j$ . Esto significa que existe una solución óptima  $S^*$ , y un conjunto  $T \subseteq G[n\dots n] \times B[j\dots m]$ , tal que  $S^* = S \sqcup T$ . Pero  $G[n\dots n] = \emptyset$ , y luego  $T = \emptyset$ , y luego  $S = S^*$ . Luego,  $S$  es una solución óptima.
2. Si  $j = m$ , pasa algo análogo con  $B[m\dots m] = \emptyset$ .

Luego, como sabemos que `res` =  $|S|$ , y devolvemos `res`:

```
return res
```

Nuestro algoritmo devuelve el tamaño de una solución óptima, y luego es correcto. □

## 6.7 Árboles

### Ejercicio 6.7.1

Un bosque es un grafo acíclico. Demostrar que cualquier bosque con  $n$  vértices y  $k$  árboles tiene  $n - k$  aristas.

*Demostración.* Sea  $G$  un bosque, y sean  $G_1 = (V_1, E_1), \dots, G_k = (V_k, E_k)$  sus  $k$  componentes conexas, con  $1 \leq k \leq n$ . Cada componente conexa es acíclica, pues  $G$  lo es, y es conexa.

Luego cada componente conexa es un árbol, y luego  $|E_i| = |V_i| - 1$  para todo  $1 \leq i \leq k$ .

Como ninguna arista puede cruzar componentes conexas (pues las conectaría), cada arista en  $E$  está en exactamente un  $E_i$  para algún  $i$ , y tenemos que  $E = \bigsqcup_{1 \leq i \leq k} E_i$ . Cada vértice, mientras tanto, está en exactamente una componente conexa, y luego  $V = \bigsqcup_{1 \leq i \leq k} V_i$ .

Luego,

$$\begin{aligned}|E| &= \left| \bigsqcup_{1 \leq i \leq k} E_i \right| \\&= \sum_{i=1}^k |E_i| \\&= \sum_{i=1}^k |V_i| - 1 \\&= \left( \sum_{i=1}^k |V_i| \right) - k \\&= \left| \bigsqcup_{1 \leq i \leq k} V_i \right| - k \\&= |V| - k \\&= n - k\end{aligned}$$

que es lo que queríamos demostrar. □

### Ejercicio 6.7.2

Probar que todo árbol con  $n \geq 2$  vértices tiene al menos 2 hojas.



*Demostración.* Sea  $G = (V, E)$  un árbol, con  $n = |V| \geq 2$ , y  $m = |E|$ . Luego,  $G$  es conexo y  $m = n - 1$ . Una hoja es un vértice de grado 1. Sea  $d : V \rightarrow \mathbb{N}$  la función de grado de cada vértice. No puede haber vértices  $v$  con  $d(v) = 0$ , pues como  $n > 1$ ,  $G$  no sería conexo. Luego,  $d(v) \geq 1$  para todo  $v \in V$ . Sea  $H \subseteq V$  el conjunto de hojas de  $T$ . Luego,  $d(v) \geq 2$  para todo vértice en  $V \setminus H$ .

$$\begin{aligned}\sum_{v \in V} d(v) &= \sum_{v \in H} d(v) + \sum_{v \in V \setminus H} d(v) \\&= |H| + \sum_{v \in V \setminus H} d(v) \\&\geq |H| + 2|V \setminus H| \\&= |H| + 2(|V| - |H|) \\&= 2|V| - |H|\end{aligned}$$

También sabemos que  $\sum_{v \in V} d(v) = 2|E| = 2m$ , que vale para todo grafo. Como  $G$  es un árbol,  $m = n - 1$ , y luego  $\sum_{v \in V} d(v) = 2m = 2(n - 1) = 2|V| - 2$ .

Luego,  $2|V| - 2 \geq 2|V| - |H|$ , y luego  $2 \leq |H|$ , con lo cual  $G$  tiene al menos dos hojas.  $\square$

### Ejercicio 6.7.3

Un puente es una arista que, al ser removida, aumenta el número de componentes conexas en un grafo. Mostrar que en un bosque, todas las aristas son puentes.



*Demostración.* Una forma simple de demostrar esto es usando el Ejercicio 6.7.1, y viendo que al sacar una arista, el número de componentes conexas aumenta. Vamos a dar otra demostración.

Sea  $G = (V, E)$  un bosque, y consideremos las componentes conexas  $G_1, \dots, G_k$  de  $G$ . Cada  $G_i$  es un árbol, para todo  $1 \leq i \leq k$ , por ser conexo y acíclico. Sea  $e = \{u, v\}$  una arista en  $E$ . Esa arista pertenece a exactamente un  $G_i$ , pues de no ser así,  $e$  conectaría vértices de dos componentes conexas, que no puede suceder. Sea  $G_i = (V_i, E_i)$  el árbol al que pertenece  $e$ .

Claramente el sacar  $e$  de  $G_i$  no va a cambiar nada sobre  $G_j$  con  $j \neq i$ . Al sacar una  $e$  de  $G_i$ , estamos desconectando  $u, v$ , que antes estaban conectados. Veamos que hay dos componentes conexas en  $G'_i = G_i \setminus \{e\}$ . Sea  $u$  un vértice en  $G'_i = G_i \setminus \{e\}$ .

1. Si hay un camino  $P$  entre  $w$  y  $u$  en  $G_i$  que no usa  $v$ , entonces  $P$  sigue estando en  $G'_i$ , y por tanto  $w$  pertenece en  $G'_i$  a la componente conexa de  $u$ .
2. Si no, el único camino  $P$  de  $w$  a  $u$  en  $G_i$  pasaba por  $v$ . Escribimos  $P = [w, x_1, \dots, x_r, v, \dots, u]$ . Entonces el camino  $[w, x_1, \dots, x_r, v]$  existe en  $G'_i$  entre  $w$  y  $v$  (que no pasa por  $u$ ). Luego  $w$  está en la componente conexa de  $v$ .

Si hubiera un camino en  $G'_i$  entre  $u$  y  $v$ , concatenar  $e$  a ese camino nos daría un ciclo en  $G_i$ , pero  $G_i$  era un árbol. Luego,  $u$  y  $v$  están en componentes conexas en  $G'_i$ , y todos los otros vértices de  $G'_i$  están en una de dos componentes conexas.

Luego  $G'_i$  tiene exactamente dos componentes conexas. Como antes de remover  $e$  de  $G$  teníamos una sola componente conexa en  $G_i$ , y al resto de los  $G_j$  con  $j \neq i$  no le hacemos nada al remover  $e$  de  $G$ , entonces  $G - \{e\}$  tiene exactamente una componente conexa más que  $G$ , y luego  $e$  es un puente.  $\square$

### Ejercicio 6.7.4

Sea  $G = (V, E)$  un grafo conexo pesado, con todos sus pesos distintos. Sea  $(U_1, U_2) = V$  una partición de los vértices de  $G$ . Sea  $e$  una arista de mínimo peso que cruza las particiones.

Entonces  $e$  está en algún árbol generador mínimo de  $G$ .



*Demostración.* Sea  $e = (u, v)$  tal arista. Llamemos  $w : E \rightarrow \mathbb{R}$  a la función de peso de  $G$ , y por comodidad escribamos  $w(H) = \sum_{v \in V(H)} w(v)$  dado un subgrafo  $H$  de  $G$ . Sea  $T$  un árbol generador mínimo de  $G$ .

Como  $T$  es árbol generador, existe en  $T$  un único camino  $P$  entre  $u$  y  $v$ . Como conecta  $u$  y  $v$ , con  $u \in U_1$  y  $v \in U_2$ , en algún momento  $P$  cruza las particiones. Sea  $f$  una arista en  $P$  que cruza esas particiones. El enunciado nos dice que  $w(e) \leq w(f)$ .

Tomemos ahora  $T' = T - f + e$ . Al remover  $f$ , estamos desconectando las particiones  $U_1$  y  $U_2$ , pero como  $e$  cruza esas mismas particiones, agregar  $e$  las reconecta. Luego  $T'$  sigue siendo un árbol generador de  $G$ . Sin embargo, ahora tenemos que  $w(T') = w(T) - w(f) + w(e) \leq w(T)$ . Como  $T$  es un árbol generador mínimo, y  $T'$  un árbol generador, debemos tener  $w(T) \leq w(T')$ . Luego,  $w(T) = w(T')$ , y  $T'$  es un árbol generador mínimo que incluye a  $e$ .  $\square$

### Ejercicio 6.7.5

El algoritmo de Kruskal (resp. Prim) con orden de selección es una variante del algoritmo de Kruskal (resp. Prim) donde a cada arista  $e$  se le asigna una prioridad  $q(e)$  además de su peso  $p(e)$ . Luego, si en alguna iteración del algoritmo de Kruskal (resp. Prim) hay más de una arista posible para ser agregada, entre esas opciones se elige alguna de mínima prioridad.

1. Demostrar que para todo árbol generador mínimo  $T$  de  $G$ , si las prioridades de selección están definidas por la función:

$$q_T(e) = \begin{cases} 0 & \text{si } e \in T \\ 1 & \text{si } e \notin T \end{cases}$$

entonces se obtiene  $T$  como resultado del algoritmo de Kruskal (resp. Prim) con orden de selección ejecutado sobre  $G$  (resp. cualquiera sea el vértice inicial en el caso de Prim).

2. Usando el inciso anterior, demostrar que si los pesos de  $G$  son todos distintos, entonces  $G$  tiene un único árbol generador mínimo.



### Demostración.

1. Probemos primero que el algoritmo de Kruskal encuentra todos los árboles generadores mínimos, y luego lo mismo para el algoritmo de Prim. Llamemos  $w(e) = (p(e), q_T(e))$ , donde ordenamos las tuplas por orden lexicográfico. Es decir,  $w(e) \leq w(e') \Leftrightarrow p(e) < p(e') \vee (p(e) = p(e') \wedge q_T(e) < q_T(e'))$ .

### Lema 6.7.6

El algoritmo de Kruskal con selección devuelve  $T$ .



*Demostración.* Por «Calamardo» en Telegram. Sea  $K$  el árbol generador que devuelve el algoritmo de Kruskal con selección. Asumimos que  $K \neq T$ . Luego, existe una primer arista  $e$  que el algoritmo agregó a  $K$  que no está en  $T$ . Consideremos  $T' = T \cup \{e\}$ . Este subgrafo de  $G$  contiene un único ciclo  $C$ , con  $e \in C$ . Como  $K$  es un árbol,  $C \not\subseteq K$ , puesto que un árbol no tiene ciclos. Luego existe alguna arista  $f \in C$  tal que  $f \notin K$ . Como  $f \notin K$  pero  $e \in K$ , tenemos que  $f \neq e$ . Como  $f \neq e$ , y  $f \in C \subseteq T + e$ , entonces  $f \in T$ . Consideremos entonces  $T'' = T' \setminus \{f\} = (T \cup \{e\}) \setminus \{f\}$ . Esto es un árbol generador, puesto que rompimos el único ciclo,  $C$ , que había en  $T'$ . Vemos que  $p(T'') = p(T) +$

$p(e) - p(f)$ . Como  $T$  es un árbol generador mínimo, debemos tener  $p(T'') \geq p(T)$ , con lo cual  $p(e) \geq p(f)$ .

Como  $e \notin T$ , y  $f \in T$ , tenemos  $q_T(f) = 0$ ,  $q_T(e) = 1$ . Luego,  $w(f) < w(e)$ , y el algoritmo de Kruskal con selección vio primero a  $f$ . Sean  $j$  el paso en el que el algoritmo vio (y decidió agregar) a  $e$ , e  $i$  el paso en el que el algoritmo vio a  $f$ . Tenemos que  $j > i$ . Dada una iteración  $t$ , sea  $F_t$  el conjunto de aristas que el algoritmo agregó al comenzar la iteración  $t$ . Tenemos que  $F_t \subseteq F_{t+r}$  para todo  $r \geq 1$ . Como  $j > i$ , tenemos que  $F_i \subseteq F_j$ . El algoritmo de Kruskal con selección va a agregar a una arista  $x$  si y sólo si, al verla en un paso  $k$ ,  $x$  no genera ciclos con  $F_k$ .

Como  $e$  es la primer arista que el algoritmo agrega que no está en  $T$ , entonces  $F_j \subseteq T$  (y  $F_{j+1} \not\subseteq T$ ). Como  $f \in T$ , y  $T$  es un árbol,  $f$  no genera ciclos con nadie en  $T$ , menos aún con aristas en  $F_j$ , y como  $F_i \subseteq F_j$ , menos aún va a generar ciclos con  $F_i$ . Luego, al ver a  $f$  en el paso  $i$ , el algoritmo pudo haber agregado a  $f$ , y no lo hizo, pero eso no puede pasar.

Luego,  $e$  no existe, y  $K = T$ . □

#### Lema 6.7.7

El algoritmo de Prim con selección devuelve  $T$ . ♥

*Demuestra*ón. Sea  $P$  el árbol generador que construye el algoritmo de Prim con selección. Recordemos que este algoritmo mantiene una partición  $(S_i, \overline{S}_i)$  de  $V(G)$ , y un conjunto de aristas  $F_i$ , donde  $F_i$  genera  $S_i$  (la componente conexa). Inicialmente  $S_0 = \{v_0\}$  para algún  $v_0$  arbitrario, y  $F_0 = \emptyset$ . En cada iteración, el algoritmo busca la arista de mínimo  $w$  que cruza la partición. Si esta arista es  $\{u, v\}$ , con  $u \in S_i$  y  $v \notin S_i$ , tenemos  $S_{i+1} = S_i \cup \{v\}$ , y  $F_{i+1} = F_i \cup \{e\}$ . El algoritmo se detiene cuando  $S_{n-1} = V(G)$ , y devuelve  $F_{n-1}$ , donde  $n = |V(G)|$ .

Asumimos que  $P \neq T$ . Luego, existe una primer arista  $e$  que el algoritmo agregó a la componente conexa, que no está en  $T$ . Sean  $\{u, v\} = e$  los extremos de  $e$ . Asumimos sin pérdida de generalidad que  $u$  es el extremo que está en  $S_i$ , y  $v$  el extremo que está en  $\overline{S}_i$ . Entonces, tenemos  $S_{i+1} = S_i \cup \{v\}$ ,  $F_{i+1} = F_i \cup \{e\}$ .

Como  $T$  es un árbol generador, pero  $e \notin T$ , existe un (único) camino  $Q$  en  $T$  entre  $u$  y  $v$ . Como  $u$  y  $v$  están en diferentes lados de la partición  $(S_i, \overline{S}_i)$ , entonces en  $Q$  existe alguna arista  $f$  que cruza la partición. Consideraremos entonces  $T' = (T \setminus \{f\}) \cup \{e\}$ . Esto es un árbol generador, por haber separado a  $T$  en dos pedazos cuando sacamos  $f$ , uno que genera  $S$  y otro que genera  $\overline{S}$ , y luego haberlos juntado al agregar  $e$ . Como  $T$  es un árbol generador mínimo, tenemos que  $p(T) \leq p(T') = p(T) - p(f) + p(e)$ , con lo cual  $p(e) \geq p(f)$ .

Como  $f$  está en  $T$ , tenemos que  $q_T(f) = 0$ . Como  $e$  no está en  $T$ , tenemos que  $q_T(e) = 1$ . Como  $p(f) \leq p(e)$ , y  $q_T(f) < q_T(e)$ , tenemos que  $w(f) < w(e)$ . Como ambas aristas cruzan la partición  $(S_i, \overline{S}_i)$ , al considerar la arista  $e$ , el algoritmo también consideró la arista  $f$ . Pero entonces, el algoritmo no eligió a la arista que cruza la partición, de menor  $w$ . Esto no puede suceder.

Luego,  $e$  no existe, y  $P = T$ . □

2. Lo que nos dice el ejercicio anterior es que para todo árbol generador mínimo  $T$ , ambos el algoritmo de Kruskal y el algoritmo de Prim tienen ejecuciones que devuelven  $T$ , dependiendo sólo de cómo se desempata cuando encuentran aristas del mismo peso, en cada ejecución. Es decir, el conjunto de resultados posibles del algoritmo de Kruskal sobre  $G$  (resp. Prim), es igual al conjunto de árboles generadores mínimos de  $G$

Luego, si para un grafo  $G$  en ningún momento hace falta desempatar, porque todos los pesos de las aristas de  $G$  son distintos, entonces en toda ejecución el resultado del algoritmo de Kruskal (resp. Prim) es único.

Como el conjunto de árboles generadores mínimos es igual al conjunto de resultados de correr estos algoritmos, y este último tiene un sólo elemento cuando los pesos de  $G$  son todos distintos, entonces  $G$  tiene un único árbol generador mínimo.

□

## 6.8 Caminos mínimos

### Ejercicio 6.8.1

Tenemos un mapa con ciudades y rutas entre pares de ciudades. Para cada ruta  $e$ , tenemos un número  $0 \leq p(e) < 1$  que nos indica la probabilidad de que nos caiga un rayo mientras estamos en esa ruta.

Diseñar un algoritmo que, dado un tal mapa y dos ciudades  $a$  y  $b$ , devuelva un camino entre  $a$  y  $b$ , que minimice la probabilidad de que nos caiga un rayo.

♣

*Demostración.* Si tenemos dos eventos independientes,  $x$  e  $y$ , entonces  $P(x \wedge y) = P(x)P(y)$ . En general, si tenemos eventos disjuntos  $\{x_i\}_{i \in I}$ , la probabilidad de que pasen todos es  $\prod_{i \in I} P(x_i)$ .

Consideremos ahora cualquier camino  $P = [e_1, e_2, \dots, e_k]$  desde  $a$  hasta  $b$ . La probabilidad de que no nos caiga un rayo en ningún momento es la probabilidad de que no nos caiga un rayo mientras estamos en  $e_1$ , y que no nos caiga un rayo mientras estamos en  $e_2$ , etc. Luego, esta probabilidad es  $X(P) = \prod_{e \in P} (1 - p(e))$ . Luego, si queremos maximizar la probabilidad de que no nos caiga un rayo, como la función logaritmo es monótona creciente, podemos maximizar  $\log X(P) = \sum_{e \in P} \log (1 - p(e))$ . Esto es lo mismo que minimizar  $-\log X(P) = -\sum_{e \in P} \log (1 - p(e)) = \sum_{e \in P} -\log (1 - p(e))$ .

Luego, podemos tomar cada probabilidad  $p(e)$ , y transformarla en una función de peso,  $w$ , con  $w(e) = -\log (1 - p(e))$ . Como  $0 \leq p(e) < 1$ , entonces  $1 \geq 1 - p(e) > 0$ , y luego  $0 \leq -\log (1 - p(e)) < \infty$ . Entonces vemos que  $w : \mathbb{E} \rightarrow \mathbb{R}_{\geq 0}$ .

Luego, consideramos el grafo pesado  $G = (V, E, w)$ , donde los vértices  $V$  son las ciudades, las aristas  $E$  son las rutas entre ciudades, y  $w$  nos da el peso de cada camino. Como vimos, maximizar la probabilidad de que no nos caiga un rayo es lo mismo que encontrar un camino  $P$  desde  $a$  hasta  $b$  que minimice  $\sum_{e \in P} -\log (1 - p(e))$ . Si usamos  $w$  como pesos en las aristas, esto es encontrar un camino de suma de pesos mínima entre todos los caminos entre  $a$  y  $b$ . Para esto podemos usar el algoritmo de Dijkstra para caminos mínimos.

```

from math import log
def F(Ciudades: set[int],
      Rutas: dict[int, list[int]],
      p: Callable[[int, int], float],
      a: int,
      b: int):
    def w(i, j):
        return - log(1.0 - p(i, j))
    G = (Ciudades, Rutas, w)
    return Dijkstra(G, a)[b]

```

□

### Ejercicio 6.8.2

Demostrar la correctitud del algoritmo de caminos mínimos de Dijkstra.

♦

Este va a ser un ejemplo de una demostración con menos formalismo que el típico para algoritmos con ciclos (es decir, el teorema del invariante), pero el mismo rigor. Tenemos que tener cuidado cuando hacemos esto, porque arriesgamos caer en «porque el vértice que no era el último sino el anteúltimo que en la anterior iteración cambiamos al valor que el otro vértice tenía antes de ser agregado cuando sabíamos el valor de  $d$  para todos los otros vértices», y demás oraciones incomprensibles.

*Demostración.* Recordemos el algoritmo de Dijkstra para caminos mínimos.

```

1: procedure DIJKSTRA( $G = (V, E)$ ,  $s \in V$ ,  $w : E \rightarrow \mathbb{R}_{\geq 0}$ )
2:    $Q \leftarrow V$ 
3:    $d[v] \leftarrow \infty \quad \forall v \in V \setminus \{s\}$ 
4:    $p[v] \leftarrow \text{NULL} \quad \forall v \in V$ 
5:    $d[s] \leftarrow 0$ 
6:   while  $Q \neq \emptyset$  do
7:      $u \leftarrow \arg \min_{u \in Q} d[u]$ 
8:      $Q \leftarrow Q \setminus \{u\}$ 
9:     for  $(u, v) \in E$  do
10:       if  $d[u] + w(u, v) < d[v]$  then
11:          $d[v] \leftarrow d[u] + w(u, v)$ 
12:          $p[v] \leftarrow u$ 
13:       end
14:     end
15:   end
16:   return  $d, p$ 
17: end

```

Vamos a demostrar y usar los dos invariantes del algoritmo.

### Lema 6.8.3

En todo momento, para todo  $v \in V$ , si  $d[v] < \infty$ , entonces  $d[v]$  es la longitud de algún camino desde  $s$  a  $v$ . En particular,  $d[v] \geq \delta(s, v)$ , con  $\delta(x, y)$  la distancia en  $G$  entre los vértices  $x$  e  $y$ .

Además, si  $d[v] < \infty$  y  $v \neq s$ , entonces  $p[v]$  es un vértice inmediatamente anterior a  $v$  en algún camino desde  $s$  a  $y$  con longitud  $d[v]$ . Si  $d[v] = \infty$  o  $v = s$ , entonces  $p[v] = \text{NULL}$ .

#### Demostración.

- Antes de comenzar el ciclo,  $d[s] = 0$ , y  $d[v] = \infty$  para todo  $v \in V \setminus \{s\}$ . Como hay un camino de longitud 0 de  $s$  a  $s$ , y  $\delta(s, s) = 0$  al ser todas las aristas de peso positivo,  $d[s] = \delta(s, s)$ . Para todos los otros vértices no hace falta probar nada sobre  $d$  (pues el antecedente en «si  $d[v] < \infty$ , entonces  $d[v]$  es la longitud de algún camino entre  $s$  y  $v$ » es falso para ellos), probamos el caso base. Vemos también que  $p[v] = \text{NULL}$  para todo  $v \in V$ , luego lo que dijimos sobre  $p$  es correcto.
- Ahora en cualquier iteración del ciclo, cuando asignamos  $d[v] \leftarrow d[u] + w(u, v)$ , por hipótesis inductiva  $d[u]$  es la longitud de algún camino  $P$  desde  $s$  hasta  $u$ . Luego, como hay una arista  $(u, v) \in E$ , de peso  $w(u, v)$ , tenemos un camino  $Q = P + [(u, v)]$  de costo  $d[u] + w(u, v)$  desde  $s$  hasta  $v$ , y luego  $d[v]$  es la longitud de algún camino entre  $s$  y  $v$ . Asimismo, este camino tiene a  $u$  como antecesor directo de  $v$  en este camino de longitud  $d[v]$ , y luego nuestra asignación de  $p[v] \leftarrow u$  es correcta.

Como estas son las únicas dos formas de modificar  $d$  y  $p$ , probamos que esto se cumple en cualquier iteración del algoritmo.  $\square$

### Lema 6.8.4

Cuando sacamos a  $u$  de  $Q$ ,  $d[u] = \delta(s, u)$ .

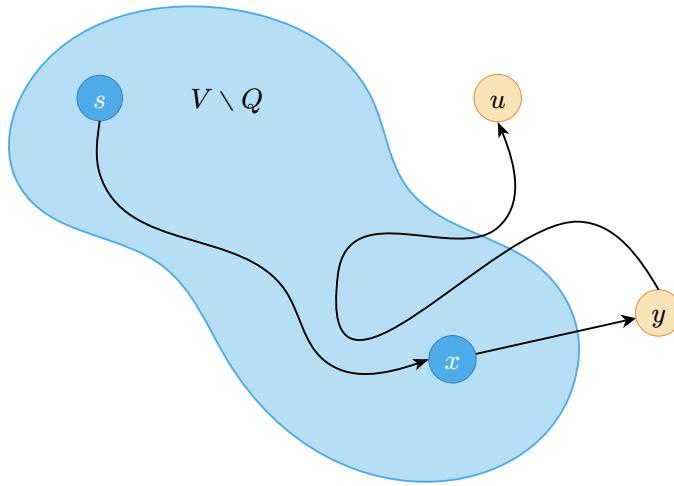
*Demostración.* Vamos a probar esto por inducción en el número de iteraciones del ciclo. Sea  $P(i)$ : Al comenzar la  $i$ -ésima iteración del ciclo, tenemos  $d[v] = \delta(s, v)$  para todo  $v \in V \setminus Q$ .

1.  $P(0)$ . Al comenzar el ciclo, no hay nada que probar, pues  $Q = V$ , entonces no hay nadie en  $V - Q$ .
2.  $P(1)$ . El único vértice que sacamos de  $Q$  en la primer iteración fue  $s$ , y luego al comenzar la segunda,  $Q = V \setminus \{s\}$ . En este caso,  $d[s] = 0$ , y no lo modificamos porque en la primer iteración no vemos aristas  $(s, s)$  ( $G$  es un grafo, no multigrafo). Como  $\delta(s, s) = 0 = d[s]$ , vale  $P(1)$ .
3. Paso inductivo. Sabemos  $P(k)$  para todo  $k \leq i$ , queremos ver que vale  $P(i + 1)$ . Sea  $u$  el vértice que estamos sacando en la  $i$ -ésima iteración. Este es el vértice que, al sacar de  $Q$  en la  $i$ -ésima iteración, estamos «agregando» a  $V \setminus Q$ . Luego, como para

todos los otros vértices  $v$  en  $V \setminus Q$  sabemos que  $d[v] = \delta(s, v)$  por el Lema 6.8.3, y no van a cambiar más porque en cada iteración sólo pueden decrecer, lo único que nos queda probar es que para  $u$ , también tenemos que  $d[u] = \delta(s, u)$ .

Si  $d(s, u) = \infty$ , es decir no hay ningún camino entre  $s$  y  $u$ , terminamos, pues por el Lema 6.8.3, sabemos que  $d[u] = \infty$ .

De otra forma, consideraremos un camino  $P$  de longitud mínima desde  $s$  hasta  $u$ . Sea  $y$  el primer vértice en  $P$  que está en  $Q$ , y sea  $x$  su predecesor en  $P$  (podríamos tener  $y = u$ , o  $x = s$ ).



Como  $y$  aparece no-después que  $u$  en  $P$ , y  $P$  es un camino mínimo con todas las aristas de peso no-negativo, tenemos que  $\delta(s, y) \leq \delta(s, u)$ . Como  $u = \arg \min_{u \in Q} d[u]$ , y sabemos que  $y \in Q$ , entonces  $d[u] \leq d[y]$ . Por el Lema 6.8.3, sabemos que  $\delta(s, u) \leq d[u]$ .

Como  $x \in V \setminus Q$ , fue removido de  $Q$  en alguna iteración  $j \leq i$ . Podemos usar  $P(j)$ , entonces, para ver que  $d[x] = \delta(s, x)$  al removerlo. Durante la iteración  $j$ , nos aseguramos que  $d[y] \leq d[x] + w(x, y) = \delta(s, x) + w(x, y)$ . Como  $s \rightsquigarrow x \rightarrow y$  es parte de un camino mínimo  $P$ , entonces  $s \rightsquigarrow x \rightarrow y$  es también un camino mínimo, y entonces  $\delta(s, x) + w(x, y) = \delta(s, y)$ . Como  $d[y] \leq \delta(s, y)$ , y por el Lema 6.8.3 sabemos que  $d[y] \geq \delta(s, y)$ , tenemos que  $d[y] = \delta(s, y)$  al final de la iteración  $j$ , y luego por el Lema 6.8.3, sigue valiendo en la iteración  $i + 1$ .

Luego sabemos que  $\delta(s, y) \leq \delta(s, u) \leq d[u] \leq d[y]$ , y  $d[y] = \delta(s, y)$ . Luego,  $\delta(s, y) = \delta(s, u) = d[u] = d[y]$ . Luego  $d[u] = \delta(s, u)$  al terminar la iteración  $i$ , es decir, vale también al comenzar la iteración  $i + 1$ , y luego vale  $P(i + 1)$ .

□

Como el algoritmo termina cuando  $Q = \emptyset$ , y empieza con  $Q = V$ , habremos al final sacado de  $Q$  todos los vértices  $v \in V$ . Por el Lema 6.8.4, Al momento de sacar cada  $v \in V$  de  $Q$ , tendremos  $d[v] = \delta(s, v)$ , y luego  $d[v]$  es la longitud de un camino mínimo entre  $s$  y  $v$ . Por el Lema 6.8.3, como cada vez que actualizamos  $d[v]$  sólo lo hacemos decrecer, y siempre es la

longitud de un camino entre  $s$  y  $v$ , nunca puede ser actualizado a algo menor a  $\delta(s, v)$ . Vemos entonces que luego de sacar a  $v$  de  $Q$ ,  $d[v]$  nunca más se actualiza, y permanece en  $\delta(s, v)$ .

Luego, al terminar el algoritmo, tenemos  $d[v]$  para todo  $v \in V$ , y es igual a  $\delta(s, v)$ . Más aún, para todo  $v \in V$ ,  $p[v]$  es o bien `NULL` (cuando  $s = v$ , o cuando  $d[v] = \infty$ ), o es el antecesor directo de  $v$  en algún camino mínimo desde  $s$  hasta  $v$ .  $\square$

### Ejercicio 6.8.5

Sea  $G = (V, E)$  un grafo dirigido pesado con pesos positivos, y  $s, t \in V$ . Queremos encontrar el camino de menor peso entre  $s$  y  $t$ . De todos los caminos con menor peso, queremos el que menos aristas tenga.

Dar un algoritmo que devuelva un tal camino. Demostrar que es correcto.



*Demostración.* Para entender cómo resolver este ejercicio es crucial ver la demostración de por qué funciona el algoritmo de Dijkstra, que damos en el Ejercicio 6.8.2.

El invariante que mantiene el algoritmo de Dijkstra es que para todos los vértices que sacamos de la cola  $Q$ , sabemos exactamente su distancia desde  $s$ , el vértice inicial. Mantenemos este invariante removiendo de  $Q$  el vértice  $u$  con menor distancia estimada  $d[u]$ , y actualizando el estimado de las distancias para todos vecinos  $v$  de  $u$  ( $d[v] \leftarrow \min(d[v], d[u] + w(u, v))$ ).

Si queremos comparar dos cosas, en orden, la estructura típica para esto son los pares, con su orden lexicográfico. En este caso, si queremos comparar primero por distancia (suma de pesos), y habiendo empates, por número de aristas, podemos definir:

- $\delta(s, v)$  como la suma de los pesos en un camino de mínima distancia desde  $s$  hasta  $v$ ,
- $\kappa(s, v)$  como el mínimo número de aristas entre todos los caminos de mínima distancia entre  $s$  y  $v$ ,
- $\varepsilon(s, v) = (\delta(s, v), \kappa(s, v))$

Vemos que  $\varepsilon$  se comporta de manera similar a  $\delta$ . Si tenemos un camino de mínima distancia y mínimo número de aristas entre todos los caminos mínimos  $P = [s, \dots, x, y]$ , entre  $s$  y  $y$ , entonces no solo  $\delta(s, y) = \delta(s, x) + w(x, y)$  (que usamos en la demostración del algoritmo de Dijkstra), sino que  $\varepsilon(s, y) = \varepsilon(s, x) + (w(x, y), 1)$ , donde definimos la suma componente-a-componente.

Veamos si podemos adaptar el algoritmo original a esta nueva noción de «distancia».

### ⓘ Nota

Como estamos cambiando el algoritmo, tenemos que demostrar que este nuevo algoritmo es correcto con respecto a la nueva especificación que estamos pidiendo. No podemos usar la demostración de un algoritmo distinto, y sólo decir que «sigue valiendo», primero porque no lo demostramos (y en general es muy difícil decir que «siguen valiendo» absolutamente todas las deducciones que se hicieron en otra demostración), y segundo porque el anterior algoritmo cumple un objetivo *distinto* al nuevo.

Tenemos que ir oración por oración de la demostración anterior, ver qué sigue valiendo, y qué no y hay que cambiar y probar. En particular, vamos a necesitar un nuevo lema acá, el Lema 6.8.7.

```

1: procedure DIJKSTRA( $G = (V, E)$ ,  $s \in V$ ,  $w : E \rightarrow \mathbb{R}_{\geq 0}$ )
2:    $Q \leftarrow V$ 
3:    $d[v] \leftarrow (\infty, \infty) \forall v \in V \setminus \{s\}$ 
4:    $p[v] \leftarrow \text{NULL} \forall v \in V$ 
5:    $d[s] \leftarrow (0, 0)$ 
6:   while  $Q \neq \emptyset$  do
7:      $u \leftarrow \arg \min_{u \in Q} d[u]$ 
8:      $Q \leftarrow Q \setminus \{u\}$ 
9:     for  $(u, v) \in E$  do
10:       if  $d[u] + (w(u, v), 1) < d[v]$  then
11:          $d[v] \leftarrow d[u] + (w(u, v), 1)$ 
12:          $p[v] \leftarrow u$ 
13:       end
14:     end
15:   end
16:   return  $d, p$ 
17: end

```

#### Definición 6.8.6

Llamamos a un camino  $[s, \dots, u]$   **$\varepsilon$ -óptimo** cuando tiene longitud mínima entre todos los caminos desde  $s$  hasta  $u$ , y dentro de todos esos caminos de longitud mínima, tiene el mínimo número de aristas.

#### Lema 6.8.7 (Subestructura óptima de caminos $\varepsilon$ -óptimos)

Sea  $P = [s, \dots, x, y]$  un camino  $\varepsilon$ -óptimo, y sea  $Q = [s, \dots, x]$  su prefijo. Entonces  $Q$  es  $\varepsilon$ -óptimo.

En particular,  $\varepsilon(s, x) \leq \varepsilon(s, y)$ .

*Demostración.* Vamos a probar esto por inducción. Por conveniencia notamos  $w(X) = \sum_{e \in X} w(e)$ , para cualquier camino  $X$ .

Como  $P$  es  $\varepsilon$ -óptimo, entonces tiene mínina distancia hasta  $y$ , y luego  $w(P) = \delta(s, y)$ . Asimismo, vemos que  $w(P) = w(Q) + w(x, y)$ . Como  $Q$  es un camino desde  $s$  hasta  $x$ , y  $\delta(s, x)$  es la mínima distancia desde  $s$  hasta  $x$ , entonces  $w(Q) \geq \delta(s, x)$ .

Si  $w(Q) > \delta(s, x)$ , podemos tomar cualquier camino mínimo  $Q'$  entre  $s$  y  $x$ , con  $w(Q') = \delta(s, x)$ . Entonces,  $Q' + (x, y)$  es un camino entre  $s$  e  $y$ , con distancia  $w(Q') + w(x, y) < w(Q) + w(x, y) = w(P)$ , pero  $P$  era de mínima distancia entre  $s$  e  $y$ , entonces  $Q'$  no puede existir. Luego,  $w(Q) = \delta(s, x)$ .

Como  $Q$  es un camino de mínima longitud desde  $s$  hasta  $x$ , y  $\kappa(s, x)$  es el mínimo número de aristas entre todos los caminos mínimos desde  $s$  hasta  $x$ , entonces  $|Q| \geq$

$\kappa(s, x)$ , con  $|Q|$  el número de aristas de  $Q$ . Si  $|Q| > \kappa(s, x)$ , entonces sea  $Q'$  cualquier camino de mínima longitud entre  $s$  y  $x$ , y de entre ellos, uno de mínimo número de aristas (es decir,  $\kappa(s, x)$  aristas). Tenemos  $w(Q') = w(Q)$ , pues ambos son caminos mínimos entre  $s$  y  $x$ . Luego  $Q' + (x, y)$  es un camino de longitud  $w(Q') + w(x, y) = w(Q) + w(x, y) = w(P)$ , y número de aristas  $|Q'| + 1 = \kappa(s, x) + 1 < |Q| + 1 = |P|$ . Pero  $P$  era un camino  $\varepsilon$ -óptimo, entonces esto no puede pasar. Luego,  $Q'$  no puede existir, y tenemos  $|Q| = \kappa(s, x)$ . Luego, como  $w(Q) = \delta(s, x)$ , y  $|Q| = \kappa(s, x)$ , tenemos que  $Q$  es  $\varepsilon$ -óptimo.  $\square$

#### Lema 6.8.8

En todo momento, para todo  $v \in V$ , si  $d[v] = (a, b) < (\infty, \infty)$ , entonces existe un camino entre  $s$  y  $v$ , con longitud  $a$ , y  $b$  aristas. En particular,  $d[v] \geq \varepsilon(s, v)$ .

Además, si  $d[v] \neq (\infty, \infty)$  y  $v \neq s$ , entonces  $p[v]$  es un vértice inmediatamente anterior a  $v$  en algún camino desde  $s$  a  $y$  con longitud  $d[v][0]$  y número de aristas  $d[v][1]$ . Si  $d[v] = (\infty, \infty)$  o  $v = s$ , entonces  $p[v] = \text{NULL}$ .



#### Demostración.

- Antes de comenzar el ciclo,  $d[s] = (0, 0)$ , y  $d[v] = (\infty, \infty)$  para todo  $v \in V \setminus \{s\}$ . Como hay un camino de longitud 0 de  $s$  a  $s$ , y  $\delta(s, s) = 0$  al ser todas las aristas de peso positivo,  $d[s] = (\delta(s, s), \kappa(s, s)) = \varepsilon(s, s)$ . Para todos los otros vértices no hace falta probar nada sobre  $d$  (pues el antecedente en «si  $d[v] < (\infty, \infty)$ , entonces ...» es falso para ellos). Vemos también que  $p[v] = \text{NULL}$  para todo  $v \in V$ , luego lo que dijimos sobre  $p$  es correcto.
- Ahora en cualquier iteración del ciclo, cuando cambiamos  $d[v] = d[u] + (w(u, v), 1)$ , por hipótesis inductiva, notando  $d[u] \leftarrow (a, b)$ ,  $a$  es la longitud de algún camino  $P$  desde  $s$  hasta  $u$ , y  $b$  es el número de aristas de  $P$ . Luego, como hay una arista  $(u, v) \in E$ , de peso  $w(u, v)$ , tenemos un camino  $Q = P + [(u, v)]$  de costo  $d[u] + w(u, v)$  desde  $s$  hasta  $v$ , con  $b + 1$  aristas. Luego, existe un camino  $Q$  tal que  $d[v] = (a', b')$ , con  $a'$  la longitud de  $Q$ , y  $b'$  el número de aristas de  $Q$ . En  $P$ , el antecesor directo de  $v$  es  $u$ , y luego nuestra asignación  $p[v] \leftarrow u$  es correcta.

Como estas son las únicas dos formas de modificar  $d$  y  $p$ , probamos que esto se cumple en cualquier iteración del algoritmo.  $\square$

#### Lema 6.8.9

Cuando sacamos a  $u$  de  $Q$ ,  $d[u] = \varepsilon(s, u)$ .



*Demostración.* Vamos a probar esto por inducción en el número de iteraciones del ciclo. Sea  $P(i)$ : Al comenzar la  $i$ -ésima iteración del ciclo, tenemos  $d[v] = \varepsilon(s, v)$  para todo  $v \in V \setminus Q$ .

1.  $P(0)$ . Al comenzar el ciclo, no hay nada que probar, pues  $Q = V$ , entonces no hay nadie en  $V - Q$ .
2.  $P(1)$ . El único vértice que sacamos de  $Q$  en la primer iteración fue  $s$ , y luego al comenzar la segunda,  $Q = V \setminus \{s\}$ . En este caso,  $d[s] = (0, 0)$ , y no lo modificamos porque en la primer iteración no vemos aristas  $(s, s)$  ( $G$  es un grafo, no multigrafo). Como  $\delta(s, s) = 0$ , y  $\kappa(s, s) = 0$ , vale  $P(1)$ .
3. Paso inductivo. Sabemos  $P(k)$  para todo  $k \leq i$ , queremos ver que vale  $P(i + 1)$ . Sea  $u$  el vértice que estamos sacando en la  $i$ -ésima iteración. Este es el vértice que, al sacar de  $Q$  en la  $i$ -ésima iteración, estamos «agregando» a  $V \setminus Q$ . Luego, como para todos los otros vértices  $v$  en  $V \setminus Q$  sabemos que  $d[v] = \varepsilon(s, v)$  por el Lema 6.8.8, y no van a cambiar más porque en cada iteración sólo pueden decrecer, lo único que nos queda probar es que para  $u$ , también tenemos que  $d[u] = \varepsilon(s, u)$ .

Si  $\delta(s, u) = \infty$ , es decir no hay ningún camino entre  $s$  y  $u$ , usamos el contrarecíproco del Lema 6.8.8, y al no haber ningún camino, entonces  $d[u] = (\infty, \infty)$ .

De otra forma, consideraremos un camino  $\varepsilon$ -óptimo  $P$  desde  $s$  hasta  $u$ . Sea  $y$  el primer vértice en  $P$  que está en  $Q$ , y sea  $x$  su predecesor en  $P$  (podríamos tener  $y = u$ , o  $x = s$ ).

Como  $P$  es  $\varepsilon$ -óptimo, por el Lema 6.8.7,  $\varepsilon(s, y) \leq \varepsilon(s, u)$ . Como  $u = \arg \min_{u \in Q} d[u]$ , y sabemos que  $y \in Q$ , entonces  $d[u] \leq d[y]$ . Por el Lema 6.8.8, sabemos que  $\varepsilon(s, u) \leq d[u]$ .

Como  $x \in V \setminus Q$ , fue removido de  $Q$  en alguna iteración  $j \leq i$ . Podemos usar  $P(j)$ , entonces, para ver que  $d[x] = \varepsilon(s, x)$  al removerlo. Durante la iteración  $j$ , nos aseguramos que  $d[y] \leq d[x] + (w(x, y), 1) = \varepsilon(s, x) + (w(x, y), 1)$ . Por Lema 6.8.7, como  $s \rightsquigarrow x \rightarrow y$  es parte de un camino  $\varepsilon$ -óptimo  $P$ , entonces  $s \rightsquigarrow x \rightarrow y$  es también  $\varepsilon$ -óptimo, y entonces  $\varepsilon(s, x) + (w(x, y), 1) = \varepsilon(s, y)$ . Como  $d[y] \leq \varepsilon(s, y)$ , y por el Lema 6.8.8 sabemos que  $d[y] \geq \varepsilon(s, y)$ , tenemos que  $d[y] = \varepsilon(s, y)$  al final de la iteración  $j$ , y luego por el Lema 6.8.8, sigue valiendo en la iteración  $i + 1$ .

Luego sabemos que  $\varepsilon(s, y) \leq \varepsilon(s, u) \leq d[u] \leq d[y]$ , y  $d[y] = \varepsilon(s, y)$ . Luego,  $\varepsilon(s, y) = \varepsilon(s, u) = d[u] = d[y]$ . Luego  $d[u] = \varepsilon(s, u)$  al terminar la iteración  $i$ , es decir, vale también al comenzar la iteración  $i + 1$ , y luego vale  $P(i + 1)$ .

□

Como el algoritmo termina cuando  $Q = \emptyset$ , y empieza con  $Q = V$ , habremos al final sacado de  $Q$  todos los vértices  $v \in V$ . Por el Lema 6.8.9, Al momento de sacar cada  $v \in V$  de  $Q$ , tendremos  $d[v] = \varepsilon(s, v)$ , y luego  $d[v]$  es la longitud de un camino mínimo entre  $s$  y  $v$ . Por el Lema 6.8.8, como cada vez que actualizamos  $d[v]$  sólo lo hacemos decrecer, y siempre es un par con la longitud de un camino entre  $s$  y  $v$  y su número de aristas, nunca puede ser

actualizado a algo menor a  $\varepsilon(s, v)$ . Vemos entonces que luego de sacar a  $v$  de  $Q$ ,  $d[v]$  nunca más se actualiza, y permanece en  $\varepsilon(s, v)$ .

Luego, al terminar el algoritmo, tenemos  $d[v]$  para todo  $v \in V$ , y es igual a  $\varepsilon(s, v)$ . Más aún, para cada  $v \in V$ ,  $p[v]$  es o NULL cuando  $v = s$  o  $d[v] = (\infty, \infty)$ , o el antecesor inmediato de  $v$  en algún camino  $\varepsilon$ -óptimo desde  $s$  hasta  $v$ .  $\square$

#### Ejercicio 6.8.10

Sea  $G = (V, E)$  un grafo dirigido pesado por  $w : E \rightarrow \mathbb{R}_{>0}$ , y  $s \in V$ .

Dar un algoritmo que calcule el número de caminos de mínimo peso entre  $s$  y  $v$ , para cada  $v \in V$ .



*Demostración.* La idea de este algoritmo es primero encontrar el grafo dirigido acíclico  $G'$  de caminos mínimos desde  $s$ . Luego, queremos encontrar el número de caminos en  $G'$ , desde  $s$  hasta  $v$ , para cada  $v$ . Para la segunda parte, podemos usar un simple algoritmo de programación dinámica. Para la primer parte, tenemos dos opciones:

- Podemos modificar el algoritmo de Dijkstra. Normalmente, el algoritmo devuelve un array  $p$ , donde  $p[v]$  nos da un vértice inmediatamente anterior a  $v$  en un camino mínimo desde  $s$  hasta  $v$ . Podemos modificar esto para que  $p[v]$  guarde un conjunto de *todos* los vértices que están inmediatamente antes que  $v$ , en algún camino mínimo desde  $s$  hasta  $v$ . La modificación sería que al iterar cada arista  $(u, v) \in E$ , incidente al vértice  $u$  que sacamos de  $Q$ , hacemos:

```

1: if  $d[v] < d[u] + w(u, v)$  then
2:    $d[v] \leftarrow d[u] + w(u, v)$ 
3:    $p[v] \leftarrow \{u\}$ 
4: else if  $d[v] = d[u] + w(u, v)$  then
5:    $p[v] \leftarrow p[v] \cup \{u\}$ 
6: end

```

Esta es una modificación útil, pero como vimos en el ejercicio anterior, tenemos que volver a demostrar que el algoritmo de Dijkstra con esta modificación es correcto con respecto a su nueva especificación, que va a hablar sobre *todos* los posibles antecesores inmediatos de cada vértice  $v$ .

Otra modificación que podríamos hacer es una que sólo cuente cuántos caminos hay:

```

1: if  $d[v] < d[u] + w(u, v)$  then
2:    $d[v] \leftarrow d[u] + w(u, v)$ 
3:    $p[v] \leftarrow p[u]$ 
4: else if  $d[v] = d[u] + w(u, v)$  then
5:    $p[v] \leftarrow p[v] + p[u]$ 
6: end

```

donde  $p[s] = 1$  es el valor inicial, y para todo otro vértice  $v$ ,  $p[v] = 0$  inicialmente.

Esto requiere una demostración aún más extensa. Tendremos que argumentar que el conjunto de caminos desde  $s$  hasta  $v$ , se partitiona de forma disjunta en los caminos desde  $s$  hasta  $u$ , y luego la arista  $(u, v)$ , para todos los vértices  $u$  tal que  $\delta(s, v) = \delta(s, u) + w(u, v)$ .

- Podemos usar el algoritmo de Dijkstra canónico. Luego, podemos reconstruir  $G'$ , usando sólo  $d$ , el array de distancias que devuelve el algoritmo de Dijkstra. Hay que escribir más código para esto, pero el código cuya correctitud debemos demostrar es simple y corto.

Como usamos el primer método en el ejercicio anterior, en este vamos a usar el segundo.

El algoritmo que vamos a usar es el siguiente:

```
from collections import defaultdict
def F(G, w, s):
    d, _ = Dijkstra(G, w, s)
    (V, E) = G
    preds = defaultdict(list)
    for (u, v) in E:
        if d[v] == d[u] + w(u, v):
            preds[v].append(u)

    n = len(V)
    x = [-1 for _ in range(n)]
    x[s] = 1
    def rec(v):
        if x[v] == -1:
            x[v] = sum(rec(u) for u in preds[v])
        return x[v]

    for v in range(n): rec(v)
    return x
```

Demostremos que este algoritmo es correcto. La función `rec` es meramente usar programación dinámica top-down, en la siguiente función recursiva:

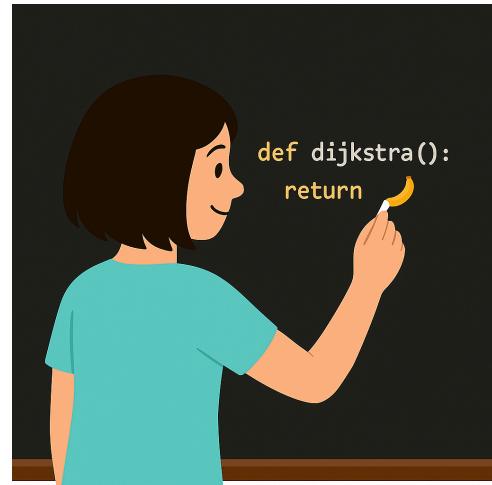
$$f(v) = \begin{cases} 1 & \text{si } v = s \\ \sum_{\substack{(u,v) \in E, \\ d[v] = d[u] + w(u,v)}} f(u) & \text{si no} \end{cases}$$

Y al correr `f(v)` para todo `v in range(n)`, estamos asegurándonos de llenar el cache `x` para todos los vértices, con lo cual devolvemos un array cuya  $i$ -ésima coordenada contiene  $f(i)$ . La semántica que le asignamos a la función  $f$  es que  **$f(i)$  devuelve el número de caminos mínimos desde  $s$  hasta  $i$** .

### ⚠️ Advertencia

Recordemos que no tiene sentido probar que una función «es correcta» sin decir cuál es su semántica. Una función que siempre devuelve el string "banana" es correcta si nuestra semántica es «una función que siempre devuelve el string "banana"».

**Siempre** que vamos a probar que una función es correcta, **tenemos** que definir cuál es su semántica. Por esto ven especificación antes de correctitud formal de algoritmos.



Probemos que  $f$  es correcta con respecto a esa semántica. Como es una función recursiva, vamos a usar inducción. ¿En qué vamos a hacer inducción? En lo que está decreciendo en cada llamada: El máximo número de aristas en un camino mínimo entre  $s$  y  $v$ . Por comodidad, llamemos  $C(v)$  al conjunto de caminos mínimos entre  $s$  hasta  $v$ , y llamemos  $M(v) = \max_{P \in C(v)} \{|P|\}$ , el máximo número de aristas en cualquier camino mínimo desde  $s$  hasta  $v$ .

Sea  $P(i)$ : Para todo  $v \in V$ , tal que  $M(v) = i$ ,  $f(v) = |C(v)|$ .

1. Caso base,  $i = 0$ . Si  $v$  es un vértice tal que  $M(v) = 0$ , entonces hay un camino de longitud 0 desde  $s$  hasta  $v$ . Luego,  $v$  es  $s$ . Como hay un único camino desde  $s$  hasta  $s$ , vemos que  $|C(s)| = |\{[]\}| = 1$ . Como  $f$  precisamente devuelve 1,  $f$  es correcta para este caso.
2. Paso inductivo. Sea  $i > 0 \in \mathbb{N}$ . Podemos asumir  $P(k)$  para todo  $k < i$ , y queremos ver  $P(i)$ . Sea  $v \in V$  tal que  $M(v) = i$ . Como  $M(v) = i > 0$ , sabemos que  $v \neq s$ , pues  $M(s) = 0$ . Luego, todo camino mínimo  $Q$  desde  $s$  hasta  $v$ , pasa por algún vértice  $u$  anterior a  $v$ . Notar que  $u$  puede ser  $s$  mismo, si hay un camino mínimo desde  $s$  hasta  $v$  que es puramente  $[(s, v)]$ . También, como los caminos mínimos tienen subestructura óptima, el prefijo de  $Q$  tiene que ser un camino mínimo desde  $s$  hasta ese vértice  $u$ . Notemos que tendremos  $\delta(s, u) + w(u, v) = \delta(s, v)$ , por definición de  $\delta$  y que  $Q$  y su prefijo hasta  $u$  son caminos mínimos. Finalmente, por la postcondición del algoritmo de Dijkstra, sabemos que  $d[v] = \delta(s, v)$  para todo  $v \in V$ .

$$\begin{aligned} C(v) &= \{Q \mid Q \in C(v)\} \\ &= \{Q' + [(u, v)] \mid u \in V, Q' \in C(u), Q' + [(u, v)] \in C(v)\} \\ &= \{Q' + [(u, v)] \mid u \in V, Q' \in C(u), \delta(s, u) + w(u, v) = \delta(s, v)\} \\ &= \{Q' + [(u, v)] \mid u \in V, Q' \in C(u), d[u] + w(u, v) = d[v]\} \end{aligned}$$

Luego,

$$\begin{aligned}
|C(v)| &= |\{Q' + [(u, v)] \mid u \in V, Q' \in C(u), d[u] + w(u, v) = d[v]\}| \\
&= \sum_{\substack{u \in V \\ Q' \in C(u) \\ d[u] + w(u, v) = d[v]}} 1 \\
&= \sum_{\substack{u \in V \\ d[u] + w(u, v) = d[v]}} |C(u)|
\end{aligned}$$

Si probamos que  $M(u) < M(v)$  para todo  $u$  en esa sumatoria, probamos que  $f$  es correcta, porque podemos usar la hipótesis inductiva  $P(M(u))$  para saber que  $f(u) = |C(u)|$ , y llegar a

$$|C(v)| = \sum_{\substack{u \in V \\ d[u] + w(u, v) = d[v]}} f(u)$$

que probaría  $P(i)$ , pues el lado derecho es la definición exacta de  $f$ , y estaríamos probando  $|C(v)| = f(v)$ . Sea  $u \in V$  tal que  $d[u] + w(u, v) = d[v]$ , y tomemos un camino  $Q$  de máximo número de aristas en  $C(u)$ . Por cómo definimos  $M(u)$ , tenemos que  $M(u) = |Q|$ . Vemos que  $Q' = Q + [(u, v)]$  es un camino en  $C(v)$ , y que  $|Q'| = |Q| + 1$ . Como  $M(v) = \max_{R \in C(v)} \{|R|\}$ , y  $Q'$  es un tal  $R$ , tenemos que  $M(v) \geq |Q'| = |Q| + 1 > M(u)$ . Luego  $M(u) < M(v)$ , y entonces podemos usar la hipótesis inductiva,  $P(M(u))$ , que nos deja concluir que  $f$  es correcta.  $\square$

### Ejercicio 6.8.11

Sea  $G = (V, E)$  un grafo dirigido pesado, y  $s \in V$ .

Queremos encontrar el camino más ligero entre  $s$  y cada  $v \in V$ . Dar una modificación del algoritmo de Bellman-Ford tal que:

- Si  $G$  no tiene ciclos de peso negativo alcanzables desde  $s$ , devuelve la distancia entre  $s$  y  $v$ , para todo  $v \in V$ .
- Si  $G$  tiene un ciclo de peso negativo alcanzable desde  $s$ , devuelve un tal ciclo.

Demostrar que es correcto. El algoritmo debería seguir haciendo  $O(|V| |E|)$  operaciones en el peor caso.



*Demostración.* Recordemos el invariante del algoritmo de Bellman-Ford para caminos mínimos en grafos pesados.

```

1: procedure BELLMAN-FORD( $G = (V, E)$ ,  $s \in V$ ,  $w : E \rightarrow \mathbb{R}$ )
2:    $d[v] \leftarrow \infty \forall v \in V$ 
3:    $d[s] \leftarrow 0$ 
4:    $p[v] \leftarrow \text{NULL} \forall v \in V$ 
5:   for  $i = 1$  to  $|V| - 1$  do
6:     for  $(u, v) \in E$  do
7:       if  $d[v] > d[u] + w(u, v)$  then
8:          $d[v] \leftarrow d[u] + w(u, v)$ 
9:          $p[v] \leftarrow u$ 
10:      end
11:    end

```

```

12:   end
13:   return  $d, p$ 
14: end

```

El invariante de este algoritmo es que al terminar la  $i$ -ésima iteración del ciclo externo, para todo  $v \in V$ ,  $d[v]$  es la distancia mínima desde  $s$  hasta  $v$  usando sólo caminos de a lo sumo  $i$  aristas. Luego, al terminar el algoritmo, tenemos que  $d[v]$  es la distancia mínima desde  $s$  hasta  $v$ , usando a lo sumo  $|V| - 1$  aristas, para todo  $v \in V$ .

Si no hay ciclos de peso negativo alcanzables desde  $s$ , entonces la distancia «usando a lo sumo  $|V| - 1$  aristas» es lo mismo que la distancia, dado que si nuestro camino tiene más de  $|V| - 1$  aristas, tiene algún ciclo, que si fuera de peso no-negativo podríamos simplemente remover, y no-empeorar la distancia.

Si hay ciclos de peso negativo alcanzables desde  $s$ , entonces las distancias hacia algunos vértices no están bien definidas.

#### Lema 6.8.12

Hay un ciclo negativo alcanzable desde  $s$  si y sólo si al hacer una iteración más del ciclo externo, en algún momento cambiamos  $d$ .

#### Demostración.

- $\Leftarrow$ ) Sea  $n = |V|$ . Por el invariante del algoritmo, al terminar la  $n - 1$ -ésima iteración, tenemos en  $d[v]$  las distancias desde  $s$  hasta  $v$ , usando a lo sumo  $n - 1$  aristas. Si hacemos una iteración más, y cambiamos  $d[v]$ , entonces ocurrió que  $d[v] > d[u] + w(u, v)$  para algún  $u$  tal que  $(u, v) \in E$ . Por el invariante del algoritmo, sabemos que  $d[v]$  es la distancia mínima desde  $s$  hasta  $u$ , usando a lo sumo  $n - 1$  aristas. Luego, si  $d[v] > d[u] + w(u, v)$ , entonces el mejor camino desde  $s$  hasta  $v$  usando  $n - 1$  aristas es más pesado que un camino que va hasta  $u$  usando a lo sumo  $n - 1$  aristas, y luego usa la arista  $(u, v)$ . Este camino necesariamente debe tener más de  $n - 1$  aristas, pues por inducción  $d[v]$  ya tiene la menor distancia entre todos los caminos que usan a lo sumo  $n - 1$  aristas.

Un camino  $P$  de más de  $n - 1$  aristas debe tener algún ciclo  $C$ . Si ese ciclo fuese de peso no-negativo, podríamos removerlo, y obtener  $P - C$ , que sigue empezando en  $s$ , terminando en  $v$ , y tiene peso igual a  $P$ , es decir  $d[u] + w(u, v)$ . Pero entonces  $P - C$  es un camino de a lo sumo  $n - 1$  aristas, y tiene peso  $d[u] + w(u, v) < d[v]$ , lo cual no puede pasar porque  $d[v]$  era la mínima distancia entre  $s$  y  $v$  usando caminos de a lo sumo  $n - 1$  aristas.

Luego  $P$  debe tener algún ciclo de peso negativo, y como  $P$  empieza en  $s$  y termina en  $v$ , entonces  $v$  es alcanzable desde  $s$  por un ciclo  $C$  de peso negativo.

- $\Rightarrow$ ) Sea  $C = [v_0, \dots, v_k = v_0]$  un ciclo de peso negativo alcanzable desde  $s$ . Luego,  $\sum_{i=1}^k w(v_{i-1}, v_i) < 0$ . Supongamos que no actualizamos ningún  $d$  en la  $n$ -ésima iteración. Entonces, en esa iteración tenemos que  $d[v_i] \leq d[v_{i-1}] + w(v_{i-1}, v_i)$  para todo  $1 \leq i \leq k$ . Sumando todas estas desigualdades, tenemos

$$\sum_{i=1}^k d[v_i] \leq \sum_{i=1}^k d[v_{i-1}] + \sum_{i=1}^k w(v_{i-1}, v_i)$$

Vemos también que como  $v_0 = v_k$ , la suma  $\sum_{i=1}^k d[v_i]$  es igual a la suma  $\sum_{i=1}^k d[v_{i-1}]$ , y en ambas aparecen todos los vértices de  $C$  una sola vez. Luego, restando esta suma de cada lado, obtenemos

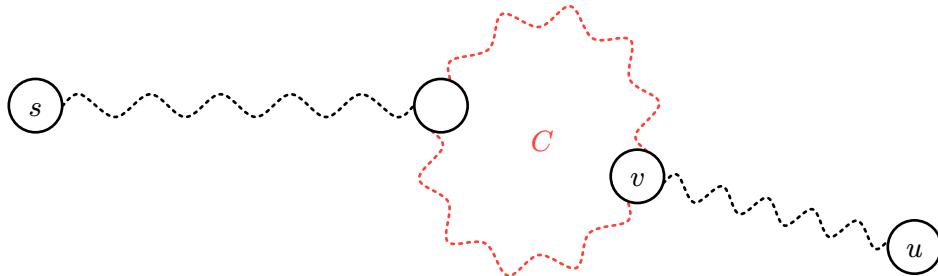
$$0 \leq \sum_{i=1}^k w(v_{i-1}, v_i)$$

Pero esto no puede pasar, pues sabemos que  $C$  tiene peso negativo. Luego, en la  $n$ -ésima iteración, tuvimos que haber actualizado  $d[v]$  para algún  $v \in C$ .

□

Esto nos dice que para saber si hay algún ciclo de peso negativo alcanzable desde  $s$ , podemos hacer una iteración más del ciclo externo, y si en algún momento cambiamos  $d[v]$ , entonces  $v$  es alcanzable desde  $s$  por un ciclo de peso negativo. Si no cambiamos nada, entonces no hay ciclos de peso negativo alcanzables desde  $s$ . Más aún, si ese ciclo de peso negativo  $C$  existe, entonces debemos actualizar  $d[v]$  en la  $n$ -ésima iteración para algún  $v \in C$ .

Notemos que no necesariamente vamos a cambiar *sólo* los  $d[v]$  donde  $v \in C$ . Como recorremos  $e \in E$  en algún orden, puede ser que primero actualizamos  $d[v]$  para algún  $v \in C$ , y luego actualizamos  $d[u]$  para algún  $u$  alcanzable desde  $v$ . Lo que vamos a saber es que si actualizamos  $d[u]$ , entonces  $u$  es alcanzable desde algún ciclo negativo, que a su vez es alcanzable desde  $s$ .



Luego, si actualizamos  $d[u]$  en la  $n$ -ésima iteración, podemos seguir la cadena de padres,  $p$ , y vamos a llegar hasta algún vértice (en este caso  $v$ ) que pertenece a  $C$ .

Cuando actualizamos  $d[v]$  en la última iteración, como el mejor camino desde  $s$  hasta  $v$  ahora tiene un ciclo (pues tiene más de  $n - 1$  aristas), si seguimos la cadena de padres  $p$ , vamos a encontrar que  $v \rightsquigarrow v$ , pues el mejor camino usa un ciclo que involucra a  $v$ . Luego, podemos retroceder  $n$  pasos usando  $p$ , y vamos a caer en  $C$ , porque la distancia entre  $v$  y  $u$  es a lo sumo  $n - 1$  aristas, y luego siguiendo  $p[u], p[p[u]], \dots$ , vamos a llegar a  $v$ . Una vez que llegamos a  $v$ , vamos a pasar por  $C$  una y otra vez si seguimos siguiendo  $p$ .

Luego, para recuperar un ciclo de peso negativo alcanzable desde  $s$ , podemos usar el siguiente algoritmo:

1: **procedure** BELLMAN-FORD-NEGATIVE-CYCLE( $G = (V, E)$ ,  $s \in V$ ,  $w : E \rightarrow \mathbb{R}$ )

```

2:    $d[v] \leftarrow \infty \forall v \in V$ 
3:    $d[s] \leftarrow 0$ 
4:    $p[v] \leftarrow \text{NULL} \forall v \in V$ 
5:   for  $i = 1$  to  $|V| - 1$  do
6:     for  $(u, v) \in E$  do
7:       if  $d[v] > d[u] + w(u, v)$  then
8:          $d[v] \leftarrow d[u] + w(u, v)$ 
9:          $p[v] \leftarrow u$ 
10:      end
11:    end
12:  end
13:   $z \leftarrow \text{NULL}$ 
14:  for  $(u, v) \in E$  do
15:    if  $d[v] > d[u] + w(u, v)$  then
16:       $z \leftarrow v$ 
17:    end
18:  end
19:  if  $z = \text{NULL}$  then
20:    return  $d, p$ 
21:  end
22:  for  $i = 1$  to  $|V|$  do
23:     $z \leftarrow p[z]$ 
24:  end
25:   $z_0 \leftarrow z$ 
26:   $C \leftarrow [z_0]$ 
27:  while  $p[z] \neq z_0$  do
28:     $z \leftarrow p[z]$ 
29:     $C \leftarrow C + [z]$ 
30:  end
31:  return  $C$ 
32: end

```

□

### Ejercicio 6.8.13

Es nuestro primer día trabajando en un banco, en el departamento de comercio internacional. Sabemos que hay  $n$  monedas en el mundo, y tenemos una tabla  $Q \in \mathbb{R}^{n \times n}$ , que nos dice que si tenemos 1 unidad de la moneda  $i$ , se puede intercambiar por  $0 < Q_{i,j}$  unidades de la moneda  $j$ .

Queremos saber si existe una secuencia  $S = (s_1, s_2, \dots, s_k)$  de intercambios de monedas que podemos hacer, tal que al intercambiar una moneda  $s_1$  por la moneda  $s_2$ , y luego  $s_2$  por  $s_3, \dots$ , y luego  $s_k$  por  $s_1$ , terminamos con más dinero que con el que empezamos.

Dar un algoritmo que determine si es posible hacer esto. El algoritmo debería tardar tiempo  $O(n^3)$  en el peor caso. Demostrar que es correcto.

*Solución.* Podemos construir un grafo dirigido pesado  $G = (V, E)$ , donde  $V = \{1, \dots, n\}$  son las monedas, hay una arista entre todo par de vértices, y el peso de la arista  $(v_i, v_j)$  es  $w(v_i, v_j) = -\log Q_{i,j}$ .

Si conseguimos un ciclo de peso negativo en  $G$ , entonces sabremos que existe una cadena de transacciones que podemos hacer, yendo de la moneda  $a_1$  a  $a_2$ , dots,  $a_k$ , y finalmente a  $a_1$ , tal que:

$$\begin{aligned} w(a_k, a_1) + \sum_{i=1}^{k-1} w(a_i, a_{i+1}) &< 0 \\ \log Q_{a_k, a_1} + \sum_{i=1}^{k-1} \log Q_{a_i, a_{i+1}} &> 0 \\ Q_{a_k, a_1} \times \prod_{i=1}^{k-1} Q_{a_i, a_{i+1}} &> 1 \end{aligned}$$

que es precisamente lo que significa obtener más dinero que con el que empezamos.

Podemos transformar  $A$  en  $B$ , con  $B_{i,j} = -\log A_{i,j}$ , y luego usar el algoritmo de Floyd-Warshall en  $B$ . Luego, si  $B_{i,i} < 0$  para algún  $i$ , entonces existe un ciclo de peso negativo (que incluye a  $v_i$ ). Como vimos arriba, la existencia este ciclo implica que existe una manera de intercambiar monedas que nos deja con más dinero del que empezamos.

#### Ejercicio 6.8.14

Sea  $G = (V, E)$  un grafo dirigido pesado, con  $V = \{v_1, \dots, v_n\}$ , y sin ciclos de peso negativo. Se tiene una matriz  $A \in \mathbb{R}^{n \times n}$ , donde  $A_{i,j}$  es la distancia entre  $v_i$  y  $v_j$  en  $G$ .

Se agrega un vértice  $v_{n+1}$  a  $G$ , con algunas aristas, obteniendo  $G'$ , con  $n+1$  vértices.

Dar un algoritmo que calcule las distancias entre **todo** par de vértices en  $G'$ . El algoritmo debe hacer  $O(n^2)$  operaciones en el peor caso. Demostrar que es correcto.

*Solución.* Vamos a construir una matriz  $B \in \mathbb{R}^{(n+1) \times (n+1)}$ , donde  $B_{i,j}$  es la distancia entre  $v_i$  y  $v_j$  en  $G'$ . Realizamos el siguiente procedimiento:

```

1: procedure ADD-VERTEX( $A \in \mathbb{R}^{n \times n}$ ,  $w : E \rightarrow \mathbb{R}$ )
2:    $B \in \mathbb{R}^{(n+1) \times (n+1)} \leftarrow \text{null}$ 
3:   for  $l = 1$  to  $n$  do
4:      $B_{n+1,l} \leftarrow \min_{1 \leq j \leq n} w(n+1, j) + A_{j,l}$ 
5:      $B_{l,n+1} \leftarrow \min_{1 \leq j \leq n} w(j, n+1) + A_{l,j}$ 
6:   end
7:    $B_{n+1,n+1} \leftarrow \min_{1 \leq j \leq n} B_{n+1,j} + B_{j,n+1}$ 
8:   for  $i = 1$  to  $n$  do
9:     for  $j = 1$  to  $n$  do
10:       $B_{i,j} \leftarrow \min(A_{i,j}, B_{i,n+1} + B_{n+1,j})$ 
11:    end
12:  end
13: return  $B$ 
14: end
```

*Demostración.* Queremos demostrar que el algoritmo correctamente computa las distancias entre todo par de vértices en  $G'$ . Inicialmente, sabemos que  $A$  contiene las mínimas distancias entre todo par de vértices en  $G$ .

Veamos por qué cada una de las partes del algoritmo es correcto.

1. El camino más corto para ir desde  $v_{n+1}$  hasta  $v_l$ , para cada  $1 \leq l \leq n$ , empieza con alguna arista  $(v_{n+1}, v_j)$ , con  $1 \leq j \leq n$ , y luego va desde  $v_j$  hasta  $v_l$  usando el camino más corto en  $G$  (es decir, sin volver a pasar por  $v_{n+1}$ , pues  $v_{n+1}$  no está en  $G$ , sino en  $G'$ ). Por eso es correcto asignar a  $B_{n+1,l}$  el valor  $\min_{1 \leq j \leq n} w(n+1, j) + A_{j,l}$ , dado que  $A$  contiene las distancias en  $G$ .
2. Algo análogo sucede con  $B_{l,n+1}$ .
3. La mejor forma de ir desde  $v_{n+1}$  hasta  $v_{n+1}$  es ir desde  $v_{n+1}$  hasta algún vértice  $v_j$  con  $1 \leq j \leq n$ , usando el camino más corto en  $G$ , y luego volver a  $v_{n+1}$ , usando el camino más corto en  $G$ . Por eso es correcto asignar a  $B_{n+1,n+1}$  el valor  $\min_{1 \leq j \leq n} B_{n+1,j} + B_{j,n+1}$ .
4. La mejor forma de ir desde  $v_i$  hasta  $v_j$ , con  $1 \leq i, j \leq n$ , puede o usar o no usar  $v_k$ . Si no lo usa, entonces es simplemente  $A_{i,j}$ . Si lo usa, entonces va desde  $v_i$  hasta  $v_k$ , usando el camino más corto en  $G$ , y luego desde  $v_k$  hasta  $v_j$ , también usando el camino más corto en  $G$ . Luego, la mejor forma de ir desde  $v_i$  hasta  $v_j$  es  $\min(A_{i,j}, B_{i,n+1} + B_{n+1,j})$ , que es lo que asignamos a  $B_{i,j}$ .

Este algoritmo inmediatamente nos da un algoritmo para distancias mínimas entre todo par de vértices en un grafo. Empezamos con un subgrafo generador por el primer vértice, y una matriz trivial de distancias (un único escalar, 0). Luego,  $n - 1$  veces agregamos un vértice más, con las aristas que corresponde.

Nuestro algoritmo usa  $O(n^2)$  operaciones en todos los casos, al ser una simple composición de ciclos. Luego, el algoritmo general para caminos mínimos entre todo par de vértices, que hace ADD-VERTEX  $n - 1$  veces, va a usar  $O(n^3)$  operaciones en el peor caso. Este algoritmo se conoce como el algoritmo de Dantzig[4]. □

## 6.9 Planaridad

### Definición 6.9.1

Un grafo  $G$  es **planar** cuando existe una forma de dibujarlo en el plano, de tal forma que las aristas no se crucen. ♣

### Teorema 9

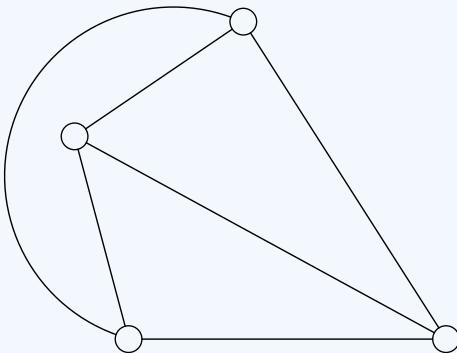
En un grafo planar con  $m > 1$  aristas, toda cara tiene al menos 3 aristas en su borde. Esto incluye la cara exterior. ♥

### Ejercicio 6.9.2 (Formula de Euler)

Sea  $G$  un grafo conexo y planar, con  $n$  vértices,  $m$  aristas, y dibujado con  $r$  caras (incluyendo una cara exterior).

Probar que  $r = m - n + 2$ .

Por ejemplo, el siguiente grafo tiene  $n = 4$  vértices,  $m = 6$  aristas, y  $r = 4$  caras:



*Demostración.* Vamos a probar esto por inducción en  $r$ , el número de caras. Sea  $P(r)$  : Todo grafo  $G = (V, E)$  conexo y planar con  $r$  caras cumple que  $|V| - |E| + r = 2$ .

1. Caso base,  $P(1)$ . Sea  $G = (V, E)$  un grafo conexo planar con  $r = 1$  caras, y sean  $n = |V|$  y  $m = |E|$ . Como  $r = 1$ , no podemos tener ciclos, pues un ciclo generaría una cara, y ya tenemos una (la externa). Luego, al ser  $G$  conexo y acíclico,  $G$  es un árbol. Luego,  $m = n - 1$ , y tenemos  $n - m + r = n - (n - 1) + 1 = 2$ , que muestra  $P(1)$ .
2. Paso inductivo. Sabemos  $P(r)$ , queremos probar  $P(r + 1)$ . Sea  $G = (V, E)$  un grafo conexo y planar, con  $r + 1 \geq 2$  caras. Si  $G$  fuera un árbol,  $r = 1$ , luego  $G$  no es un árbol, y al ser conexo, tiene un ciclo  $C$ . Sea  $e \in C$  cualquiera arista en  $C$ . Sea ahora  $G' = (V, E - \{e\})$ . Como  $e$  pertenecía a un ciclo,  $G'$  sigue siendo conexo, con  $|E - \{e\}| = |E| - 1$ . Al haber roto un ciclo sacando  $e$ ,  $G'$  tiene una cara menos que  $G$ . Luego, usando  $P(r)$ , vemos que  $|V| - (|E| - 1) + r = 2$ . Esto nos dice que  $|V| - |E| + (r + 1) = 2$ , que es prueba  $P(r + 1)$ .

Luego probamos  $P(r)$  para todo  $r \in \mathbb{N}$ ,  $r \geq 1$ . Como todo grafo planar tiene un número positivo de caras, esto prueba que todo grafo conexo planar cumple la ecuación.  $\square$

### Ejercicio 6.9.3

Sea  $G$  un grafo planar de  $n$  vértices y  $m$  aristas, con  $n \geq 3$ .

Probar que  $m \leq 3n - 6$ .

*Demostración.* Vamos a probar esto para grafos conexos. Si nos dan un grafo no conexo, podemos agregarle aristas hasta hacerlo conexo, sin perder planaridad, y vamos a ver que la desigualdad sigue valiendo. Como estamos aumentando  $m$ , sin cambiar  $n$ , valía antes de agregar las aristas.

Sea entonces  $G = (V, E)$  un grafo planar conexo, con  $n = |V|$ ,  $m = |E|$ , y  $r$  caras. Si  $m \leq 2$ , esto es cierto pues  $3n - 6 \geq 3$ . De otra manera, podemos usar el Teorema 9, y sabemos que toda cara tiene al menos 3 aristas en su borde. Para cada  $i \in \mathbb{N}$ ,  $f_i$  el número de caras con exactamente  $i$  aristas en  $G$ . Luego, la siguiente suma va a contar a cada arista dos veces (una vez en cada una de las dos caras que parte, o dos veces en la misma cara si la arista está enteramente en una cara):

$$\begin{aligned} 2m &= 1f_2 + 2f_3 + 3f_4 + \dots \\ &= \sum_{i \in \mathbb{N}} if_i \\ &= \sum_{\substack{i \in \mathbb{N}, \\ i \geq 3}} if_i \\ &\geq \sum_{i \in \mathbb{N}} 3f_i \\ &= 3 \sum_{i \in \mathbb{N}} f_i \\ &= 3r \end{aligned}$$

Por el teorema de Euler,  $n - m + r = 2$ . Luego,  $r = m - n + 2$ , y luego  $3r = 3m - 3n + 6$ . Usando la desigualdad de arriba,  $3m - 3n + 6 \leq 2m$ . Simplificando,  $m - 3n + 6 \leq 0$ , y luego  $m \leq 3n - 6$ .  $\square$

#### Ejercicio 6.9.4

Sea  $G$  un grafo conexo y planar, y sea  $\delta(G)$  el mínimo grado entre todos los vértices de  $G$ .

Probar que  $\delta(G) \leq 5$ .



*Demostración.* Sea  $G = (V, E)$  un grafo conexo y planar, con  $n = |V|$ ,  $m = |E|$ , y  $r$  caras.

Si  $n \leq 2$ , la propiedad es obvia.

Si  $n \geq 3$ , contemos:

$$2m = \sum_{v \in V} d_G(v) \geq 6n$$

La primera igualdad es porque en todo grafo tenemos  $\sum_{v \in V} d_G(v) = 2m$ , y la segunda es porque  $d_G(v) \geq 6$  para todo  $v \in V$ , por el enunciado.

Por el ejercicio anterior, vemos que  $m \leq 3n - 6$ . Luego,  $2m \leq 6n - 12$ . Pero entonces tenemos  $6n \leq 2m \leq 6n - 12$ , que no tiene sentido.

Luego no puede ser que todos los vértices tienen grado al menos 6, y tiene que haber algún vértice con grado a lo sumo 5.  $\square$

## 6.10 Coloreo

### Definición 6.10.1

Sea  $G = (V, E)$  un grafo. Un **coloreo de  $G$  con  $k$  colores** es una función  $f : V \rightarrow \{1 \dots k\}$ , tal que para todo  $(u, v) \in E$ ,  $f(u) \neq f(v)$ .

Si existe un coloreo de  $G$  con  $k$  colores, se dice que  $G$  es  **$k$ -coloreable**.

El **número cromático** de  $G$  es el mínimo  $k$  tal que  $G$  es  $k$ -coloreable, y se denota  $\chi(G)$ .

### Ejercicio 6.10.2

Dado un  $n \in \mathbb{N}$ , el grafo  $C_n$  tiene  $n$  vértices  $v_1, \dots, v_n$ , y hay una arista entre  $v_i$  y  $v_{(i+1) \bmod n}$  para todo  $0 \leq i \leq n$ .

¿Para cuáles  $n \in \mathbb{N}$ ,  $C_n$  es 2-coloreable?

*Demostración.* Sea  $C_n = (V, E) = [v_1, \dots, v_n]$  el ciclo de  $n$  vértices. Supongamos que tenemos  $f : V \rightarrow \{0, 1\}$  que colorea a  $C_n$  con 2 colores. Luego, claramente los valores de  $f(v_1), f(v_2), f(v_3), \dots$  tienen que alternarse, pues si no no sería un coloreo válido. La única restricción va a ser qué pasa cuando el ciclo vuelve a  $v_1$ . Supongamos sin pérdida de generalidad que  $f(v_1) = 1$ , entonces vamos a tener  $f(v_i) = i \bmod 2$ . Entonces, si tuvieramos un ciclo de longitud impar, tendríamos  $f(v_n) = n \bmod 2 = 1$ , pero  $f(v_1) = 1$ , y  $\{v_n, v_1\} \in E$ . Luego, si queremos que  $C_n$  sea 2-coloreable, debemos tener  $n \equiv 0 \pmod{2}$ .  $\square$

### Ejercicio 6.10.3

Probar que el número cromático de todo árbol es menor o igual a 2.

*Demostración.* Sea  $T$  un árbol con  $n$  vértices.

- Si  $n = 1$ , puedo colorear al único vértice de un color, y tengo un 1-coloreo. No existen 0-coloreos, luego el número cromático de  $T$  es uno.
- Si  $n > 1$ , entonces al ser  $T$  bipartito, podemos tomar las dos particiones  $U, W$  de sus vértices, colorear  $U$  de un color y  $W$  de otro, y obtenemos un 2-coloreo válido. Como un árbol es conexo, y hay más de un vértice, debe existir al menos una arista  $e = \{u, w\}$ , con  $u \in U$  y  $w \in W$ . Esa arista previene que  $u$  y  $w$  tengan el mismo color, luego no existen 1-coloreos. Como se puede colorear con 2 colores, y no con 1 color, el número cromático de  $T$  es 2.

$\square$

### Ejercicio 6.10.4 (Teorema de Ramsey)

Probar que toda forma de colorear las **aristas** de  $K_6$  con dos colores, azul y rojo, tiene un triángulo rojo o un triángulo azul.

*Demostración.* Sea  $G = (V, E) = K_6$  el grafo completo en 6 vértices, y sea  $f : E \rightarrow \{\bullet, \circ\}$  una manera de colorear  $E$  con dos colores. Sea  $v \in V$ . Como  $d_G(v) = 5$ , tiene o 3 aristas incidentes  $\bullet$ , o aristas incidentes  $\circ$ . Sin pérdida de generalidad, asumamos que tiene tres aristas incidentes  $\bullet$ . Sean  $x, y, z$  los vértices que unen a estas aristas con  $v$ . Luego,  $f(\{v, x\}) = f(\{v, y\}) = f(\{v, z\}) = \bullet$ . Si  $f(\{x, y\}) = f(\{y, z\}) = f(\{z, x\}) = \bullet$ , encontramos un triángulo  $\bullet$ , y terminamos. De otra forma, existe una arista  $\circ$  entre algunos de esos tres vértices. Sin pérdida de generalidad, asumamos que es  $f(\{x, y\}) = \circ$ . Entonces tenemos un triángulo  $\bullet, (\{v, x\}, \{x, y\}, \{y, v\})$ , y terminamos.  $\square$

#### Ejercicio 6.10.5

Sea  $G$  un grafo de  $n$  vértices, denotemos  $\alpha(G)$  el máximo tamaño de un conjunto independiente en  $G$ .

Probar que  $\frac{n}{\alpha} \leq \chi$ .



*Demostración.* Sea  $G = (V, E)$ . Nombramos por comodidad  $\chi = \chi(G)$ , y  $\alpha = \alpha(G)$ , y  $n = |V|$ . y consideremos un colooreo óptimo  $f : V \rightarrow \{1, \dots, \chi\}$ . Sean  $S_i = f^{-1}(i) \subseteq V$  los conjuntos de cada color, para  $1 \leq i \leq \chi$ . Si  $u, v \in S_i$ , entonces  $\{u, v\} \notin E$ , puesto que  $f$  es un colooreo válido. Luego  $S_i$  es un conjunto independiente.

Como  $V = \bigsqcup_{1 \leq i \leq \chi} S_i$ , entonces  $n = \sum_{i=1}^{\chi} |S_i|$ . Como cada  $S_i$  es un conjunto independiente,  $|S_i| \leq \alpha$ . Luego,  $n \leq \sum_{i=1}^{\chi} \alpha = \chi\alpha$ , y luego  $\frac{n}{\alpha} \leq \chi$ .  $\square$

#### Ejercicio 6.10.6

Sea  $G = (V, E)$  un grafo con  $n = |V|$  vértices, y  $\overline{G}$  su complemento.

Demostrar que  $\chi(G) + \chi(\overline{G}) \leq n + 1$ .



*Demostración.* Por inducción. Si  $n = 1$ ,  $G = \overline{G}$ , y  $\chi(G) = \chi(\overline{G})$ , luego  $\chi(G) + \chi(\overline{G}) = 1 + 1 = 2 \leq 1 + 1$ . En el paso inductivo, si  $n > 1$ , podemos tomar  $G$  y sacarle un vértice  $v$ , obteniendo  $G' = G - v$ . También llamamos  $\overline{G}' = \overline{G} - v$ . Usamos la hipótesis inductiva en  $n - 1$ , y obtenemos un colooreo con  $k$  colores de  $G'$ , y un colooreo con  $l$  colores de  $\overline{G}'$ , tal que  $k + l \leq (n - 1) + 1 = n$ .

1. Si  $d_G(v) < k$ , extendemos el colooreo de  $G'$  a un colooreo de  $G$  con  $k$  colores. Coloreamos  $\overline{G}$  usando el colooreo de  $l$  colores de  $\overline{G}'$ , y usando un nuevo color para  $v$  que agregamos. Luego,  $\chi(G) + \chi(\overline{G}) \leq (n - 1 + 1) + 1 = n + 1$ , que es lo que queríamos demostrar.
2. Si  $d_G(v) \geq k$ , entonces  $d_{\overline{G}}(v) \leq n - k = l - 1 < l$ , y podemos extender el  $l$ -colooreo de  $\overline{G}'$  a un  $l$ -colooreo de  $\overline{G}$ . En  $G$ , usamos el  $k$ -colooreo de  $G'$ , y usamos un color nuevo para  $v$ . Luego,  $\chi(\overline{G}) + \chi(G) \leq l + k + 1 = n + 1$ , que es lo que queríamos demostrar.



### Ejercicio 6.10.7

Sea  $G = (V, E)$  un grafo con  $n = |V|$  vértices, y  $\bar{G}$  su complemento.

Demostrar que  $\chi(G) + \chi(\bar{G}) \geq 2\sqrt{n}$ .



*Demostración.* Consideremos  $G \cup \bar{G} = K_n$ , el grafo completo. Si  $\chi(G)$ -coloreamos  $G$  con  $c$ , y  $\chi(\bar{G})$ -coloreamos  $\bar{G}$  con  $c'$ , entonces podemos construir un  $\chi(G)\chi(\bar{G})$ -coloreo de  $K_n$ , coloreando a cada vértice  $v \in V$  con  $(c(v), c'(v))$ . Como  $\chi(K_n) = n$ , entonces  $n \leq \chi(G)\chi(\bar{G})$ .

Ahora bien:

$$\begin{aligned}\left(\chi(G) - \chi(\bar{G})\right)^2 &\geq 0 \\ \chi(G)^2 - 2\chi(G)\chi(\bar{G}) + \chi(\bar{G})^2 &\geq 0 \\ \chi(G)^2 + 2\chi(G)\chi(\bar{G}) + \chi(\bar{G})^2 &\geq 4\chi(G)\chi(\bar{G}) \\ \left(\chi(G) + \chi(\bar{G})\right)^2 &\geq 4\chi(G)\chi(\bar{G}) \geq 4n \\ \left(\chi(G) + \chi(\bar{G})\right)^2 &\geq 4n \\ \chi(G) + \chi(\bar{G}) &\geq \sqrt{4n} \\ \chi(G) + \chi(\bar{G}) &\geq 2\sqrt{n}\end{aligned}$$



### Ejercicio 6.10.8

Sea  $G = (V, E)$  un grafo con  $n = |V|$  vértices, y  $\bar{G}$  su complemento.

Demostrar que  $n \leq \chi(G)\chi(\bar{G}) \leq \left(\frac{n+1}{2}\right)^2$ .

Para la segunda desigualdad, puede serles útil la desigualdad aritmética-geométrica.



*Demostración.* Consideremos el grafo  $H = (V \times V, E')$ , con  $E' = \{\{(u, v), (x, y)\} \mid \{u, x\} \in E \vee \{v, y\} \notin E\}$ . Es decir, un par  $(u, v)$  está conectado a un par  $(x, y)$ , exactamente cuando  $u$  está conectado a  $x$  en  $G$ , o  $v$  está conectado a  $y$  en  $\bar{G}$ .

Sea  $c$  un  $\chi(G)$ -coloreo para  $G$ , y  $c'$  un  $\chi(\bar{G})$ -coloreo para  $\bar{G}$ . Obtenemos entonces un coloreo  $f$  con  $\chi(G)\chi(\bar{G})$  colores para  $H$ , donde  $f((u, v)) = (c(u), c'(v))$ , para cualquier  $(u, v) \in V \times V$ . Veamos que  $f$  es un coloreo válido. Si  $\{(u, x), (v, y)\} \in E'$ , entonces o bien  $\{u, v\} \in E$ , o bien  $\{x, y\} \notin E$ . En el primer caso,  $c(u) \neq c(v)$ , y en el segundo,  $c'(x) \neq c'(y)$ . Luego, el color asignado a  $(u, x)$  no puede ser igual al color asignado a  $(x, y)$ . Luego,  $f$  es un coloreo válido, y  $\chi(G) \leq \chi(G)\chi(\bar{G})$ .

Ahora bien, tomemos el conjunto de vértices de  $H$ ,  $F = \{(v, v) \mid v \in V\}$ . Sea cualquier par de vértices en  $F$ ,  $a = (v, v)$ , y  $b = (w, w)$ . Entonces o bien  $\{v, w\} \in E$ , o bien  $\{v, w\} \notin E$ . En el

primer caso,  $\{(v, v), (w, w)\} \in E'$  porque  $(v, w) \in E$ . En el segundo,  $\{(v, v), (w, w)\} \in E'$  porque  $(v, w) \notin E$ . Luego, para cada par de vértices distintos en  $F$ , tenemos una arista entre ellos en  $H$ . Luego  $F$  es una clique en  $H$ , de  $n$  vértices, y luego todo colooreo de  $H$  debe asignarle  $n$  colores distintos a los vértices de  $F$ . Luego,  $\chi(G) \geq n$ .

Juntando ambas ecuaciones, vemos que  $n \leq \chi(G)\chi(\overline{G})$ .

Para ver que  $\chi(G)\chi(\overline{G}) \leq (\frac{n+1}{2})^2$ , podemos usar la desigualdad aritmética-geométrica. Esta nos dice que para todo  $x, y$  reales no-negativos, tenemos  $\frac{x+y}{2} \geq \sqrt{xy}$ . Aplicando esto a  $x = \chi(G)$ ,  $y = \chi(\overline{G})$ , sabemos que  $\frac{\chi(G)+\chi(\overline{G})}{2} \geq \sqrt{\chi(G)\chi(\overline{G})}$ , o lo que es lo mismo,  $\left(\frac{\chi(G)+\chi(\overline{G})}{2}\right)^2 \geq \chi(G)\chi(\overline{G})$ . Por la desigualdad del Ejercicio 6.10.6, sabemos que  $\chi(G) + \chi(\overline{G}) \leq n + 1$ .

Luego,  $\chi(G)\chi(\overline{G}) \leq (\frac{n+1}{2})^2$ , que es lo que queríamos demostrar.  $\square$

### Ejercicio 6.10.9

Sea  $G$  un grafo, denotemos  $\Delta(G)$  el máximo grado de cualquier vértice en  $G$ .

Probar que  $\chi \leq \Delta + 1$ .



*Demostración.* Vamos a demostrar esto con un algoritmo, el algoritmo greedy para colooreo.

```
def color(V: list[int], E: dict[int, list[int]]) -> list[int]:
    (V, E) = G
    n = len(V)
    Delta = max(len(vs) for vs in E.values())
    all_colors = set(range(Delta + 1))
    colors = [-1 for _ in range(n)]
    for (i, v) in enumerate(V):
        used = set(colors[w] for w in E[v] if colors[w] != -1)
        colors[i] = min(all_colors - used)
    return colors
```

La variable `Delta` representa  $\Delta(G)$ , el máximo grado entre todos los vértices. `all_colors` es el conjunto  $\{0, 1, \dots, \Delta\}$ . El invariante que mantenemos es que al terminar la  $i$ -ésima iteración, le asignamos correctamente un color entre 0 y  $\Delta$  a los primeros  $i$  vértices. Como hay sólo  $\Delta$  vecinos, `len(used)` siempre tiene como mucho  $\Delta$  elementos. Como `len(all_colors) = Delta + 1`, siempre hay algún color no usado al calcular `all_colors - used`, y por lo tanto podemos tomar el mínimo color en esa resta. Como estamos sacando `used` de `all_colors`, nunca vamos a asignarle a `colors[i]` un color idéntico a `colors[j]` con  $j \in E[v]$ . Luego, esta asignación de colores nos da un colooreo válido de los primeros  $i$  vértices, si ya teníamos un colooreo válido de los primeros  $i - 1$  vértices. Al terminar el ciclo, coloreamos los  $n$  vértices.

Este algoritmo, entonces, colorea  $G$  usando a lo sumo  $\Delta + 1$  colores, que están en `all_colors`. Como el algoritmo termina, y asigna un  $(\Delta + 1)$ -colooreo válido, entonces  $\Delta + 1 \geq \chi(G)$ , donde  $\chi(G)$  es el número cromático de  $G$ .  $\square$

### Ejercicio 6.10.10

Sea  $G = (V, E)$  un grafo, y sea  $v \in V$ .

Probar que  $\chi(G - v) \in \{\chi(G), \chi(G) - 1\}$ .



*Demostración.* Sea  $H = G - v$ . Si  $f : G \rightarrow \{1, \dots, \chi(G)\}$  es un colooreo óptimo de  $G$ , entonces  $g : H \rightarrow \{1, \dots, \chi(G)\}$ ,  $g(v) = f(v)$  es un  $\chi(G)$ -colooreo válido de  $H$ . Luego,  $\chi(H) \leq \chi(G)$ .

Sólo nos queda probar que  $\chi(H) \geq \chi(G) - 1$ . Sea  $f : V \setminus \{v\} \rightarrow \{1, \dots, \chi(H)\}$  un colooreo óptimo de  $H$ . Podemos entonces agregar  $v$  a  $H$ , con las mismas aristas que tenía  $v$  en  $G$ .

Vamos a definir:

$$g : V \rightarrow \{1, \dots, \chi(H) + 1\}$$
$$g(w) = \begin{cases} f(w) & \text{si } w \neq v \\ \chi(H) + 1 & \text{si } w = v \end{cases}$$

Como estamos agregando un vértice nuevo,  $v$ , que tiene un color distinto al color de todos los otros vértices (pues  $f(w) \leq \chi(H) \forall w \in V \setminus \{v\}$ ), entonces  $g$  no introduce conflictos de colores. Vemos que  $g$  es, entonces, un  $(\chi(H) + 1)$ -colooreo válido de  $G$ . Por lo tanto,  $\chi(G) \leq \chi(H) + 1$ , o lo que es lo mismo,  $\chi(H) \geq \chi(G) - 1$ , que es lo que queríamos demostrar.  $\square$

## 6.11 Homomorfismo e isomorfismo de grafos

### Definición 6.11.1

Sean  $G = (V, E)$ ,  $H = (V', E')$  grafos. Se dice que una función  $f : V \rightarrow V'$  es un **homomorfismo de grafos** cuando para todo  $u, v \in V$ ,  $\{u, v\} \in E \Rightarrow \{f(u), f(v)\} \in E'$ .

Un **isomorfismo de grafos** entre  $G$  y  $H$  es una función  $f : V \rightarrow V'$ , tal que para todo  $u, v \in V$ ,  $\{u, v\} \in E \Leftrightarrow \{f(u), f(v)\} \in E'$ .



### Ejercicio 6.11.2

Sean  $G, H, K$  grafos, y  $f : G \rightarrow H$ ,  $g : H \rightarrow K$  homomorfismos. Probar que  $g \circ f$  es un homomorfismo.



*Demostración.* Para ver que  $g \circ f$  es un homomorfismo, tenemos que tomar dos vértices  $u, v \in V(G)$ , y ver que si  $\{u, v\} \in E(G)$ , entonces  $\{g(f(u)), g(f(v))\} \in E(K)$ .

Sean  $u, v$  tales que  $\{u, v\} \in E(G)$ . Entonces como  $f$  es un homomorfismo,  $\{f(u), f(v)\} \in E(H)$ . Como  $g$  es un homomorfismo,  $\{g(f(u)), g(f(v))\} \in E(K)$ , que es lo que queríamos demostrar.  $\square$



### Ejercicio 6.11.3

Sean  $G, H$  grafos, y  $f : G \rightarrow H$  un isomorfismo. Probar que si  $G$  es conexo, entonces  $H$  también lo es.



*Demostración.* Sean  $(V_H, E_H) = H$ , y  $(V_G, E_G) = G$ . Vamos a probar que para todo par de vértices en  $H$ , hay un camino entre ellos. Sean  $u, v \in V_H$  cualquier par de vértices. Como  $f$  es un isomorfismo, en particular es biyectiva, y luego existen  $x = f^{-1}(u)$ ,  $y = f^{-1}(v)$ . Como  $G$  es conexo, existe un camino  $P = [x = p_1, p_2, \dots, p_k = y]$  en  $G$ . Para todo  $1 \leq i < k$ ,  $\{p_i, p_{i+1}\} \in E_G$ . Luego, como  $f$  es un isomorfismo, para todo  $1 \leq i < k$ ,  $e = \{f(p_i), f(p_{i+1})\} \in E_H$ . Consideremos  $P' = [f(x) = f(p_1), f(p_2), \dots, f(p_k) = f(y)]$ . Como vimos, cada transición es una arista en  $E_H$ . Luego,  $P'$  es un camino en  $H$ , entre  $f(x) = u$ , y  $f(y) = v$ .  $\square$

#### Ejercicio 6.11.4

Sean  $G = (V, E)$  y  $H = (V', E')$  grafos.

Notemos por  $\chi(G)$  el número de colojo de un grafo  $G$ , y  $\omega(G)$  el tamaño de la clique máxima en  $G$ .

Mostrar que si hay un homomorfismo  $f : G \rightarrow H$ , entonces:

1.  $\omega(G) \leq \omega(H)$
2.  $\chi(G) \leq \chi(H)$



*Demostración.* Notemos  $G = (V, E)$ ,  $H = (W, F)$ .

1. Sea  $K \subseteq V$  una clique máxima en  $G$ . Luego  $|K| = \omega(G)$ . Sea  $K' = \{f(w) \mid w \in K\} \subseteq W$ . Sean  $u, v \in K$ . Entonces  $\{u, v\} \in E$ , y por lo tanto,  $\{f(u), f(v)\} \in F$ .  $H$  es un grafo, y no pseudografo,  $f(u) \neq f(v)$ . Luego, todos los vértices en  $K$  van a parar a vértices distintos en  $K'$ , y luego  $f|_K : K \rightarrow K'$  es biyectiva. Entre cada par de vértices  $x, y \in K'$ , vamos a encontrar entonces una arista  $\{f^{-1}(x), f^{-1}(y)\}$  en  $G$ , pues  $f^{-1}(x)$  y  $f^{-1}(y)$  son vértices distintos en  $K$ , una clique. Luego  $K'$  también es una clique. Como  $f|_K$  es una biyección, tenemos que  $|K'| = |K| = \omega(G)$ . Luego,  $H$  contiene una clique de tamaño  $\omega(G)$ , y entonces la clique máxima de  $H$  tiene tamaño como mínimo  $\omega(G)$ . Luego,  $\omega(G) \leq \omega(H)$ .
2. Sea  $g : W \rightarrow \{1, \dots, \chi(H)\}$  un colojo óptimo de  $H$ . Vamos a ver que  $g \circ f$  es un  $\chi(H)$ -colojo de  $G$ . Claramente le asigna a cada vértice de  $H$  un número entre 1 y  $\chi(H)$ , por cómo definimos  $g$ . Ahora tomemos cualquier par de vértices  $u, v \in V(G)$ , tal que  $\{u, v\} \in E(G)$ . Luego, como  $f$  es un homomorfismo,  $\{f(u), f(v)\} \in E(H)$ . Como  $g$  es un colojo, entonces  $g(f(u)) \neq g(f(v))$ . Luego, el colojo que creamos,  $g \circ f$ , es un colojo válido para  $G$ , con  $\chi(H)$  colores. Entonces  $\chi(G) \leq \chi(H)$ .



## 6.12 Circuitos y caminos

### Definición 6.12.1

Sea  $G$  un grafo. Un camino en  $G$  se dice **camino euleriano** cuando pasa por todas las aristas de  $G$  exactamente una vez.

Un ciclo de  $G$  se dice **ciclo euleriano** cuando pasa por todas las aristas de  $G$  exactamente una vez.

Si  $G$  tiene un ciclo euleriano, se dice que  **$G$  es euleriano**.



### Definición 6.12.2

Sea  $G$  un grafo. Un camino en  $G$  se dice **camino Hamiltoniano** cuando pasa por todos los vértices exactamente una vez.

Un ciclo de  $G$  se dice **ciclo Hamiltiniano** cuando pasa por todos los vértices exactamente una vez.

Si  $G$  tiene un ciclo Hamiltoniano, entonces se dice que  **$G$  es Hamiltoniano**



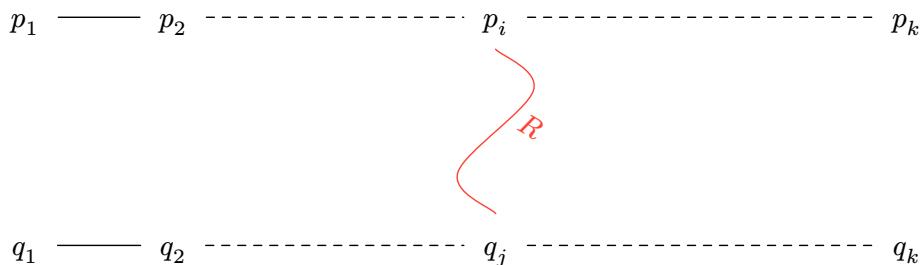
### Ejercicio 6.12.3

Sea  $G$  un grafo conexo, y sean  $P$  y  $Q$  dos caminos de longitud máxima en  $G$ . Probar que  $P$  y  $Q$  comparten al menos un vértice.



*Demostración.* Por contradicción, asumamos que no. Luego,  $P = [p_1, \dots, p_k]$  y  $Q = [q_1, \dots, q_k]$  no comparten vértices. Notar que como ambos tienen longitud máxima, ambos tienen que tener la misma longitud, que llamamos  $k$ .

Como  $G$  es conexo, existen vértices  $p_i$  y  $q_j$ , tal que hay un camino  $R$  entre  $p_i$  y  $q_j$ . Sin pérdida de generalidad, podemos asumir que  $R$  no tiene vértices en  $P$  ni en  $Q$ , salvo  $p_i$  y  $q_j$  (siempre podemos quedarnos con el sub-camino de  $R$  que cruce los caminos, sin tocar a nadie de  $P$  ni  $Q$  en el medio).



Sin pérdida de generalidad, podemos asumir que  $i \geq \lceil \frac{k}{2} \rceil$ . Si esto no sucediera, simplemente revertimos el orden de los vértices de  $P$ . Por el mismo motivo, podemos asumir que  $j \geq \lceil \frac{k}{2} \rceil$ . Vemos que  $|R| \geq 1$ , pues si no,  $P$  y  $Q$  compartirían vértices.

Consideremos ahora el camino  $\alpha = [p_1, p_2, \dots, p_i] + R + [q_j, \dots, q_1]$ . Este camino tiene longitud  $|\alpha| = i + |R| + j \geq 2\lceil \frac{k}{2} \rceil + |R| \geq k + |R| > k$ . Esto no puede suceder, pues  $P$  y  $Q$

eran caminos de longitud máxima,  $k$ , y acabamos de encontrar un camino de longitud mayor a  $k$ .

Luego no puede pasar que  $P$  y  $Q$  sean disjuntos, y tienen que compartir algún vértice.  $\square$

#### Ejercicio 6.12.4

Probar que si un grafo tiene algún vértice de grado impar, entonces no tiene un circuito euleriano.

*Demostración.* Un ciclo euleriano usa cada arista exactamente una vez, y termina donde comienza. Cada vez que tomamos una arista  $u \rightarrow v$  en el camino, visitando a  $v$ , inmediatamente tomamos  $v \rightarrow w$ , saliendo de  $v$ . Luego, cada vez que usamos una arista hacia  $v$ , usamos otra que sale desde  $v$ . Como las aristas incidentes a  $v$  se usan siempre de a pares, y al terminar el ciclo todas las aristas son usadas exactamente una vez, si llamamos  $f(v)$  al número de veces que  $C$  visita  $v$ , tenemos  $d_G(v) = 2f(v)$ . Luego,  $d_G(v)$  es par, para todo  $v \in V$ .  $\square$

#### Ejercicio 6.12.5

Probar que si un grafo tiene un camino euleriano, entonces tiene exactamente dos vértices con grado impar.

*Demostración.* Sea  $G = (V, E)$  un grafo, tal que más de dos vértices en  $V$  tienen grado impar.

Supongamos que  $G$  tiene un camino euleriano,  $P$ , desde  $u$  a  $w$ . Sea  $v \in V$ ,  $v \neq u$ ,  $v \neq w$ . Si  $v$  aparece  $k$  veces en  $P$ , entonces como  $P$  usa todas las aristas de  $G$  exactamente una vez, y cada vez que  $P$  pasa por  $v$  toca dos de sus aristas incidentes,  $v$  debe tener grado  $2k$ . En el caso de  $u$  y  $w$ , la primera vez que  $P$  visita a  $u$ , y la última vez que visita a  $w$ ,  $P$  sólo usa una de sus aristas incidentes.

Luego los únicos dos vértices que tienen grado impar son  $u$  y  $w$ , y todos los otros vértices tienen grado par. Notar que como  $P$  es un camino euleriano, y no un ciclo, sabemos que  $u$  y  $w$  son distintos.  $\square$

#### Ejercicio 6.12.6

Sea  $G = (V, E)$  un grafo dirigido, tal que para todo par de vértices  $u, v \in V$ , exactamente uno de  $(u, v)$  o  $(v, u)$  está en  $E$ .

Probar que  $G$  contiene un camino Hamiltoniano.

*Demostración.* Vamos a probar esto por inducción. Por comodidad, defino  $Q(G = (V, E)) :$   $\forall u, v \in V. ((u, v) \in E) \neq ((v, u) \in E)$ . Definimos entonces  $P(n) :$  Para todo grafo dirigido  $G$

con  $n$  vértices tal que  $Q(G)$ ,  $G$  tiene un camino Hamiltoniano. Por comodidad, dendo un subconjunto  $X \subseteq V$  de vértices de un grafo  $G = (V, E)$ , definimos  $G[X] = (X, \{(a, b) \in E \mid a \in X, b \in X\})$ , el subgrafo inducido por  $X$  en  $G$ .

1. Caso base,  $P(0)$ . Esto es trivialmente cierto porque no hay ningún grafo sin vértices, luego no existe ningún tal  $G$ .
2. Caso base,  $P(1)$ . También trivial, si  $V = \{v\}$ ,  $[v]$  un camino Hamiltoniano.
3. Paso inductivo. Sea  $n \in \mathbb{N}, n \geq 2$ . Asumimos que vale  $P(k)$  para todo  $k \in \mathbb{N}, k < n$ , queremos probar  $P(n)$ . Sea  $G = (V, E)$  un grafo dirigido, con  $|V| = n$ . Como  $n \geq 2$ , sea cuquier  $v \in V$ , y  $W = V \setminus \{v\}$ . Sea  $W^\rightarrow = \{w \in W \mid (w, v) \in E\}$ , y  $W^\leftarrow = \{w \in W \mid (v, w) \in E\}$ . Como para todo  $w \in W \subseteq V$  sabemos que o bien  $(w, v) \in E$ , o bien  $(v, w) \in E$ , y ocurre exactamente una de las dos, entonces vemos que  $W = W^\rightarrow \sqcup W^\leftarrow$ , la unión disjunta.

Como  $W = V \setminus \{v\}$ , tenemos  $|W| = |V| - 1 = n - 1 < n$ . Como ambos  $W^\rightarrow \subseteq W$  y  $W^\leftarrow \subseteq W$ , tenemos  $|W^\rightarrow| < n$ , y  $|W^\leftarrow| < n$ , y es más,  $|W^\rightarrow| + |W^\leftarrow| = |W| = n - 1$ .

- Si  $W^\leftarrow = \emptyset$ , entonces  $W = W^\rightarrow$ , y luego  $|W^\rightarrow| = |W| = n - 1 \geq 1$ . Construimos  $G' = G[W^\rightarrow]$ , con  $n - 1$  vértices. Como  $Q(G)$ , sigue valiendo  $Q(G')$ . Ahora usamos  $P(n - 1)$  para obtener un camino Hamiltoniano  $R = [r_1, \dots, r_{n-1}]$  en  $G'$ . Como es Hamiltoniano, no repite vértices.  $R$  existe en  $G$ , por ser  $G'$  subgrafo de  $G$ . Como  $r_{n-1} \in W^\rightarrow$ , por definición  $(r_{n-1}, v) \in E$ . Sea  $R' = [r_1, \dots, r_{n-1}, v]$ . Como  $R$  es un camino en  $G'$ , y  $v \notin V(G')$ , entonces  $v \notin R$ . Luego  $R'$  es un camino de longitud  $n$  en  $G$  que no repite vértices, y luego es Hamiltoniano en  $G$ .
- Si  $W^\rightarrow = \emptyset$ , pasa algo análogo, tomando el subgrafo inducido por  $W^\leftarrow$  en  $G$ .
- Si ninguna de las dos particiones está vacía, sea  $k = |W^\rightarrow|, t = |W^\leftarrow|$ , con  $1 \leq k, t < n$ . Tomemos  $G^\rightarrow = G[W^\rightarrow], G^\leftarrow = G[W^\leftarrow]$ , y como seguimos teniendo  $Q(G^\rightarrow)$  y  $Q(G^\leftarrow)$  por ser subgrafos inducidos de  $G$ , usamos  $P(k)$  y  $P(t)$ , y vemos que ambos tienen un camino Hamiltoniano. Sean  $R^\rightarrow = [x_1, \dots, x_k]$  y  $R^\leftarrow = [y_1, \dots, y_t]$  tales caminos en  $G^\rightarrow$  y  $G^\leftarrow$ , respectivamente. Por ser Hamiltonianos, sabemos que no repiten vértices, y sabiendo que  $|W^\rightarrow| + |W^\leftarrow| = n - 1$ , sabemos que  $k + t = n + 1$ . Como  $W^\rightarrow \cap W^\leftarrow = \emptyset$ , vemos que  $R^\rightarrow$  y  $R^\leftarrow$  no comparten vértices. Como  $x_k \in W^\rightarrow$ , tenemos que  $(x_k, v) \in E$ . Como  $y_1 \in W^\leftarrow$ , tenemos que  $(v, y_1) \in E$ . Luego,  $R = R^\rightarrow + [v] + R^\leftarrow = \left[ \underbrace{x_1, \dots, x_k}_\in W^\rightarrow, v, \underbrace{y_1, \dots, y_t}_\in W^\leftarrow \right]$  es un camino en  $G$  de longitud  $k + t + 1 = n - 1 + 1 = n$ . Como  $v \notin W^\rightarrow$  y  $v \notin W^\leftarrow$ ,  $R$  no repite vértices, y luego es Hamiltoniano en  $G$ .

□

### Ejercicio 6.12.7 (Teorema de Bondy-Chvátal[5])

Sea  $G$  un grafo de  $n$  vértices, y sean  $u$  y  $v$  vértices no adyacentes, tal que  $d_G(u) + d_G(v) \geq n$ .

Probar que si  $G + \{u, v\}$  es Hamiltoniano, entonces  $G$  también lo es.

♣

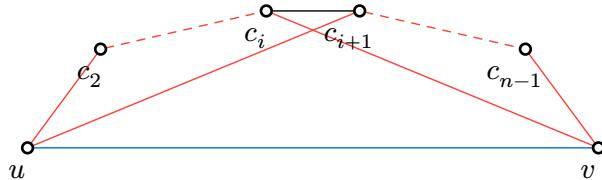
*Demostración.* Llamemos  $G = (E, V)$ , y  $n = |V|$ . Sea  $C$  un ciclo Hamiltoniano de  $G + \{u, v\}$ . Si  $\{u, v\} \notin C$ , entonces  $C$  también es un ciclo Hamiltoniano en  $G$ , y terminamos. Luego,

asumamos que  $\{u, v\} \in C$ . Consideremos el camino  $C - \{u, v\} = [c_1 = u, c_2, \dots, c_n = v]$ .

Sean  $S_u = \{i \mid 1 \leq i \leq n-1, \{u, c_{i+1}\} \in E\}$ , y  $S_v = \{i \mid 1 \leq i \leq n-1, \{c_i, v\} \in E\}$ .

Vemos que  $|S_u| = d_G(u)$ , y  $|S_v| = d_G(v)$ . Luego,  $|S_u| + |S_v| \geq n$ , por hipótesis, mientras que luego  $|S_u \cup S_v| \leq n-1$ , dado que en sus definiciones,  $i \in [0, \dots, n-1]$ . Luego,  $S_u \cap S_v \neq \emptyset$ .

Sea entonces  $i \in S_u \cap S_v$ .



Como  $c_1 = u$  y  $c_n = v$  no son adjacientes en  $G$ , vemos que  $2 \leq i \leq n-2$ . Pero ahora,  $C' = [c_1 = u, c_2, \dots, c_i, v = c_n, c_{n-1}, \dots, c_{i+1}, c_1 = u]$  es un ciclo Hamiltoniano en  $G$ , y luego  $G$  es Hamiltoniano.  $\square$

### Ejercicio 6.12.8 (Teorema de Dirac)

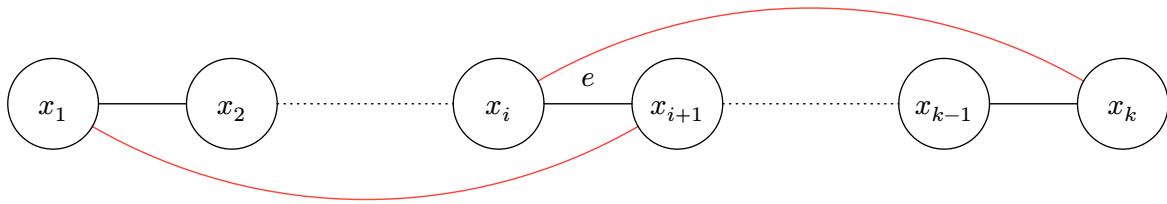
Si  $G$  tiene  $n \geq 3$  vértices y mínimo grado  $\delta(G) \geq \frac{n}{2}$ , entonces  $G$  es Hamiltoniano.



*Demostración.* Sea  $C$  la componente conexa más chica de  $G = (V, E)$ , y  $v \in C$ . Como  $\delta(G) \geq \frac{n}{2}$ , tenemos  $d(v) \geq \frac{n}{2}$ . Como al menos  $1 + \frac{n}{2}$  vértices (i.e. más de la mitad de los vértices) están en la componente conexa más chica, tiene que haber sólo una componente conexa. Luego  $G$  es conexo.

Sea  $P = [x_1, \dots, x_k]$  un camino de máxima longitud en  $G$ . Como  $\delta(G) \geq \frac{n}{2}$ , entonces  $d(x_k) \geq \frac{n}{2}$ , y  $d(x_1) \geq \frac{n}{2}$ . Si algún vecino de  $x_1$  o  $x_k$  no estuviera en  $P$ , podríamos extender alguno de los extremos de  $P$ , obteniendo un camino de mayor longitud de  $P$ . Por ende todos los vecinos de  $x_1$ , y todos los vecinos de  $x_k$  están en  $P$ .

Probemos que existen dos vértices,  $x_j$  y  $x_{j+1} \in V$ , tal que  $\{x_j, x_k\} \in E$ , y  $\{x_{j+1}, x_1\} \in E$ . Sea  $A = \{\{x_t, x_{t+1}\} \mid \{x_{t+1}, x_1\} \in E\}$ , y  $B = \{\{x_t, x_{t+1}\} \mid \{x_t, x_k\} \in E\}$ . Vemos que  $|A| = d(x_1) \geq \frac{n}{2}$ , y  $|B| = d(x_k) \geq \frac{n}{2}$ . Como  $P$  tiene a lo sumo  $n-1$  aristas, y  $|A| + |B| \geq n$ , por el principio del palomar existe alguna arista que está en ambos  $|A|$  y  $|B|$ . Esa arista, luego, es de la forma  $e = \{a, b\}$  con  $\{a, x_k\} \in E$  (pues  $e \in A$ ) y  $\{x_1, b\} \in E$  (pues  $e \in B$ ).



Sea  $C = [x_1, x_2, \dots, x_i, x_k, x_{k-1}, \dots, x_{i+1}, x_1]$ , de longitud  $k+1$ .  $C$  es un ciclo.

Si no fuera Hamiltoniano, entonces como  $G$  es conexo, existiría un vértice  $y \in V$ , adyacente a algún  $x_j \in C$ , con  $y \notin C$ . Luego existe  $\{x_j, y\} \in E$ . Pero entonces podríamos tomar  $C$ , sacarle una arista incidente a  $x_j$ , y agregarle  $e$ , y obtendríamos un *camino* en  $G$ , de longitud  $k+1$ , que no puede pasar pues  $P$ , de longitud  $k$ , era de longitud máxima.

Luego  $C$  es un ciclo Hamiltoniano en  $G$ , y luego  $G$  es Hamiltoniano. □

## 6.13 Flujo

### Definición 6.13.1

Una **red de flujo**  $G$  es una 5-tupla  $G = (V, E, s, t, c)$ , tal que  $(V, E)$  es un grafo,  $s \in V, t \in V$ , y  $c : E \rightarrow \mathbb{R}_{\geq 0}$ .

Un **flujo** en tal red es una función  $f : E \rightarrow \mathbb{R}_{\geq 0}$ , tal que:

- $0 \leq f(e) \leq c(e)$  para todo  $e \in E$ .
- Para todo  $v \in V \setminus \{s, t\}$ ,  $\sum_{u \in V} f(u, v) = \sum_{u \in V} f(v, u)$ .

En tal caso decimos que el **valor de  $f$**  es  $|f| = \sum_{v \in V} f(s, v) - \sum_{v \in V} f(v, s)$ . Un flujo se dice **máximo** cuando tiene valor máximo entre todos los flujos en  $G$ .

Dada una tal red de flujo, un **corte** es una partición de  $V$  en  $(S, T)$ , tal que  $s \in S$ , y  $t \in T$ .

Dado un flujo  $f$  en esa red, el **flujo neto de  $(S, T)$**  se define como  $f(S, T) =$

$\sum_{u \in S, v \in T} f(u, v) - \sum_{u \in S, v \in T} f(v, u)$ . La **capacidad de  $(S, T)$**  se define como  $c(S, T) = \sum_{u \in S, v \in T} c(u, v)$ .



### Ejercicio 6.13.2

Sea  $G = (V, E, s, t, c)$  una red de flujo, y  $f$  un flujo máximo en  $G$ . Sea  $\lambda \in \mathbb{R}_{>0}$ . Definimos  $c'$  como  $c'(x) = \lambda c(x)$  para todo  $x \in E$ .

Sea  $f'$  un flujo máximo en  $G' = (V, E, s, t, c')$ .

Demostrar que  $|f'| = \lambda |f|$ .



*Demostración.* Vamos a construir una biyección explícita  $\varphi$  entre flujos en  $G$ , y flujos en  $G'$ .

Sea  $f : E \rightarrow \mathbb{R}_{\geq 0}$  un flujo en  $G$ . Definimos  $\varphi(f) : E \rightarrow \mathbb{R}_{\geq 0}$  como  $\varphi(f)(e) = \lambda f(e)$  para todo  $e \in E$ . Notemos que  $\varphi(f)$  es un flujo en  $G'$ , pues como  $f$  es un flujo en  $G$ , sabemos que  $0 \leq f(e) \leq c(e)$  para todo  $e \in E$ , y luego  $0 \leq \varphi(f)(e) = \lambda f(e) \leq \lambda c(e) = c'(e)$  para todo  $e \in E$ . Además, como  $f$  es un flujo, tenemos que para todo  $v \in V \setminus \{s, t\}$ , se cumple que

$$\begin{aligned} \sum_{u \in V} f(u, v) &= \sum_{u \in V} f(v, u), \text{ y luego } \sum_{u \in V} \varphi(f)(u, v) = \lambda \sum_{u \in V} f(u, v) = \\ \lambda \sum_{u \in V} f(v, u) &= \sum_{u \in V} \varphi(f)(v, u), \text{ y luego } \varphi(f) \text{ es un flujo en } G'. \end{aligned}$$

Notamos que  $\varphi$  es monótona en valores. Es decir, si  $|f| \geq |g|$ , entonces  $|\varphi(f)| = \lambda |f| \geq \lambda |g| = |\varphi(g)|$ . Luego, si  $f$  es un flujo máximo en  $G$ , entonces  $\varphi(f)$  es un flujo máximo en  $G'$ . Entonces si  $f'$  es el valor de cualquier flujo máximo en  $G'$ , tenemos que  $|f'| = |\varphi(f)| = \lambda |f|$ , y luego  $|f'| = \lambda |f|$ , que es lo que queríamos demostrar. □

### Ejercicio 6.13.3

Sea  $G = (V, E, s \in V, t \in V, c : E \rightarrow \mathbb{R}_{\geq 0})$  una red de flujo. Sea  $f : E \rightarrow \mathbb{R}_{\geq 0}$  un flujo en  $G$ . Sea  $(S, T)$  un corte en  $G$ .

Mostrar que  $|f| = f(S, T) \leq c(S, T)$ .



*Demostración.* Como  $f$  es un flujo, sabemos que para todo  $u \in V \setminus \{s, t\}$ , tenemos

$\sum_{v \in V} f(u, v) = \sum_{v \in V} f(v, u)$ . Esto es lo mismo que decir  $0 = \sum_{v \in V} f(u, v) - \sum_{v \in V} f(v, u)$ . Por definición,  $|f| = \sum_{v \in V} f(s, v) - \sum_{v \in V} f(v, s)$ . Si a esta última ecuación le sumamos la primera, por todo  $u \in S \setminus \{s\}$ , tenemos que

$$\begin{aligned} |f| &= \sum_{v \in V} f(s, v) - \sum_{v \in V} f(v, s) + \sum_{u \in S \setminus \{s\}} \left( \sum_{v \in T} f(u, v) - \sum_{v \in T} f(v, u) \right) \\ &= \sum_{v \in V} f(s, v) - \sum_{v \in V} f(v, s) + \sum_{u \in S \setminus \{s\}} \sum_{v \in T} f(u, v) - \sum_{u \in S \setminus \{s\}} \sum_{v \in T} f(v, u) \\ &= \sum_{v \in V} \left( f(s, v) \sum_{u \in S \setminus \{s\}} f(u, v) \right) - \sum_{v \in V} \left( f(v, s) + \sum_{u \in S \setminus \{s\}} f(v, u) \right) \\ &= \sum_{v \in V} \sum_{u \in S} f(u, v) - \sum_{v \in V} \sum_{u \in S} f(v, u) \end{aligned}$$

Ahora usamos que  $V = S \sqcup T$ ,

$$\begin{aligned} &= \left( \sum_{v \in S} \sum_{u \in S} f(u, v) + \sum_{v \in T} \sum_{u \in S} f(u, v) \right) - \left( \sum_{v \in S} \sum_{u \in S} f(v, u) - \sum_{v \in T} \sum_{u \in S} f(v, u) \right) \\ &= \sum_{v \in T} \sum_{u \in S} f(u, v) - \sum_{v \in T} \sum_{u \in S} f(v, u) \\ &= f(S, T) \end{aligned}$$

Para ver una cota superior de  $f(S, T)$ , basta usar sólo la primer sumatoria:

$$\begin{aligned} |f| &= f(S, T) \\ &= \sum_{v \in T} \sum_{u \in S} f(u, v) - \sum_{v \in T} \sum_{u \in S} f(v, u), \text{ al ser } f(v, u) \geq 0 \quad \forall v, u \in V \\ &\leq \sum_{v \in T} \sum_{u \in S} f(u, v) \\ &\leq \sum_{v \in T} \sum_{u \in S} c(u, v) \\ &= \sum_{u \in S, v \in T} c(u, v) \\ &= c(S, T) \end{aligned}$$



### Ejercicio 6.13.4

Sea  $G = (V, E, s, t, c)$  una red de flujo, y  $f$  un flujo máximo en  $G$ . Sea  $n = |V|, m = |E|$ . Ahora elegimos un  $e \in E$ , y construimos  $G' = (V, E, s, t, c')$ , tal que:

$$c'(x) = \begin{cases} c(x) & \text{si } x \neq e \\ c(x) + 1 & \text{si } x = e \end{cases}$$

Dar un algoritmo que encuentre un flujo máximo en  $G'$ . El algoritmo debe hacer  $O(n + m)$  operaciones en el peor caso. Demostrar que el algoritmo es correcto.



*Demostración.* Primero vamos a citar dos teoremas que ven en la teórica.

#### Teorema 10

Sea  $G$  una red de flujo, y  $f$  un flujo en  $G$ . Sea  $G_f$  la red residual.  $f$  es un flujo máximo en  $G$  si y sólo si  $G_f$  no tiene caminos de aumento.



Por el teorema de flujo máximo - corte mínimo, como  $f$  es un flujo máximo, existe un corte  $(S, T)$  en  $G$  tal que  $|f| = c(S, T)$ . Veamos ahora cuál es la capacidad del corte  $(S, T)$  en  $G'$ . Si  $e$  no cruza la partición  $(S, T)$ , entonces  $c'(S, T) = c(S, T)$ . Si  $e$  cruza la partición, entonces  $c'(S, T) = c(S, T) + 1$ . En ambos casos, tenemos que  $c(S, T) \leq c'(S, T) \leq c(S, T) + 1 = |f| + 1$ .

Como  $f$  es un flujo en  $G$ , también es un flujo en  $G'$ , puesto que las condiciones de sumatoria en cada vértice se siguen cumpliendo, y las condiciones de capacidad también (pues meramente *agregamos* capacidad a  $e$ ). Podemos entonces construir la red residual de  $f$ ,  $G'_f$ .

Esto nos sugiere correr una iteración más del algoritmo de Edmonds-Karp en  $G'_f$ , y si encontramos un camino de aumento  $P$ , podemos aumentar el flujo en  $f$  en  $c(P)$ , obteniendo un flujo  $f'$ , y luego  $|f'| = |f| + c(P) = |f| + \min_{e \in P} c(e)$ . Como vimos,  $|f'| \leq |f| + 1$ , entonces  $c(P)$  va a ser 1, si  $P$  existe, pues todos los flujos encontrados por Edmonds-Karp son de enteros, siendo las capacidades de  $G'$  enteros. Si no encontramos un camino, por el Teorema 10  $f$  es un flujo máximo en  $G'$ , y luego  $|f'| = |f|$ . □

## 7 Ejercicios

### Ejercicio 7.1

Sean  $x, y \in \mathbb{R}_{\geq 0}$ . Probar que  $\frac{x+y}{2} \geq \sqrt{xy}$ , y que la igualdad vale si y sólo si  $x = y$ .



### Ejercicio 7.2

Sea  $T : \mathbb{N} \rightarrow \mathbb{N}$  una función que cumple que para todo  $n > 0$ ,  $T(n) = 4T\left(\frac{n}{3}\right) + O(n \log n)$ . Probar que  $T \in O(n^{1.5})$ .



### Ejercicio 7.3

Se tienen dos arreglos de  $n$  naturales,  $A$  y  $B$ .  $A$  está ordenado de manera creciente, y  $B$  de manera decreciente. Ningún valor aparece más de una vez en el mismo array. Para cada posición  $i$ , consideraremos la diferencia absoluta entre los valores de los arrays,  $|A[i] - B[i]|$ . Se desea buscar el mínimo valor posible de dicha cuenta. Por ejemplo, si los arrays son  $A = [1, 2, 3, 4]$ , y  $B = [6, 4, 2, 1]$ , los valores de las diferencias son  $[5, 2, 1, 3]$ , y el resultado es 1.

1. Diseñar un algoritmo basado en divide-and-conquer que resuelva este problema.
2. Demostrar que es correcto.
3. Dar una cota superior ajustada de su complejidad temporal asintótica.

### Ejercicio 7.4

Demostrar la correctitud del algoritmo de Dijkstra para caminos mínimos.

### Ejercicio 7.5

Demostrar la correctitud del algoritmo de Kruskal para árboles generadores mínimos.

### Ejercicio 7.6

Demostrar la correctitud del algoritmo de Prim para árboles generadores mínimos.

### Ejercicio 7.7

Se tiene un algoritmo  $A$  que multiplica dos matrices de  $n \times n$  en tiempo  $O(n^\omega)$ . Sea  $G = (V, E)$  un grafo dirigido pesado, con  $|V| = n$ .

Dar un algoritmo que determine el camino de peso mínimo entre todo par de vértices en  $G$ , en tiempo  $O(n^\omega \log n)$ . Demostrar que es correcto.

## Bibliografía

- [1] Euclides, *Elementos*. 301AD.
- [2] Gregory H. Moore, *Zermelo's Axiom of Choice: Its Origins, Development, and Influence*, 1.<sup>a</sup> ed. Springer, 1982.
- [3] William Kuszmaul y Charles Eric Leiserson, «Floors and Ceilings in Divide-and-Conquer Recurrences», *Symposium on Simplicity in Algorithms*, pp. 133-141, 2021, doi: 10.1137/1.9781611976496.15.
- [4] George Bernard Dantzig, «All shortest routes in a graph», Operations Research House, Stanford University, nov. 1966.
- [5] Václav Chvátal, «On Hamilton's ideals», *Journal of Combinatorial Theory, Series B*, vol. 12, pp. 163-168, 1972.
- [6] Thomas H. Cormen, Charles Eric Leiserson, Ronald Linn Rivest, y Clifford Stein, *Introduction to Algorithms*, 3.<sup>a</sup> ed. The MIT Press, 2009.