



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Tesis de Licenciatura

Propiedades lógicas de estructuras aleatorias

Numa Grinberg

Director: Sergio Abriola

Fecha de Presentación: 16 de Abril de 2026

Índice general

Agradecimientos	4
Introducción	5
1. Preliminares lógicos	7
1.1. El misterio de las estructuras finitas	7
1.2. Los juegos de Ehrenfeucht-Fraisse	8
1.3. Los teoremas de Localidad	16
1.3.1. El teorema de Hanf	16
1.3.2. Aplicación: Decibilidad de la lógica de primer orden monádica	19
1.3.3. El teorema de Gaifamn	22
1.4. Variantes de los juegos E-F	26
1.4.1. Juegos E-F en lenguaje con funciones	27
1.4.2. Juego E-F Unilateral	27
1.4.3. Juego E-F Biyectivo	28
1.4.4. Juegos E-F de segundo orden	30
1.4.5. Juegos con fichas	33
2. Ley 0-1 de Fagin para grafos aleatorios y otras estructuras	40
2.1. Grafos Aleatorios	40
2.1.1. La ley de Fagin	40
2.1.2. Cotas para la convergencia de la ley de Fagin	45
2.1.3. Ley de Fagin no etiquetada	51
2.2. Una generalización: Las estructuras paramétricas	56
2.2.1. Introducción	56
2.2.2. La ley de Oberschelp	59
2.2.3. Cotas para la convergencia de la ley de Oberschelp	61
2.3. Estructuras universales	64
2.3.1. El grafo de Rado	64
2.3.2. El caso general	71
3. Ley de convergencia para grafos muy sparse	75
3.1. Introducción	75
3.2. Preliminares probabilísticos	76
3.2.1. El método de los momentos, ó, Cómo aprendí a dejar de preocuparme y a amar la esperanza	76
3.2.2. Cuatro versiones de la ley de convergencia Poisson	81
3.3. Grafos muy Sparse	87
3.3.1. Estructura local	87
3.3.2. Ley de convergencia	96
3.4. Espacios con funciones	101

4. No convergencia y no decidibilidad en grafos ordenados y otras estructuras	104
4.1. Grafos ordenados	104
4.1.1. Introducción	104
4.1.2. La maratón de danza	105
4.1.3. El método de los cerebros de Boltzmann	108
4.1.4. Ley de poca variación	115
4.2. Otras estructuras ordenadas	118
4.2.1. Conjuntos con múltiples ordenes	118
4.2.2. Conjuntos ordenados con permutaciones	123
4.2.3. Conjuntos ordenados con funciones	125
4.2.4. Grafos sparse ordenados	128
4.3. Lógica de segundo orden monádica sobre grafos	130
Conclusiones	136
A. Algunas definiciones elementales	137

Agradecimientos

A Sergio, por haber aceptado dirigir este proyecto ecléctico y ayudar a convertirlo en algo medianamente legible. A Inés y Santiago, por aceptar ser jurados del trabajo y por sus comentarios.

A la Universidad de Buenos Aires, por la excelente formación recibida y por darme trabajo tanto tiempo.

A todos los amigos que conocí durante la carrera, por haber hecho tan felices estos años. A Casiel, Frodo, Viole, Joa, Lauti, Cami, Juan, Juan, Juli, Juli, Nahue, Nahue ¹, Zenón, Gonza, Iván. A Boris, por todas las salidas al cine. A Valen, por haberme enseñado tantas cosas (incluyendo, crucialmente, como usar LaTeX). A Nacho, por todas las charlas sobre la matemática o la vida que se extendieron hasta la madrugada.

A Pepa, luz de mi vida, por todos los momentos que pasamos juntos y todos los que nos quedan pasar.

¹Dejo abierto el interrogante de si los pares de nombres repetidos hacen referencia a dos personas distintas o meramente a dos facetas distintas de una única persona, que todavía no conocemos.

Introducción

Desde sus modestos orígenes como un pequeño arbolito en las costas del Eufra-tes o del Mediterráneo, la matemática creció hasta convertirse en un secuoya gigantesca, con ramas de kilómetros de longitud que se extienden en el cielo en toda dirección posible. La vista es majestuosa, cuando no abrumadora, pero una consecuencia no tan agraciada de ese crecimiento es que las ramas están cada vez mas lejos entre si. Cada vez los descubrimientos del algebristas son mas ininteligibles para el analista y viceversa. Se ha discutido mucho quien fue el último matemático universal, pero independientemente de si se piense que fue Euler, Gauss, Poincare o Hilbert, todos están de acuerdo que fue generacio-nes atrás. El árbol creció demasiado para que la vista de un solo humano pueda abarcarlo por completo.

Por esa razón, son tan felices esos extraños momentos en los que, desafiando toda concepción previa sobre la geometría del árbol, se descubren nuevas cone-xiones entre ramas que se crean lejanas. Esta tesis esta dedicada a una pequeña y muy bella intersección entre la lógica, la teoría de probabilidad y la combina-toria iniciada en los 70s por las investigaciones de Ronald Fagin y, del otro lado del telón de acero, por Yuri Glebskii, aunque este último hecho fue por muchos años desconocido en occidente.

Un tópico central en la combinatoria y teoría de probabilidad es el estudio de propiedades asintóticas de estructuras aleatorias. El ejemplo mas celebre y fruc-tífero de dicha naturaleza es la teoría de grafos aleatorios iniciada por Paul Erdős y Alfred Renyi en la decada del 50. La conexión con la lógica empieza en [12], donde Fagin considera, dado un lenguaje relacional \mathcal{L} fijo, la variable aleatoria $S_{\mathcal{L}}(n)$, que tiene una distribución uniforme sobre el conjunto de \mathcal{L} -estructuras sobre $\{1, \dots, n\}$ y demuestra que vale el siguiente teorema:

Teorema 0.1. (Ley 0-1 de Fagin)[12] Sea \mathcal{L} un lenguaje y ϕ una oración sobre \mathcal{L} . Luego:

$$\lim_{n \rightarrow \infty} Pr(S_{\mathcal{L}} \models \phi) \in \{0, 1\}$$

Notemos que el resultado anterior se puede interpretar de dos maneras distin-tas. Desde un punto de vista probabilístico, es un meta-resultado, un teorema que responde a una clase muy amplia de problemas de probabilidad asintóti-cos. Pero desde un punto de vista lógico, es un resultado de inexpresabilidad, que da limitaciones muy fuertes a la clase de propiedades que son expresables mediante la lógica de primer orden: propiedades que valen casi siempre o que valen casi nunca. En línea con lo anterior, Fagin[13] interpreta el teorema como una formalización del principio de que para toda oración de primer orden ϕ , a lo sumo una de las dos oraciones $\phi, \neg\phi$ es interesante. Por ejemplo, la colección de conjuntos equipados con una operación binaria que obedece los axiomas de

un grupo es ciertamente una clase interesante, pero difícilmente alguien encuentre algo de interés que decir de la colección de conjuntos equipados con una operación binaria que **no** obedece los axiomas de un grupo.

El teorema de Fagin maravillo a numerosos matemáticos y en las décadas siguientes dio lugar a una rica área de investigación dentro de la teoría de modelos finitos. De manera muy esquemática, el problema a estudiar es el siguiente: dado una clase M de estructuras, una variable aleatoria $S(n)$ con cierta distribución sobre las estructuras de \mathcal{C} y una lógica L que exprese propiedades de \mathcal{C} , clasificar el comportamiento de el siguiente límite:

$$\lim_{n \rightarrow \infty} Pr(S(n) \models \phi) \quad \text{para } \phi \in L$$

En esta tesis estudiaremos los principales resultados de esa naturaleza, viendo como el problema anterior varía enormemente según que clase de estructuras \mathcal{C} , que distribución $S(n)$ y que lógica L se consideran. Daremos énfasis a un tópico muy poco estudiado en la literatura, el de las **versiones efectivas** de los teoremas del tipo anterior, es decir, cuantificar cuan rápido converge las probabilidades y cual es la complejidad algorítmica de decidir el valor del límite, problemas de particular interés para poder aplicar dichos resultados como meta-resultados probabilísticos.

En el capítulo 1 presentaremos las herramientas de teoría de modelos finitos que usaremos en el resto de la tesis: Los juegos de Ehrenfeucht-Fraïssé, con sus diversas variantes, y los teoremas de localidad de Hanf y Gaïfman. Daremos además su motivación histórica y numerosas aplicaciones.

En el capítulo 2 demostraremos primero un caso particular de la ley de Fagin, para grafos aleatorios, y luego una versión mucho más general del teorema demostrada originalmente por Oberschelp. Demostraremos también nuevas cotas inferiores y superiores bastante precisas para la velocidad de convergencia del teorema de Fagin y el de Oberschelp, y una versión de la ley de Fagin para grafos aleatorios no etiquetados.

El capítulo 3 lo dedicaremos al teorema de convergencia demostrado por Lynch sobre grafos aleatorios muy sparse, dando una demostración alternativa. Además, demostraremos nuevas cotas inferiores en la velocidad de convergencia y la complejidad computacional de dicho resultado. Por último, demostraremos también un caso particular de un problema relacionado y también estudiado por Lynch, la ley de convergencia en espacios con funciones.

Para finalizar, el capítulo 4 está dedicado a teoremas de carácter negativo, casos donde el comportamiento es clasificable e indecidible. Demostraremos los resultados de Shelah y Compton sobre grafos ordenados aleatorios, y los resultados de Shelah, Kauffman y Le Bars sobre grafos aleatorios con lógica de segundo orden monádica, dando nuevas demostraciones y aplicaciones. Además, generalizaremos el método de Shelah y Compton a nuevas clases de estructuras, incluyendo los grafos sparse ordenados y los conjuntos equipados con un orden y una permutación.

Capítulo 1

Preliminares lógicos

1.1. El misterio de las estructuras finitas

La teoría de modelos finitos estudia las propiedades de modelos de cardinal finito, a diferencia de la teoría de modelos, que estudia modelos de cualquier cardinal. Aunque a primera vista puede parecer una restricción trivial, por el contrario, requiere un cambio radical de enfoque. Eso se debe a que casi todos los resultados básicos de teoría de modelos fallan en el caso finito, por lo que resultados similares requerirán de nuevas técnicas para ser demostrados. Según Fagin, esto se debe a la naturaleza misma de las demostraciones: "*in almost all of the deeper theorems in model theory, the key to the proof is to construct the right kind of model...These models almost invariably turn out to be infinite structures*"[13]. Consideremos, por ejemplo, uno de los teoremas más importantes de teoría de modelos:

Teorema 1.1. (Compacidad)[20] *Sea T un conjunto de formulas de primer orden sobre lenguaje \mathcal{L} . Si para todo subconjunto finito S de T hay una \mathcal{L} -estructura M_S tal que $M_S \models T$, existe \mathcal{L} -estructura M tal que $M \models T$.*

Si reemplazamos « \mathcal{L} -estructura» por « \mathcal{L} -estructura finita» el teorema deja de ser cierto. El contraejemplo es muy simple: Sea $\phi_i = \langle \text{hay } i \text{ distintos elementos} \rangle$ y sea $T = \{\phi_1, \phi_2, \dots\}$. Cualquier subconjunto finito de T tiene un modelo finito: $\{\phi_1, \dots, \phi_n\}$ es satisfecho por el conjunto de $n+1$ elementos, pero por otro lado es evidente que T no tiene ningún modelo finito. La misma construcción sirve como contraejemplo al teorema de completitud de Gödel.

Otra comparación interesante nos la da el problema con el que empieza la lógica moderna: el *Entscheidungsproblem* (Problema de decisión). Consiste en, dada ϕ en lenguaje \mathcal{L} , decidir si para toda \mathcal{L} -estructura M vale que $M \models \phi$. Como demostró Turing, el problema es recursivamente enumerable (pues un algoritmo puede buscar entre toda las combinaciones de símbolos una demostración, y en caso de ser cierto, por el teorema de completitud eventualmente encontrara una) pero no co-recursivamente enumerable. Por otro lado, en el *Entscheidungsproblem finito*, que consiste en decidir si para toda L -estructura finita M vale que $M \models \phi$, la situación es inversa: Es co-recursivamente enumerable (porque un algoritmo puede buscar todas las estructuras finitas hasta hallar un contraejemplo) pero no recursivamente enumerable[5]. Este último resultado lo demostró Trakhtenbrot en 1950.

Una de las pocas herramientas que si es válida cuando nos restringimos a estructuras finitas son los juegos de Ehrenfeucht-Fraïssé. Fueron introducidos

por Roland Fraisse[15], que lo describió de forma puramente algebraica como el método *Back and Forth*, que generalizaba resultados clásicos como el famoso teorema de Cantor que establece que todos los ordenes lineales densos contables sin extremos son isomorfos. Mas tarde, Andrzej Ehrenfeucht[7] les dio la semántica de juegos que ahora es estándar. Desde entonces se convirtieron en uno de los métodos mas usados para demostrar propiedades de modelos finitos y los usaremos extensamente durante la tesis.

1.2. Los juegos de Ehrenfeucht-Fraisse

Antes de definir el juego, necesitamos definir el **rango de cuantificador**. El rango de cuantificador de una fórmula de primer orden es una medida de la complejidad de la formula, midiendo informalmente la «profundidad» de cuantificadores.

Definición 1.2. El rango de cuantificador de una fórmula ϕ se define de forma recursiva:

- $qr(\phi) = 0$ si ϕ es una fórmula atómica
- $qr(\neg\phi) = qr(\phi)$
- $qr(\phi \wedge \psi) = qr(\phi \vee \psi) = \max(qr(\phi), qr(\psi))$
- $qr(\forall x \phi) = qr(\exists x \phi) = qr(\phi) + 1$

Ejemplo 1.3. Veamos dos ejemplos en el lenguaje de grafos:

- $qr(\exists x \forall y \neg(x \sim y)) = qr(\text{“hay un vértice aislado”}) = 2$
- $qr(\forall x \forall y ((x \sim y) \vee (\exists z x \sim z \wedge z \sim y))) = qr(\text{“el diámetro es como mucho 2”}) = 3$

Observación 1.4. Si ϕ esta en forma normal prenex, $qr(\phi)$ es la cantidad de cuantificadores.

Observación 1.5. $qr(\exists!x \phi(x)) = qr(\phi) + 2$

Un hecho muy importante es que en todo lenguaje \mathcal{L} consistente en un numero finito de relaciones y símbolos de constante, hay una cantidad finita de oraciones con rango de cuantificador k hasta equivalencia lógica, como muestra la siguiente desigualdad:

Teorema 1.6. Sea \mathcal{L} un lenguaje con C símbolos de constante y a_n relaciones de aridad n para $n > 0$. Sea $f_{\mathcal{L}}(k)$ la cantidad de oraciones sobre \mathcal{L} de rango de cuantificador k hasta equivalencia lógica. Luego, tengo la siguiente desigualdad:

$$f_{\mathcal{L}}(k) \leq 2^{2^{\sum a_n \cdot (k+c)^n}}$$

Donde la torre tiene $2 \cdot k + 2$ exponentes.

Demostración. Por inducción en k . Las formulas con rango de cuantificador 0 son combinaciones booleanas de oraciones atómicas, es decir, de oraciones de la forma $R(c_1, \dots, c_k)$ donde R es una relación de \mathcal{L} (posiblemente la igualdad) y c_1, \dots, c_k son símbolos de constante de \mathcal{L} . Es fácil ver que hay $\sum a_n \cdot C^n$ oraciones atómicas, y por tanto hay $2^{2^{\sum a_n c^n}}$ combinaciones booleanas distintas de oraciones atómicas.

Para $k > 0$, sabemos que toda oración de rango de cuantificador k se escribe como combinación booleana de oraciones de la forma $\exists x \phi$, donde ϕ tiene rango de cuantificador $k - 1$. Como hay 2^{2^n} funciones booleanas en un conjunto de n elementos, se sigue la siguiente recursión:

$$f_{\mathcal{L}}(k+1) \leq 2^{2^{f_{\mathcal{L}}(k)}}$$

De lo que se deduce la proposición. □

Notemos que la cota superior anterior crece **extremadamente** rápido con k , de forma no elemental. La siguiente proposición muestra que para lenguajes con cierta complejidad, el crecimiento es efectivamente no elemental respecto a k .

Proposición 1.7. *Sea \mathcal{L} un lenguaje con una relación de aridad mayor o igual a 2. Sea $f_{\mathcal{L}}(k)$ la cantidad de oraciones sobre \mathcal{L} de rango de cuantificador k hasta equivalencia lógica. Luego, tengo la siguiente desigualdad:*

$$f_{\mathcal{L}}(k) \geq 2^{2^{\dots^k}}$$

Donde la torre tiene $k - 1$ exponentes

Demostración. Sea $R \in \mathcal{L}$ un símbolo de relación binario. Construiremos, para todo k , un conjunto S_k de fórmulas con una variable libre y un conjunto Z_k de oraciones. S_1 consistirá únicamente de la fórmula $x = x$. S_{k+1} consistirá de las siguientes formulas para todo subconjunto $U \subset S_k$:

$$\phi_U(x) = \left[\bigwedge_{\psi \in U} \exists y : R(y, x) \wedge \psi(y) \right]$$

Ahora construiremos para todo k un conjunto Z_k de oraciones con rango de cuantificador k de la siguiente manera:

$$\psi_V = \left[\bigwedge (\exists x : \phi(x) \wedge (\forall y \neg R(x, y))) \right]$$

Es fácil ver que $|S_{k+1}| = 2^{|S_k|}$ y que $|Z_{k+1}| = 2^{|S_k|}$, y mas aun, que toda oración de Z_k es satisfacible y no son equivalentes entre si. □

Definición 1.8. Sean M y N dos \mathcal{L} -estructuras. Decimos que son **k -equivalentes** (denotado $M \equiv_k N$) si cumplen las mismas oraciones ϕ con $qr(\phi) \leq k$. Decimos que son **elementariamente equivalentes** (denotado $M \equiv N$) si $M \equiv_k N$ para todo k

Definición 1.9. Mas en general, dada una lógica L con una noción de rango de cuantificador, decimos que dos \mathcal{L} -estructuras M y N son (L, k) -equivalentes, denotado $M \equiv_{L,k} N$, si cumplen las mismas oraciones $\phi \in L$ con $qr(\phi) \leq k$ y que son L -equivalentes, denotado $M \equiv_L N$ si cumplen las mismas oraciones de L .

Observación 1.10. \equiv_k es una relación de equivalencia.

Observación 1.11. Si M y N son \mathcal{L} -estructuras finitas y $M \equiv N$ se sigue que son isomorfas. No es cierto, sin embargo, para L -estructuras en general: Se puede ver que, sobre el lenguaje de ordenes totales, $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$, aunque ciertamente no son isomorfas.

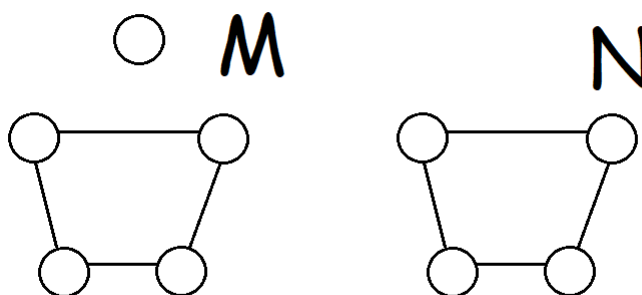
Ahora ya podemos definir el Juego de Ehrenfeucht-Fraisse:

Definición 1.12. Sea \mathcal{L} un lenguaje relacional con símbolos de constante c_i y sea $k > 0$. El **juego de Ehrenfeucht-Fraisse** se juega entre dos \mathcal{L} -estructuras M y N . Los jugadores son el **Spoiler**, cuyo objetivo es mostrar una diferencia entre las estructuras y el **Duplicator**, cuyo objetivo es mostrar que son similares. La Dinámica del juego es simple: en cada turno, el Spoiler elije un elemento de una de las estructuras, no previamente elegido, y el Duplicator responde eligiendo un elemento de la otra estructura. Tras los k turnos, quedan determinados elementos m_1, \dots, m_k de M y n_1, \dots, n_k de N . Si la correspondencia $m_i \longleftrightarrow n_i$ y $c_i^M \longleftrightarrow c_i^N$ define un isomorfismo parcial gana el Duplicator, sino gana el Spoiler.

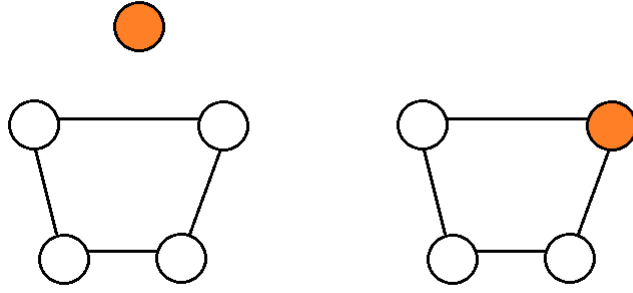
Observación 1.13. La definición contempla el "juego" de 0 turnos: Ninguno de los dos jugadores hace ningun movimiento y si la correspondencia de elementos destacados $c_{M,i} \longleftrightarrow c_{N,i}$ define un isomorfismo parcial gana el Duplicator, en caso contrario, gana el Spoiler.

Ejemplo 1.14. Veamos un ejemplo sencillo de juego de Ehrenfeucht-Fraisse en grafos:

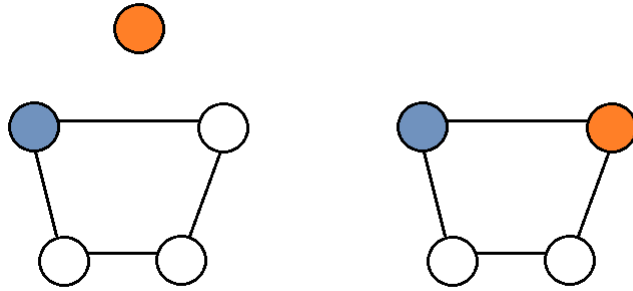
Sean M y N los grafos mostrados a continuación. Veamos que el Spoiler puede ganar en 2 turnos.



Empieza seleccionando el vértice superior de M , marcado con naranja. El Duplicator responde seleccionando a su vez un vértice de N , sin perdida de generalidad el superior de la derecha.



Ahora, el Spoiler juega en el grafo N y selecciona un vecino del vértice naranja, marcado con celeste. Como el vértice naranja del grafo M no tiene vecinos, ninguna selección que haga el Duplicator constituye un isomorfismo parcial y por tanto pierde.



Proposición 1.15. Si M y N son \mathcal{L} -estructuras isomorfas, el Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïssé entre M y N de k turnos para todo k .

Demostración. Dado isomorfismo $f : M \rightarrow N$, el Duplicator tiene la siguiente estrategia ganadora: En cada turno, si el Spoiler elige $m \in M$, el Duplicator responde $f(m) \in N$ y si el Spoiler elige $n \in N$, elige $f^{-1}(n) \in M$. Como f es un isomorfismo, en particular lo será restringido a m_1, \dots, m_k . \square

Teorema 1.16. (Teorema fundamental de los juegos E-F)[5] Sea \mathcal{L} lenguaje relacional y sean M y N dos \mathcal{L} -estructuras finitas. Luego, son equivalentes:

- El Duplicator tiene una estrategia ganadora en el juego de k turnos entre M y N
- $M \equiv_k N$

Demostración. \rightarrow) Veamos que dada ϕ con $qr(\phi) \leq k$ tal que $M \models \phi$ y $N \not\models \phi$, el Spoiler tiene estrategia ganadora en k turnos, por inducción en k . Si $k = 0$, ϕ es una fórmula solo con elementos destacados, luego no hay un isomorfismo entre los elementos destacados, por lo que el Spoiler gana el juego automáticamente.

Si $k > 0$, como ϕ es combinación booleana de formulas de la forma $\exists x \psi$, podemos asumir sin pérdida de generalidad que $\phi = \exists x \psi(x)$. Como $M \models \phi$,

hay un $m_0 \in M$ tal que $M[x = m_0] \models \psi$, pero como $N \not\models \phi$ para todo $n_0 \in N$ vale que $N[x = n_0] \not\models \psi$. Por hipótesis inductiva, el Spoiler tiene una estrategia ganadora en el juego de $k - 1$ turnos entre $M[x = m_0]$ y $N[x = n_0]$ para todo n_0 , luego tiene una estrategia ganadora en el juego de k turnos entre M y N que empieza eligiendo $m_0 \in M$ en el primer turno.

\leftarrow) Supongamos que el Spoiler tiene estrategia ganadora en k turnos y construiremos una ϕ tal que $M \models \phi$ y $N \not\models \phi$. Por inducción en k : Si $k=0$ no hay isomorfismo parcial entre los elementos destacados, eso se puede expresar con una fórmula ϕ sin cuantificadores, es decir, con $qr(\phi) = 0$.

Sea $k > 0$. Supongamos que la estrategia ganadora del Spoiler empieza eligiendo $m_1 \in M$. Para cualquier respuesta $n \in N$ que elija el Duplicator, el Spoiler tiene estrategia ganadora entre $M[x = m_1]$ y $N[x = n]$. Luego, por hipótesis inductiva, existe una fórmula ϕ_n con $qr(\phi_n) \leq k - 1$ tal que $M[x = m_1] \models \phi_n(x)$ pero $N[x = n] \not\models \phi_n(x)$. Sea ψ la fórmula $\exists x : \bigwedge_{n \in N} \phi_n(x)$. Es fácil corroborar que $M \models \psi$ y $N \not\models \psi$ y que $qr(\psi) = k$.

El otro caso es dual: Si la estrategia ganadora del Spoiler empieza eligiendo $n_0 \in N$, Para cualquier respuesta $m \in M$ que elija el Duplicator, el Spoiler tiene estrategia ganadora entre $M[x = m]$ y $N[x = n_0]$. Luego, por hipótesis inductiva, existe una fórmula ϕ_m con $qr(\phi_m) \leq k - 1$ tal que $M[x = m] \models \phi_m(x)$ pero $N[x = n_0] \not\models \phi_m(x)$. Sea ψ la fórmula $\forall x (\bigvee_{m \in M} \phi_m(x))$. Es fácil corroborar que $M \models \psi$ pero $N \not\models \psi$ y que $qr(\psi) = k$. □

Ejemplo 1.17. En el ejemplo anterior, vimos dos grafos M y N tal que el Spoiler ganaba el juego entre ellos en dos turnos. Por el teorema fundamental de los juegos E-F, dicha estrategia se corresponde con una ϕ de rango de cuantificador menor a dos tal que los distingue. Es fácil ver que la ϕ en cuestión es $(\exists x \forall y \neg(x \sim y))$

Observación 1.18. (El caso infinito) Si el lenguaje \mathcal{L} es finito, el teorema anterior es válido para estructuras infinitas, con exactamente la misma demostración: \rightarrow) Es claro, pero en \leftarrow) aparece una intersección potencialmente infinita. No es problema, porque como vimos en la proposición 1.6, para un lenguaje \mathcal{L} finito hay finitas oraciones ϕ con $qr(\phi) \leq k$, luego la intersección es equivalente a una intersección finita.

Sin embargo, si el lenguaje \mathcal{L} **no** es finito, el teorema deja de ser válido para estructuras infinitas: Es posible construir pares de estructuras M y N tal que $M \equiv N$ pero el Spoiler tenga una estrategia ganadora en el juego de Ehrenfeucht-Fraïssé entre M y N . La intuición es la siguiente: cualquier oración sobre el lenguaje \mathcal{L} solo puede usar finitos símbolos de \mathcal{L} , pero para que el Duplicator gane el juego tiene que lograr un isomorfismo parcial que involucra a todos los símbolos de \mathcal{L} , que es una tarea más difícil. Por otro lado, construir explícitamente pares de estructuras M y N que sirvan como contraejemplo no es una tarea trivial. La forma estándar es mediante un teorema profundo de teoría de modelos, el **teorema de omisión de tipos**. [20]

El teorema anterior muestra una de las propiedades fundamentales de la lógica de primer orden: **la lógica de primer orden solo es capaz de expresar**

propiedades locales, es decir, propiedades que se pueden demostrar o refutar mirando un numero fijo de elementos a la vez.

El siguiente lema nos permitirá usar los juegos E-F como herramienta para demostrar resultados de inexpressabilidad:

Lema 1.19. [21][5] *Sea \mathcal{L} un lenguaje relacional y \mathcal{C} una clase de L -estructuras finitas. Si para todo $k \in \mathbb{N}$ existe $M \in \mathcal{C}$ Y $N \notin \mathcal{C}$ tal que $M \equiv_k N$, la clase \mathcal{C} no es expresable en lógica de primer orden.*

Demostración. Supongamos que existe oración ϕ en lenguaje L que expresa la clase \mathcal{C} . Sea $k = qr(\phi)$. Por el teorema fundamental de los juegos E-F, si $M \in \mathcal{C}$ y $N \notin \mathcal{C}$ vale que $M \not\equiv_k N$, lo que contradice nuestra hipótesis. \square

Veamos algunos ejemplos de resultados de inexpressabilidad:

Proposición 1.20. *No existe ninguna oración en la lógica de primer orden de grafos finitos que exprese la propiedad «El grafo tiene una cantidad par de vértices».*

Demostración. Dado $k > 0$, sea M el grafo completo de k vértices y sea N el grafo completo de $k + 1$ vértices. Vamos a ver que $M \equiv_k N$ y como claramente M tiene una cantidad par de vértices y N una cantidad impar, se sigue por el lema 1.19 que la propiedad «El grafo tiene una cantidad par de vértices» no es expresable en lógica de primer orden de grafos.

Para ver que $M \equiv_k N$ vamos a usar el teorema fundamental de los juegos E-F, mostrando que el Duplicator tiene una estrategia ganadora en el juego entre M y N de k turnos. De hecho, vale algo mucho mas fuerte, **cualquier** estrategia que tome el Duplicator en dicho juego es ganadora. Esto se debe a que todo subconjunto de k elementos de tanto M como N es isomorfo al grafo completo de k vértices, luego al terminar el juego siempre hay un isomorfismo parcial entre los elementos elegidos. \square

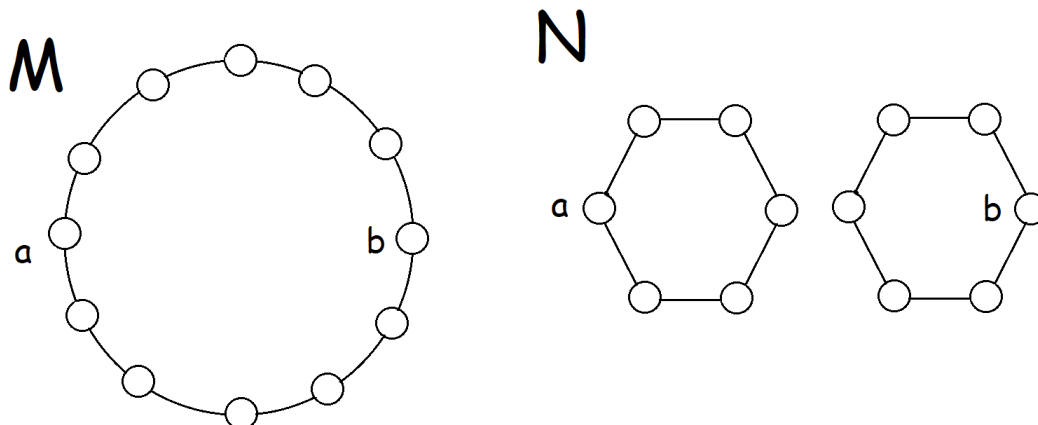
Proposición 1.21. *No existe ninguna oración en la lógica de primer orden de grafos finitos que exprese la propiedad «El grafo tiene una cantidad par de aristas»*

Demostración. Dado $k > 0$, sea M el grafo completo de $4k$ vértices y N el grafo completo de $4k + 1$ vértices. Por un argumento idéntico al presentado en la demostración anterior se sigue que $M \equiv_k N$. Por otro lado, M tiene $\frac{(4k) \cdot (4k+1)}{2} = 2 \cdot k \cdot (4k + 1)$ aristas, una cantidad par, mientras que N tiene $\frac{(4k+1)(4k+2)}{2} = (4k + 1) \cdot (2k + 1)$ aristas, una cantidad impar. Se sigue por el lema 1.19 que la propiedad «El grafo tiene una cantidad par de aristas» no es expresable por lógica de primer orden de grafos. \square

Pero la demostraciones anteriores son engañosamente sencillas. En general, demostrar resultados de inexpressabilidad mediante los juegos requiere considerablemente mas trabajo. La proposición siguiente es el ejemplo prototípico.

Proposición 1.22. [5] *No hay ninguna fórmula $\phi(a, b)$ en la lógica de primer orden de grafos finitos que exprese la propiedad "hay un camino de a a b "*

Demostración. Sea M una bola de $4 \cdot 3^k + 2$ nodos y sea N unión disjuntos de dos bolas de $2 \cdot 3^k + 2$ nodos, con elementos a y b destacados como indica la figura. Demostraremos que $M \equiv_k N$ y como M tiene un camino de a a b pero N no, se sigue por el lema 1.19 que dicha propiedad no es expresable en lógica de primer orden.



Para demostrar que $M \equiv_k N$ usaremos el teorema fundamental de los juegos E-F, dando una estrategia ganadora para el Duplicator en el juego entre M y N de k turnos. La estrategia consiste en, el turno n , responder eligiendo un elemento que preserve la distancia hasta 3^{k-n} entre los elementos elegidos. De esa forma, el turno k , la correspondencia $\{m_1, \dots, m_k\} \longleftrightarrow \{n_1, \dots, n_k\}$ preserva la distancia hasta 1, es decir, los vértices adyacentes, por lo que es efectivamente una estrategia ganadora. Notemos que como solo nos interesan las distancias hasta 3^{k-1} y todas las bolas tienen diámetro de al menos 3^k , las distancias que nos interesan se realizan en sentido horario o antihorario pero no en ambas. Veamos ahora que es siempre posible para el Duplicator mantener la estrategia, por inducción en el número de turno n .

Es fácil ver que el Duplicator puede mantener la estrategia en el primer turno. Supongamos ahora que el Duplicator consiguió mantener la estrategia hasta el turno $n - 1$. En el turno n , el Spoiler elige elemento x en alguna de las estructuras. Entre todos los elementos elegidos en turnos anteriores (y los elementos destacados a y b), Sean y y z los elementos más cercanos a x a la izquierda y derecha respectivamente, y sean y' y z' sus correspondientes en la otra estructura. Notemos que para ver que el Duplicator tiene una respuesta válida, basta ver que existe un x' entre y' y z' tal que $d(y, x) = d(x', y')$ (o ambos son mayores a) y $d(y, z) = d(y', z')$ (o ambos son mayores a), ya que para cualquier otro w elegido en turnos anteriores tenemos que $d(w, x) = d(w, y) + d(y, x)$ o $d(w, x) = d(w, z) + d(z, x)$ y por hipótesis inductiva, los elementos elegidos en los turnos anteriores respetan las distancias hasta 3^{k-n} .

Como el Duplicator mantuvo la estrategia hasta el turno $n - 1$, hay dos posibilidades:

Caso I: $d(y, z) = d(y', z')$. En este caso, como el intervalo (y, z) es de igual longitud a (y', z') , es posible hallar un x' tal que $d(y, x) = d(y', x')$ y $d(x, z) = d(x', z')$.

Caso II: $d(y, z) > 3^{k+1-n}$ y $d(y', z') > 3^{k+1-n}$. Como $d(y, z) = d(y, x) + d(x, z)$ se sigue que o bien $d(y, x) > 3^{k-n}$ o bien $d(x, z) > 3^{k-n}$. Asumamos, sin pérdida de generalidad, que $d(y, x) > 3^{k-n}$. Hay que considerar a su vez dos subcasos:

Caso II.a: $d(x, y) > 3^{k-n}$ y $d(y, z) < 3^{k-n}$. En ese caso, elegimos el único x' entre y' y z' tal que $d(x', z') = d(x, z)$. Se sigue que: $d(y', x') = d(y', z') - d(x', z') > 3^{k+1-n} - 3^{k-n} > 3^{k-n}$, por lo que mantiene la estrategia ganadora.

Caso II.b: $d(x, y) > 3^{k-n}$ y $d(y, z) > 3^{k-n}$. En ese caso, elegimos como x' el elemento exactamente a mitad de camino entre y' y z' . Se sigue que $d(x', y') = \frac{1}{2} \cdot 3^{k+1-n} > 3^{k-n}$ y $d(x', z') = \frac{1}{2} \cdot 3^{k+1-n} > 3^{k-n}$, luego mantiene la estrategia ganadora. \square

Los juegos de Ehrenfeucht-Fraïssé son una herramienta muy versátil y su uso no se limita a demostrar resultados de expresibilidad. A continuación los usaremos para dar una demostración sencilla de que dos construcciones básicas de \mathcal{L} -estructuras, el producto y el coproducto, preservan la k -equivalencia, hecho que se suele demostrar como consecuencia del teorema de Feferman-Vaught.[32]

Definición 1.23. Sea \mathcal{L} un lenguaje relacional y sean M y N dos \mathcal{L} -estructuras. El **producto** de M y N es la \mathcal{L} -estructura dada en $M \times N$ por $M \times N \models R((m_1, n_1), \dots, (m_k, n_k))$ si y solo si $M \models R(m_1, \dots, m_k)$ y $N \models R(n_1, \dots, n_k)$.

Definición 1.24. Sea \mathcal{L} un lenguaje relacional y sean M y N dos \mathcal{L} -estructuras. El **coproducto** de M y N es la \mathcal{L} -estructura en la unión disjunta dada por $M \amalg N \models R(x_1, \dots, x_k)$ si y solo si o bien todos los x_i pertenecen a M y $M \models R(x_1, \dots, x_k)$ o bien todos los x_i pertenecen a N y $N \models R(x_1, \dots, x_k)$.

Observación 1.25. El producto y el coproducto de \mathcal{L} -estructuras cumplen las propiedades universales usuales.

Teorema 1.26. Sea \mathcal{L} un lenguaje relacional y sean M_1, M_2, N_1, N_2 cuatro \mathcal{L} -estructuras tal que $M_1 \equiv_k N_1$ y $M_2 \equiv_k N_2$. Luego se sigue que:

- $M_1 \times M_2 \equiv_k N_1 \times N_2$
- $M_1 \amalg M_2 \equiv_k N_1 \amalg N_2$

Demostración. Por hipótesis, el Duplicator tiene una estrategia ganadora en el juego entre de k turnos entre M_1 y N_1 , que denotaremos τ_1 y el juego entre M_2 y N_2 de k turnos, que denotaremos por τ_2 . Vamos a ver que en ambos casos el Duplicator tiene una estrategia ganadora en el juego E-F de k turnos:

- Una partida del juego E-F de k turnos entre $M_1 \times M_2$ y $N_1 \times N_2$ se puede ver como dos partidas **simultáneas** del juego E-F, una entre M_1 y N_1 y otra entre M_2 y N_2 , ya que si al final de los k turnos quedan seleccionados $(m_1, m'_1), \dots, (m_k, m'_k) \in M_1 \times M_2$ y $(n_1, n'_1), \dots, (n_k, n'_k) \in N_1 \times N_2$ por la definición del producto la correspondencia $(m_i, m'_i) \leftrightarrow (n_i, n'_i)$ induce un isomorfismo parcial si y solo si las correspondencias $m_i \leftrightarrow n_i$ y $m'_i \leftrightarrow n'_i$ inducen isomorfismos parciales.

Luego, la estrategia ganadora del Duplicator en el juego E-F entre $M_1 \times M_2$ y $N_1 \times N_2$ consiste en aplicar las estrategias τ_1 y τ_2 simultáneamente, usando τ_1 para responder en la primer coordenada y τ_2 para responder en la segunda coordenada.

- Una partida del juego E-F de k turnos entre $M_1 \amalg M_2$ y $N_1 \amalg N_2$ se puede ver como dos partidas **paralelas** del juego E-F, una entre M_1 y N_1 y otra entre M_2 y N_2 , ya que si al final de los k turnos quedan seleccionados $m_1, \dots, m_s, m'_1, \dots, m'_t \in M_1 \amalg M_2$ y $n_1, \dots, n_s, n'_1, \dots, n'_t \in N_1 \amalg N_2$ por la definición del coproducto la correspondencia induce un isomorfismo parcial si y solo las correspondencias $m_i \leftrightarrow n_i$ y $m'_i \leftrightarrow n'_i$ inducen isomorfismos parciales.

Luego, la estrategia ganadora del Duplicator consiste en aplicar paralelamente las estrategias τ_1 y τ_2 , usando τ_1 para responder a jugadas en M_1 o N_1 y τ_2 para responder a jugadas en M_2 o N_2 .

□

1.3. Los teoremas de Localidad

1.3.1. El teorema de Hanf

La demostración anterior de que la propiedad "Hay un camino de a a b " no puede ser expresada sobre grafos finitos, por medio de juegos de Ehrenfeucht-Fraïsse, nos llevo una carilla y muy rápido degenero en complicados argumentos combinatorios. Comparémosla con la simpleza y elegancia de la demostración clásica del mismo teorema pero sin la restricción del caso finito:

Teorema 1.27. [20] *No hay ninguna fórmula en la lógica de primer orden de grafos que exprese la propiedad «hay un camino de a a b ».*

Demostración. Supongamos que $\psi(a, b)$ sea tal formula. Sea ϕ_k la siguiente oración

$$\phi_k(a, b) = [\exists x_1 (\dots) \exists x_k : (a \sim x_1) \wedge (x_1 \sim x_2) \wedge (\dots) \wedge (x_k \sim b)]$$

que expresa «Hay camino de k pasos entre a y b ». Sea T el conjunto de formulas dada por $T = \{\psi, \neg\phi_1, \neg\phi_2, \dots\}$. Es fácil ver que todo subconjunto finito de T admite modelo. Luego, por el teorema de compacidad, se sigue que T admite modelo, pero eso es absurdo: si a y b están conectados, lo están por un camino de orden k para algún k , pero como $T \models \neg\phi_k$, no hay tal camino. □

La comparación muestra la debilidad fundamental de los juegos de Ehrenfeucht-Fraïsse como herramienta: **Los juegos de Ehrenfeucht-Fraïsse son difíciles de jugar.** A diferencia de los métodos clásicos, demostrar quien es el ganador requiere una atención minuciosa al detalle y puede llegar a ser simplemente inviable para estructuras suficientemente complejas. Ni siquiera la búsqueda computacional puede ser de mucha ayuda: Se ha demostrado que el problema de decidir cual de los dos jugadores tiene una estrategia ganadora es PSPACE-completo[31], por lo que se cree fuertemente que no existe ningún algoritmo eficiente para resolverlo.

Afortunadamente, usando los juegos E-F se demostraron una serie de teoremas que formalizan la intuición mencionada anteriormente que la lógica de primer orden solo captura propiedades locales, dando condiciones que garantizan que dos estructuras dadas son k -equivalentes. Estos resultados son extremadamente utiles, porque evitan tener que hacer complicados argumentos ad-hoc para determinar el resultado de los juegos. Presentaremos aquí dos resultados que

necesitaremos mas adelante, el teorema de Hanf y el teorema de Gaifman. Un listado mas completo se encuentra en [14].

Definición 1.28. Sea L lenguaje relacional. Dada una L -Estructura M , el grafo de Gaifman $G(M)$ es el formado por los elementos de M , con $m_1 \sim m_2$ si y solo si hay una relación R de aridad k en L y tupla $c \in M^k$ tal que $M \models R(c)$ y m_1 y m_2 pertenecen a c . Es decir, $m_1 \sim m_2$ si y solo si hay alguna relación R en L que los «conecta».

Ejemplo 1.29. El grafo de Gaifman de un grafo es el mismo grafo.

Ejemplo 1.30. El grafo de Gaifman de una estructura que cuenta con un orden total \leq es el grafo completo, ya que para todo x e y o bien $x \leq y$ o $y \leq x$. Luego, ni el teorema de Hanf ni el de Gaifman no serán de ninguna ayuda. Veremos mas adelante que para trabajar con estructuras ordenadas se necesita otra noción de "localidad".

Definición 1.31. El grafo de Gaifman nos permite dar una noción de distancia a la estructura: definimos $d(m_1, m_2)$ como la distancia de m_1 a m_2 en el grafo de Gaifman y $B(m, r) = \{x \in M : d(m, x) < r\}$.

Proposición 1.32. Sea \mathcal{L} un lenguaje relacional con relaciones de aridad como mucho k . Luego la propiedad $d(m_1, m_2) \leq r$ se puede expresar con una fórmula de primer orden $\phi_r(m_1, m_2)$ de rango de cuantificador menor a $\lceil \log_2(r) \rceil + k - 2$.

Demostración. Por inducción en r . Para $r = 1$ podemos tomar la fórmula $\phi_1(m_1, m_2)$ como $\exists z_1 (\dots) \exists z : \bigvee R(c)$, donde el \bigvee es sobre todas las relaciones $R \in L$ y todas las tuplas c que contienen a m_1 y m_2 . Por definición, $qr(\phi_1) = k - 2$. Para $r > 1$ definimos $\phi_r(m_1, m_2)$ como $\exists z : \phi_{\frac{r}{2}}(m_1, z) \wedge \phi_{\frac{r}{2}}(z, m_2)$. Nos queda $qr(\phi_r) = 1 + qr(\phi_{\frac{r}{2}})$ y por hipótesis inductiva se sigue que $qr(\phi_r) \leq \lceil \log_2(r) \rceil + k - 2$ \square

Definición 1.33. Dada \mathcal{L} -estructura M , el r -tipo de $m \in M$ es la clase de isomorfismo de $B(m, r)$, con m destacado.

Observación 1.34. Si dos elementos de M tienen el mismo r -tipo, tienen el mismo s -tipo para todo $s < r$.

Teorema 1.35. (Hanf)[18][5] Sea \mathcal{L} lenguaje relacional, $k > 0$ y M y N dos \mathcal{L} -estructuras. Sea $r = 3^k$. Supongamos que todas las bolas de radio r en tanto M como N tienen menos de d elementos. Si para todo r -tipo τ se cumple una de las siguientes condiciones:

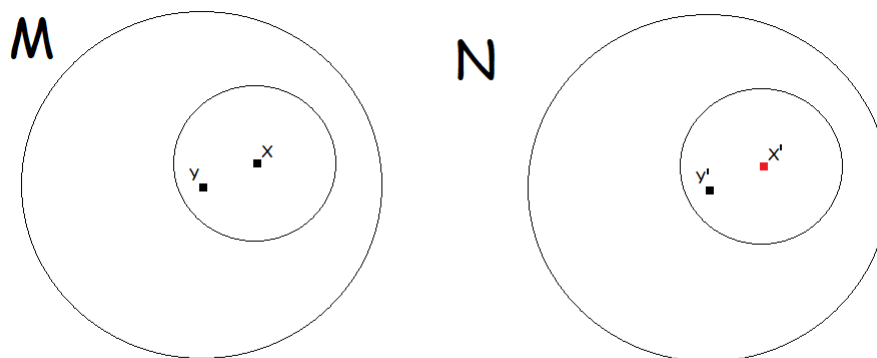
- M y N tienen la misma cantidad de elementos de r -tipo τ
- Tanto M como N tienen mas de $k \cdot d$ elementos de r -tipo τ

Se sigue que $M \equiv_k N$

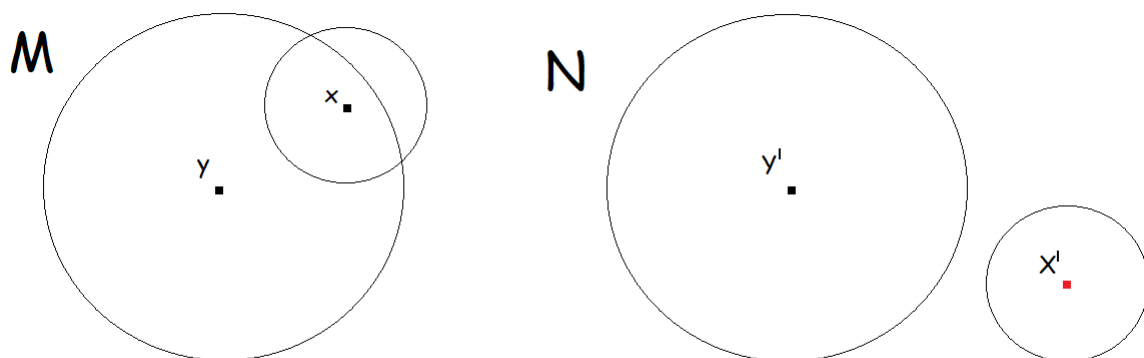
Demostración. Demostraremos que bajo dichas hipótesis el Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïsse de k turnos, de lo que se sigue que $M \equiv_k N$. Si en el turno i el Spoiler elige elemento x en alguna de las estructuras, el Duplicator deberá responder con un x' en la otra estructura tal que $B(x, 3^{k-i})$ sea isomorfo a $B(x', 3^{k-i})$ **teniendo en cuenta los elementos**

elegidos en las rondas anteriores. Es claro que si es capaz de mantener dicha estrategia los k turnos el Duplicator gana el juego, porque como para todo par de elementos destacados durante el juego sus bolas de radio 1 son isomorfas, en particular dicha correspondencia es un isomorfismo parcial.

Demostremos que el Duplicator puede mantener dicha estrategia por inducción en i . Es claro para $i = 0$, porque por las hipótesis para cualquier elemento que elija el Spoiler hay uno en la otra estructura con el mismo r -tipo. Sea $i > 0$ y supongamos que el Spoiler elija $x \in M$. Hay dos casos a analizar:



Caso I: Hay un $y \in M$ elegido en algún turno anterior a i tal que $d(x, y) \leq 3^{k-i}$. Por hipótesis Inductiva, y está en correspondencia con $y' \in N$ y $B(y, 3^{k-i+1})$ es isomorfa a $B(y', 3^{k-i+1})$. Sea f el isomorfismo. Como $B(x, 3^{k-i}) \subset B(y, 3^{k-i+1})$, se sigue que $B(x, 3^{k-i})$ es isomorfa a $B(f(x), 3^{k-i})$. Luego la estrategia del Duplicator será elegir $f(x)$. Pero esa elección depende de y ¿Que pasa si hay otro elemento z elegido en un turno anterior tal que también $d(x, z) \leq 3^{k-i}$? Si lo hubiera, $d(y, z) \leq d(y, x) + d(x, z) \leq 3^{k-i} + 3^{k-i} < 3^{k-i+1}$ y como tanto y como z fueron elegidos antes del turno $(i - 1)$ por hipótesis inductiva $B(y, 3^{k-i+1}) \cap B(z, 3^{k-i+1})$ es isomorfa a $B(y', 3^{k-i+1}) \cap B(z', 3^{k-i+1})$, luego la elección es compatible tanto con y como con z .



Caso II: No hay ningún $y \in M$ elegido en turnos anteriores a i tal que $d(x, y) \leq 3^{k-i}$. Para mantener la estrategia, el Duplicator debe responder con un x' con el mismo 3^{k-i} -tipo que x y tal que no halla ningún elemento elegidos en turnos anteriores a distancia menor a 3^{k-i} . Para ver que siempre

le será posible encontrar dicho x' , consideremos cuantos elementos con cierto 3^{k-i} -tipo y a distancia mayor a 3^{k-i} a cualquier otro elemento elegido hay. Por las hipótesis, al principio de la partida o bien ambas tienen la misma cantidad o ambas tienen más de $k \cdot 3^k$. A medida que se desarrolla el juego, esa cantidad puede reducirse, porque se pueden elegir elementos cercanos. Pero la clave es que, por la condición de isomorfismo entre las bolas, **la cantidad se reduce siempre por igual**. Luego, si empiezan siendo la misma cantidad siempre serán la misma cantidad. En particular, si el Spoiler elige una en M el Duplicator podrá responder con otra en N . Y si ambas empiezan siendo más de $d \cdot 3^k$ nunca llegarán a 0, porque cada elemento elegido "descalifica" a como mucho a d y el juego dura k rondas. \square

Observación 1.36. Enunciamos el teorema con $r = 3^k$, pero lo único que usamos de esa cantidad en la demostración es que $3^k + 3^k < 3^{k+1}$ y que $3^0 = 1$. Luego, seguirá siendo cierto con $r = (2 + \varepsilon)^k$ para cualquier $\varepsilon > 0$.

Veamos como el teorema de Hanf nos ayuda a demostrar resultados de inexpressabilidad en el caso finito dando una demostración alternativa de la proposición 1.22, que afirma que ninguna fórmula $\phi(a, b)$ en la lógica de primer orden de grafos puede expresar «hay un camino de a a b » en grafos finitos:

Demostración. Sea $k > 0$. Consideremos, de forma muy similar a la demostración de la proposición 1.22, grafos con dos vértices destacados M y N , donde M es una bola de $4 \cdot 3^k + 2$ elementos y N es la unión de dos bolas de $2 \cdot 3^k + 1$ elementos. Vamos a ver que $M \equiv_k N$ y como hay un camino de a a b en M pero no en N , se sigue por el teorema 1.19 que la propiedad de que haya un camino de a a b no es expresable.

Analicemos que 3^k -tipos aparecen en cada estructura. En M , hay, para cada $d \in \{0, \dots, 3^k\}$, exactamente 2 elementos cuyo 3^k tipo es una cadena que incluye a a a distancia d , 2 elementos cuyo 3^k tipo es una cadena que incluye a b a distancia d y no hay ningún otro 3^k -tipo. Lo mismo vale para N . Luego, por el Teorema de Hanf, se sigue que $M \equiv_k N$. \square

1.3.2. Aplicación: Decidibilidad de la lógica de primer orden monádica

Otra aplicación que muestra la utilidad del teorema de Hanf es la siguiente demostración de la decidibilidad de la lógica de primer orden monádica, el fragmento de la lógica de primer orden en el que solo hay relaciones de aridad 1, es decir, predicados. Dicho resultado, demostrado originalmente por Skolem en 1929, es de gran interés desde un punto de vista histórico, porque el paradigma que domino la lógica occidental desde la antigüedad clásica hasta la publicación del *Begriffsschrift* de Frege a finales del siglo XIX fue la lógica Aristotélica, que es equivalente a un fragmento de la lógica de primer orden monádica.

Teorema 1.37 (Skolem). [6] Sea \mathcal{L} un lenguaje relacional que contiene solo c relaciones unarias. Luego, dada oración ϕ en lenguaje \mathcal{L} con $qr(\phi) = k$, vale que ϕ es satisfacible si y solo si tiene un modelo de como mucho $2^c \cdot k$ elementos.

Demostración. Supongamos que \mathcal{L} tiene c relaciones unarias. Notemos que en una \mathcal{L} -estructura, el grafo de Gaifman es siempre vacío, luego toda bola $B(r, x)$ tiene un solo elemento. Por tanto, hay un total de 2^c r -tipos posibles para todo r , porque el r -tipo de x queda unicamente determinado por cuales de las c relaciones unarias de L cumple x .

Sea ϕ una fórmula en lenguaje \mathcal{L} con $qr(\phi) = k$. Por el teorema de Hanf, se sigue que podemos caracterizar hasta k -equivalencia las L -estructuras por la cantidad elementos que tienen de cada 1-tipo, es decir, con cada combinación posible de propiedades unarias. Como la cantidad de 1-tipos posibles es 2^c y el teorema de Hanf garantiza equivalencia si tienen la misma cantidad o mas de k , se sigue que toda clase de k -equivalencia de \mathcal{L} -estructuras contiene una estructura de menos de $2^c \cdot k$ elementos, luego, ϕ es satisfacible si y solo si tiene un modelo de menos de $2^c \cdot k$ elementos. \square

Teorema 1.38. *Sea \mathcal{L} un lenguaje relacional monádico fijo. Luego el problema de decidir si una oración ϕ en lenguaje \mathcal{L} es satisfacible es PSPACE-completo.*

Demostración. Para ver que esta en PSPACE veamos que hay un algoritmo que, dada oración ϕ en lenguaje \mathcal{L} de longitud k , decide si es satisfacible usando espacio $O(k)$.

Supongamos que \mathcal{L} tiene c relaciones unarias. Sea ϕ una oración en lenguaje \mathcal{L} de longitud k . Se sigue que $qr(\phi) \leq k$. Por el teorema 1.37 se sigue que ϕ es satisfacible si y lo es por una estructura de $2^c \cdot k$ elementos. El algoritmo enumerara todas las estructuras posibles de dicho tamaño y corroborara en cada una si cumple ϕ . Enmerar las estructuras requiere $2^c \cdot k$ espacio, al igual que comprobar en cada una si verifican ϕ .

Para ver que es PSAPCE-hard basta ver que podemos reducir el problema de determinar si una oración booleana cuantificada es cierta a determinar si una oración en el lenguaje $\mathcal{L} = \{P\}$ con una única relación unaria es satisfacible. Dada fórmula booleana cuantificadora $\phi = Q_1 p_1 (\dots) Q_k p_k \psi(p_1, \dots, p_k)$ sea $\bar{\phi} = Q_1 x_1 (\dots) Q_k x_k \psi(P(x_1), (\dots), P(x_k))$. Es fácil ver que ϕ es cierta si y solo si $\bar{\phi}$ tiene un modelo de 2 elementos. \square

Proposición 1.39. *Sea \mathcal{L} un lenguaje con c predicados unarios. Sea $f_{\mathcal{L}}(k)$ la cantidad de oraciones en \mathcal{L} de rango de cuantificador k hasta equivalencia lógica. Luego:*

$$f_{\mathcal{L}}(k) = 2^{2^c \cdot k}$$

Demostración. Por la demostración del teorema 1.37, las clases de k -equivalencia de \mathcal{L} -estructuras dependen solo de la cantidad de elementos de cada 1-tipo posible. Como hay 2^c elementos y es indiferente si una estructura tiene k o mas de k elementos de un 1-tipo dado, se sigue que hay $2^{2^c k}$ clases de k -equivalencia. Por otro lado, es fácil ver que cada una de estas es capturada por una oración de rango k , que establezca explícitamente la cantidad de elementos de cada 1-tipo. Se sigue que $f_{\mathcal{L}}(k) = 2^{2^c \cdot k}$. \square

La situación cambia radicalmente cuando pasamos a considerar el problema de decidir si una oración es satisfacible en un lenguaje monádico no fijo:

Teorema 1.40. [6] *El problema decidir, si una oración ϕ en un lenguaje monádico es satisfacible, es NEXPTIME-completo.*

Demostración. Empecemos viendo que el problema esta en NEXPTIME, dando un algoritmo no deterministico que decide el problema en tiempo exponencial. Sea ϕ una oración de longitud n . Como una oración de longitud n puede usar a lo sumo n predicados, se sigue por el teorema 1.37 que la oración ϕ es satisfacible si y solo si tiene un modelo de tamaño menor a $n \cdot 2^n$. Como la oración ϕ tiene como mucho n variables, determinar en una de esas estructuras si se cumple ϕ por fuerza bruta requiere probar todas las posibles asignaciones de las variables en los $n \cdot 2^n$ elementos, es decir, requiere como mucho $(n \cdot 2^n)^n = O(2^{n^3})$ pasos. Por otro lado, el no determinismo nos permite probar en todas dichas estructuras simultáneamente en $O(2^{n^3})$ pasos.

Para ver que el problema es NEXPTIME-hard, vamos a generar, dada una máquina de Turing no deterministica M y $n > 0$, una oración monádica $\phi_{M,n}$ de longitud $O(n^2)$ tal que $\phi_{M,n}$ es satisfacible si y solo si M llega a un estado de aceptación desde la cinta vacía en 2^n pasos. De ello se sigue que el problema de determinar si una maquina no deterministica llega al estado de aceptación en 2^{n^k} pasos se puede codificar mediante una oración de tamaño $O(n^{2k})$. Como ese problema es NEXPTIME-hard y las reducciones son polinómicas, se sigue que decidir la validez de una oración monádica es NEXPTIME-hard.

La construcción de $\phi_{M,n}$ es la siguiente. Sean P_1, \dots, P_n y Q_1, \dots, Q_n predicados en \mathcal{L} . Vamos a definir las relaciones $<_1$ y $<_2$ de longitud $O(n^2)$, que ordenan los elementos interpretando los predicados P_i y Q_i respectivamente como bits de un numero binario:

$$x \leq_1 y \Leftrightarrow \bigvee_i [P_i(y) \wedge \neg P_i(x) \wedge \bigwedge_{j>i} (P_j(x) \rightarrow P_j(y))]$$

$$x \leq_2 y \Leftrightarrow \bigvee_i [Q_i(y) \wedge \neg Q_i(x) \wedge \bigwedge_{j>i} (Q_j(x) \rightarrow Q_j(y))]$$

A continuación, definimos la oración $\phi_{grilla,n}$ como la conjunción de las siguientes cuatro oraciones. La primera nos dice que los elementos están unequivocamente determinados por los valores de P_i y Q_i , y las siguientes cuatro garantizan que cada elemento no maximal/minimal respecto al orden $<_1$ tiene un sucesor/antecesor inmediato y lo mismo para el orden $<_2$:

$$\forall x \forall y [(\bigwedge_i P_i(x) \leftrightarrow P_i(y)) \wedge (\bigwedge_i Q_i(x) \leftrightarrow Q_i(y))] \rightarrow (x = y)$$

$$\forall x (\exists y : y >_1 x) \rightarrow (\exists z : (z >_1 x) \wedge (z =_2 x) \wedge (\forall w (w >_1 x) \rightarrow (w \geq_1 z)))$$

$$\forall x (\exists y : y <_1 x) \rightarrow (\exists z : (z <_1 x) \wedge (z =_2 x) \wedge (\forall w (w <_1 x) \rightarrow (w \leq_1 z)))$$

$$\forall x (\exists y : y >_2 x) \rightarrow (\exists z : (z >_2 x) \wedge (z =_1 x) \wedge (\forall w (w >_2 x) \rightarrow (w \geq_2 z)))$$

$$\forall x (\exists y : y <_2 x) \rightarrow (\exists z : (z <_2 x) \wedge (z =_1 x) \wedge (\forall w (w <_2 x) \rightarrow (w \leq_2 z)))$$

No es difícil ver que $\phi_{grilla,n}$ fuerza a que la estructura dada por los ordenes $<_1$ y $<_2$ sea una grilla $m \times k$, cuyo tamaño máximo sera $2^n \times 2^n$. Por ultimo,

usando un nuevo conjunto de predicados podemos generar una oración ψ que sea cierta si y solo si la grilla codifica una computación de la maquina M que empieza desde la cinta vacía y llega a un estado de aceptación. Luego, la oración $\phi_{M,n} = \phi_{grilla,n} \wedge \psi_M$ es satisficible si y solo si la máquina M llega a un estado de aceptación en 2^n pasos.

Corolario 1.41. *El problema de decidir si una oración ϕ en un lenguaje monádico es valida en todas las estructuras coNEXPTIME-completo.*

Demostración. ϕ es valida en todas las estructuras si y solo si $\neg\phi$ no es satisficible. □

□

1.3.3. El teorema de Gaifamn

Aunque el teorema de Hanf es útil en muchos contextos, una limitación es que la hipótesis de isomorfismo local es muchas veces demasiado fuerte. Mas aun, por la dinámica de los juegos E-F, uno esperaría que la hipótesis de isomorfismo local se pueda relajar a una equivalencia local, pues eso basta para que el Duplicator gane localmente. Esa generalización es exactamente la que nos da el teorema de Gaifamn. Antes de poder enunciar el teorema, necesitamos definir algunos conceptos:

Definición 1.42. Sea \mathcal{L} un lenguaje relacional, ϕ una oración, M una \mathcal{L} -estructura, $m \in M$, $r > 0$.

$\phi_{B(m,r)}$ es la oración ϕ restringida a $B(m,r)$. Se define recursivamente:

- $(\exists x \psi)_{B(m,r)} = \exists x : ((d(m,x) \leq r) \wedge \psi_{B(m,r)}(x))$
- $(\neg\phi)_{B(m,r)} = \neg \phi_{B(m,r)}$
- $(\phi \wedge \psi)_{B(m,r)} = \phi_{B(m,r)} \wedge \psi_{B(m,r)}$

Observación 1.43. De la definición anterior se sigue que $(\phi \vee \psi)_{B(m,r)} = \phi_{B(m,r)} \vee \psi_{B(m,r)}$ y que $(\forall x \psi)_{B(m,r)} = \forall x ((d(x,m) \leq r) \rightarrow \psi_{B(m,r)}(x))$

Observación 1.44. Mas en general, dado lenguaje \mathcal{L} , oración ϕ , \mathcal{L} -estructura M y subconjunto $A \subset M$, podemos definir la oración ϕ_A de forma similar a la anterior, aunque solo sera expresable en lógica de primer orden si el subconjunto A lo es.

Proposición 1.45. *Para toda oración ϕ vale que $qr(\phi_{B(m,r)}) \leq qr(\phi) + \log_2(r)$.*

Demostración. Por inducción en el rango de cuantificador. Es obvio que vale para formulas sin cuantificadores. Sea $\phi = \exists x : \psi(x)$, con $qr(\phi) = k + 1$. Por definición tenemos que:

$$qr(\phi_{B(m,r)}) = qr(\exists x : (d(m,x) \leq r) \wedge (\psi_{B(m,r)}(x))) = 1 + qr((d(m,x) \leq r) \wedge (\psi_{B(m,r)}(x)))$$

$$qr(\phi_{B(m,r)}) = 1 + \max(qr(d(m,x) \leq r) , qr(\psi_{B(m,r)}(x)))$$

Sabemos que $d(m, x) \leq r$ se puede expresar con una oración de rango de cuantificador $\log_2(r)$ y por hipótesis inductiva sabemos que $qr(\psi_{B(m,r)}) \leq qr(\psi) + \log_2(r) = k + \log_2(r)$ luego nos queda:

$$qr(\phi_{B(m,r)}) \leq 1 + \max(\log_2(r), k + \log_2(r)) = 1 + k + \log_2(r) = qr(\phi) + \log_2(r)$$

□

Definición 1.46. Una ϕ se dice **local** si es de la forma $\exists x_1 \dots \exists x_s : (\bigwedge_{i,j} d(x_i, x_j) > 2 \cdot r) \wedge (\bigwedge_i \psi_{B(m,r)}(x_i))$. Coloquialmente, postula la existencia de x_1, \dots, x_s suficientemente lejos entre si tal que todos cumplen cierta propiedad local.

Para clasificar la complejidad de una oración local nos interesa no solo el rango de cuantificador, sino cuantas y cuan grandes son las bolas sobre las que cuantifica. Para eso introducimos la siguiente definición:

Definición 1.47. Decimos que una oración local ϕ es de clase (k, n, R) si es de la forma $\phi = \exists x_1 \dots \exists x_s : (\bigwedge_{i \neq j} d(x_i, x_j) > 2 \cdot r) \wedge (\bigwedge_i \psi_{B(x_i,r)}(x_i))$ con $s \leq n$, $r \leq R$ y $qr(\psi) \leq k$.

Ahora vamos a definir una generalización del tipo, que usaremos para generalizar el teorema de Hanf:

Definición 1.48. Sea \mathcal{L} lenguaje relacional. Dada \mathcal{L} -estructura M y elementos $m_1, m_2 \in M$, decimos que m_1 y m_2 tienen el mismo (r, k) -**pseudotipo** si para todo $s < r$, $B(m_1, s) \equiv_k B(m_2, s)$

Observación 1.49. Si m_1 y m_2 tienen el mismo r -tipo, tienen el mismo (r, k) -pseudotipo para todo k .

Observación 1.50. Si m_1 y m_2 tienen el mismo (r, k) -pseudotipo, tienen el mismo (r', k') -pseudotipo para todo $r' \leq r$, $k' \leq k$

Proposición 1.51. Para todo (r, k) -pseudotipo τ existe una fórmula ϕ con $qr(\phi) \leq k + \log_2(r)$ tal que $M \models \phi(m)$ si y solo si m tiene (r, k) -pseudotipo τ .

Demostración. La fórmula $\phi(x)$ es la conjunción de todas las formulas locales de radio menor/igual a r y rango de cuantificador k que el pseudotipo τ cumple, es decir, ϕ es de la forma $\bigwedge \psi_{B(x,s)}(x)$. El rango de cuantificador de ϕ es el máximo del rango de cuantificador de los $\psi_{B(x,s)}$ que por la proposición 1.45 es menor a $k + \log_2(s) \leq k + \log_2(r)$. □

Ahora que tenemos toda la terminología, podemos enunciar y demostrar el teorema de Gaifman. La demostración original de Gaifman[16] es por eliminación de cuantificadores, presentamos la demostración de Ebbinghaus[5], que usa juegos de Ehrenfeucht-Fraïsse.

Teorema 1.52. (Gaifman)[16] Sea \mathcal{L} un lenguaje relacional, $k > 0$ y sean M y N dos \mathcal{L} -estructuras que cumplen las mismas oraciones locales de tipo $(2 \cdot k^2, k+1, 7^k)$. Luego se sigue que $M \equiv_k N$.

Demostración. La idea de la demostración es muy similar a la del teorema de Hanf: mostraremos que bajo esas hipótesis, el Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïsse de k turnos, de lo que se sigue que $M \equiv_k N$. A diferencia de en el caso del Teorema de Hanf, sin embargo, no tenemos una hipótesis de isomorfismo local sino una de equivalencia local y trabajar con esa hipótesis mas débil requiere mas atención a los detalles.

En la demostración del teorema de Hanf, la estrategia dada para el Duplicator era la siguiente: Si en el turno i el Spoiler elije elemento x en alguna de las estructuras, el Duplicator debiera responder con un x' en la otra estructura tal que x y x' tengan el mismo $< 3^{k-i}$ -tipo **teniendo en cuenta los elementos elegidos en las rondas anteriores**. Ahora, relajaremos la hipótesis de isomorfismo a una de equivalencia, pero considerando bolas mas grandes: La estrategia consistirá en responder un x' en la otra estructura que tenga el mismo $(7^{k-i}, 2 \cdot (k-i)^2)$ -pseudotipo teniendo en cuenta los elementos elegidos en las rondas anteriores. Es claro que si el Duplicator puede mantener dicho estrategia al final los k turnos hay un isomorfismo parcial entre los elementos elegidos, luego gana el juego.

Supongamos que el Duplicator logro mantener la condición por los primeros s turnos y que fueron elegidos elementos $m_1, \dots, m_s \in M$ y $n_1, \dots, n_s \in N$. Denotemos con \bar{m} y \bar{n} las tuplas (m_1, \dots, m_s) y (n_1, \dots, n_s) respectivamente. Notemos que el hecho de que el Duplicator haya mantenido la estrategia garantiza que $\bigcup B(m_i, r) = B(\bar{m}, r)$ y $\bigcup B(n_i, r) = B(\bar{n}, r)$ son $2 \cdot (k-s)^2$ -equivalentes para todo $r \leq 7^{k-s}$

Sin perdida de generalidad el Spoiler elije elemento $m \in M$. Veamos que el Duplicator puede elegir un $n \in N$ que mantenga la condicion ganadora. Hay que considerar dos casos distintos:

Caso I: m esta a distancia menor a $2 \cdot 7^{k-s-1}$ de algún elemento de \bar{m} .

Vamos a construir una oración ϕ que postule la existencia de la respuesta que queremos para el Duplicator y usaremos la equivalencia local para probar que existe tal respuesta en N . Sea $\theta(x)$ la fórmula que establece esas distancias (por ejemplo, podria ser de la forma $d(x, m_1) = 8 \wedge d(x, m_2) > 7^{k-s-1} \wedge \dots$) y sea $\psi(x)$ la fórmula que establece el $(7^{k-s-1}, 2 \cdot (k-s-1)^2)$ -pseudotipo de m . Sabemos que θ tiene rango de cuantificador menor a $\log_2(7^{k-s-1}) \leq 3 \cdot (k-s-1)$ y por la proposición 1.51, que ψ tiene rango de cuantificador

$$\log_2(3^{k-s-1}) + (k-s-1)^2 < 3 \cdot (k-s-1) + 2 \cdot (k-s-1)^2$$

Sea $\phi = \exists x : (\theta(x) \wedge \psi(x))$. Es claro que $B(\bar{m}, 7^{k-s}) \models \theta(m) \wedge \phi(m)$, ya que $B(7^{k-s-1}, m) \subset B(7^{k-s}, \bar{m})$. Se sigue que $B(\bar{m}, 7^{k-s}) \models \phi$. El rango de cuantificador de esa fórmula es:

$$qr(\phi) = \max(qr(\theta), qr(\psi)) + 1 = qr(\psi) + 1 < 2 \cdot (k-s-1)^2 + 4 \cdot (k-s-1) + 1 = 2 \cdot (k-s)^2.$$

Como el Duplicator mantuvo la estrategia ganadora, sabemos que $B(7^{k-s}, \bar{m})$ y $B(7^{k-s}, \bar{n})$ son $2 \cdot (k-s)^2$ -equivalentes, luego se sigue que $B(7^{k-s}, \bar{n}) \models \phi$, es decir, existe $n \in B(\bar{n}, 7^{k-s})$ tal que $B(7^{k-s}, \bar{n}) \models \theta(n) \wedge \psi(n)$, lo que significa que si el Duplicator responde con dicho n mantiene la estrategia ganadora.

Caso II: m esta a distancia mayor a $2 \cdot 7^{k-s-1}$ de todos los elementos de \bar{m} . Para definir la estrategia del Duplicator en este caso, debemos introducir las constantes T_M, S_M, T_N y S_N . T_M es la cantidad maxima (hasta $k+1$) de elementos en M con el $(7^{k-s-1}, 2 \cdot (k-s-1)^2)$ -pseudotipo de m y a distancia mayor a $4 \cdot 7^{k-s-1}$ entre si. S_M es la cantidad maxima (hasta $k+1$) de elementos en $B(\bar{m}, 2 \cdot 7^{k-s-1})$ con el $(7^{k-s-1}, 2 \cdot (k-s-1)^2)$ -pseudotipo de m y a distancia mayor a $4 \cdot 7^{k-s-1}$ entre si. T_N es la cantidad máxima (hasta $k+1$) de elementos en N con el $(7^{k-s-1}, 2 \cdot (k-s-1)^2)$ -pseudotipo de m y a distancia mayor a $4 \cdot 7^{k-s-1}$ entre si. S_N es la cantidad maxima (hasta $k+1$) de elementos en $B(\bar{n}, 2 \cdot 7^{k-s-1})$ con el $(7^{k-s-1}, 2 \cdot (k-s-1)^2)$ -pseudotipo de m y a distancia mayor a $4 \cdot 7^{k-s-1}$ entre si.

Empecemos con algunas observaciones. Para empezar, es evidente que $S_M \leq T_M$ y $S_N \leq T_N$, ya que $B(\bar{m}, 2 \cdot 7^{k-s-1})$ y $B(\bar{n}, 2 \cdot 7^{k-s-1})$ son subconjuntos de M y N respectivamente. Tambien es fácil ver que $S_M \leq k$ y $S_N \leq k$, ya que $B(\bar{m}, 2 \cdot 7^{k-s-1})$ y $B(\bar{n}, 2 \cdot 7^{k-s-1})$ estan compuestos de como mucho k bolas de radio 7^{k-s-1} , luego pueden contener como mucho k elementos a distancias mayores a 7 entre si. Por último, se puede ver que $T_M = T_N$, ya que si definimos a α_T como la siguiente oración:

$$\exists x_1 (\dots) \exists x_T : (\bigwedge d(x_i, x_j) > 4 \cdot 7^{k-s-1}) \wedge (\bigwedge \psi(x)) \quad (1.1)$$

Por definición se sigue que T_M es el mayor T (hasta $k+1$) tal que $M \models \alpha_T$ y que T_N es el mayor T (hasta $k+1$) tal que $N \models \alpha_T$. Pero las α_T son oraciones locales de tipo $(k+1, 7k, k+1)$, y como por hipótesis sabemos que M y N cumplen las mismas oraciones locales de ese tipo se sigue que $T_M = T_N$. Como el Duplicator mantuvo la estrategia ganadora hasta el turno s , sabemos que \cdot y \cdot cumplen tambien las mismas oraciones locales de tipo $(2 \cdot k^2, k+1, 7^k)$, por un argumento idéntico se sigue que $S_M = S_N$. Tenemos que considerar dos sub-casos:

El primero es si $S_M = T_M$. En ese caso se sigue que $d(m, \bar{m}) < 4 \cdot 7^{k-s-1}$, ya que si no podemos añadir m a cualquier colección de elementos en $B(\bar{m}, 7^{k-s-1})$ con el pseudotipo dado y a distancia mayor a $4 \cdot 7^{k-s-1}$ entre si y obtendríamos una mas grande. Como $2 \cdot 7^{k-s-1} + 4 \cdot 7^{k-s-1} = 6 \cdot 7^{k-s-1} < 7^{k-s}$ se sigue que $B(m, 7^{k-s-1}) \subset B(\bar{m}, 7^{k-s})$. Luego $B(\bar{m}, 7^{k-s}) \models \exists x : \theta(x) \wedge \psi(x)$, como el Duplicator mantuvo la estrategia ganadora se sigue que $B(\bar{n}, 7^{k-s}) \models \exists x : \theta(x) \wedge \psi(x)$ y el n tal que $B(\bar{n}, 7^{k-s}) \models \theta(n) \wedge \psi(n)$ es la respuesta del Duplicator que mantiene la estrategia.

El segundo caso es si $S_M < T_M$ y es una argumentación análoga. Se sigue entonces que $B(\bar{m}, 7^{k-s}) \models \exists x_1 (\dots) \exists x_{S_M+1} : (\bigwedge d(x_i, x_j) > 4 \cdot 7^{k-s-1}) \wedge (\bigwedge \psi(x))$, luego $B(\bar{n}, 7^{k-s}) \models \exists x_1 (\dots) \exists x_{S_M+1} : (\bigwedge d(x_i, x_j) > 4 \cdot 7^{k-s-1}) \wedge (\bigwedge \psi(x))$ y por tanto, hay un n que el Duplicator puede elegir para mantener la estrategia. \square

Corolario 1.53. *Sea \mathcal{L} un lenguaje relacional y sean M y N dos \mathcal{L} -estructuras que cumplen las mismas oraciones locales. Luego se sigue que $M \equiv N$.*

La importancia del teorema anterior radica en el siguiente resultado, que nos dice que cualquier oración de primer orden se puede expresar como una combinación booleana de oraciones locales, a la que llamaremos su forma normal de Gaifman.

Teorema 1.54. (Forma normal de Gaifman) Sea ϕ una oración de primer orden, con $qr(\phi) = k$. Luego existe una oración $\tilde{\phi}$ formada como combinación booleana de formulas locales de clase $(2 \cdot k^2, k + 1, 7^k)$ tal que ϕ y $\tilde{\phi}$ son equivalentes.

Demostración. Sea ϕ una fórmula de primer orden en lenguaje \mathcal{L} y sea $k = qr(\phi)$. Para abreviar, llamemos oraciones locales buenas a las oraciones locales de tipo $(2 \cdot k^2, k + 1, 7^{k-1})$. Para cada \mathcal{L} -estructura M tal que $M \models \phi$, sea $T_M = \{\psi : \psi \text{ local buena, } M \models \psi\} \cup \{\neg\psi : \psi \text{ local buena, } M \not\models \psi\}$. El conjunto de formulas $T_M \cup \{\neg\phi\}$ es insatisfacible, pues de tener modelo habria una \mathcal{L} -estructura N que cumple la mismas formulas locales buenas que M pero tal que $N \not\models \phi$, lo que por la proposición anterior es un absurdo. Luego, por el teorema de compacidad debe haber una fórmula θ_M de la forma $\psi_1 \wedge \dots \wedge \psi_k$ donde todo $\psi_i \in T_M$ y θ_M implica a ϕ .

Ahora consideremos el conjunto de formulas $T = \{\neg\theta_M : M \models \phi\}$. El conjunto $T \cup \{\phi\}$ es insatisfacible, ya que si $M \models \phi$ por definición $M \models \theta_M$. Por el teorema de compacidad se sigue que hay una fórmula θ de la forma $\neg\theta_{M_1} \wedge \dots \wedge \neg\theta_{M_k}$ que implica a $\neg\phi$, o lo que es lo mismo, tal que ϕ implica a $\neg\theta = \theta_{M_1} \vee \dots \vee \theta_{M_k}$. Por otro lado, sabemos que cada θ_{M_i} implica a ϕ , luego se sigue que ϕ y $\neg\theta$ son equivalentes, y por definición $\neg\theta$ es una combinación booleana de formulas locales buenas. \square

Ejemplo 1.55. Veamos algunos ejemplos de formulas convertidas a su equivalente local. Consideremos en el lenguaje de grafos las formulas

$$\phi_1 = \exists x \exists y \exists z : (x \sim y) \wedge (y \sim z) \wedge (z \sim x)$$

$$\phi_2 = \forall x \exists y \forall z : (x \sim z) \vee (y \sim z)$$

Es fácil ver que ϕ_1 es equivalente a la fórmula local

$$\phi_1^* = \exists x (\exists y \exists z : y \sim z \wedge y \neq x \wedge z \neq x)_{B(x,1)}$$

Mientras que requiere un poco mas de trabajo ver que ϕ_2 es equivalente a la siguiente negación de una fórmula local

$$\phi_2^* = \neg(\exists x_1 \exists x_2 \exists x_3 \bigwedge d(x_i, x_j) \geq 2)$$

1.4. Variantes de los juegos E-F

La idea de pensar la equivalencia lógica entre dos estructuras como un juego es extremadamente fructífera y no se limita al teorema antes mencionado. En teoría de modelos nos interesa trabajar tanto con lógicas mas debiles como con lógicas mas fuertes que la de primer orden, y cada una de esas lógicas tiene asociada una variante del juego de Ehrenfeucht-Fraisse, que cumplirá que dos \mathcal{L} -estructuras M y N cumplen las mismas oraciones de dicha clase si y solo si el Duplicator gana dicha dicha versión del juego E-F entre M y N . Enunciaremos muchos de esos teoremas sin demostración porque las demostraciones son totalmente análogas a la anterior y no requieren ninguna idea nueva.

1.4.1. Juegos E-F en lenguaje con funciones

Definición 1.56. Sea \mathcal{L} un lenguaje con funciones. Su relativización \mathcal{L}^* es un lenguaje relacional que extiende a \mathcal{L} con símbolos de relación R_f de aridad $k+1$ para toda función $f \in \mathcal{L}$ de aridad k .

Toda \mathcal{L} -estructura puede verse como una \mathcal{L}^* -estructura, definiendo $M \models R_f(x_1, \dots, x_k, y)$ si y solo si $M \models f(x_1, \dots, x_k) = y$

Definición 1.57. Sea \mathcal{L} un lenguaje con funciones y ϕ una fórmula en \mathcal{L} . Denotamos con ϕ^* su traslado a \mathcal{L}^*

Ejemplo 1.58. Sea $\mathcal{L} = \{f\}$ el lenguaje con una única función unaria y sea ϕ la oración $\exists x : f(f(f(x))) = x$. Luego $\phi^*(x) = \exists y \exists z : R_f(x, y) \wedge R_f(y, z) \wedge R_f(z, x)$

Observación 1.59. Vale siempre que $qr(\phi) \leq qr(\phi^*)$ y en general la desigualdad es estricta, como muestra el ejemplo anterior.

Definición 1.60. Sea \mathcal{L} un lenguaje con funciones y sean M y N dos \mathcal{L} -estructuras. El juego de Ehrenfeucht-Fraïsse con funciones entre M y N es el juego de Ehrenfeucht-Fraïsse entre M y N visto como \mathcal{L}^* -estructuras.

Teorema 1.61. Sean M y N dos \mathcal{L} -estructuras finitas. Luego son equivalentes:

- El Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïsse con funciones de k turnos entre M y N .
- Para todo $\phi \in \mathcal{L}$ tal que $qr(\phi^*) = k$ vale que $M \models \phi$ si y solo si $N \models \phi$

Demostración. Consecuencia inmediata del teorema fundamental de los juegos de Ehrenfeucht-Fraïsse. \square

1.4.2. Juego E-F Unilateral

Definición 1.62. En el **juego de Ehrenfeucht-Fraïsse unilateral** el Spoiler está limitado a elegir elementos de una sola de las estructuras, M sin pérdida de generalidad. La dinámica y condición de victoria son idénticas al juego E-F usual.

Observación 1.63. Basta con que exista un isomorfismo de M a un subconjunto de N para que el Duplicator tenga estrategia ganadora en el juego de Ehrenfeucht-Fraïsse unilateral.

Definición 1.64. El fragmento positivo de la lógica de primer orden, denotado FO^+ , es la menor colección de fórmulas primer orden tal que:

- Si ϕ es atómica entonces $\phi \in FO^+$
- Si $\phi \in FO^+$ y $\psi \in FO^+$, entonces $\phi \wedge \psi \in FO^+$ y $\phi \vee \psi \in FO^+$
- Si $\phi \in FO^+$ entonces $\exists x \phi \in FO^+$

Teorema 1.65. Sea \mathcal{L} lenguaje relacional. Luego, son equivalentes:

- El Duplicator tiene una estrategia ganadora en el juego E-F unilateral de k turnos entre M y N
- $M \equiv_{FO^+,k} N$

Demostración. Restricción del teorema fundamental de los juegos de Ehrenfeucht-Fraisse. \square

Teorema 1.66. (Teorema de Hanf, versión existencial) Sea L lenguaje relacional, $k > 0$ y M y N L -estructuras. Sea $r = 3^k$. Supongamos que todas las bolas de radio r en tanto M como N tienen menos de d elementos. Si para todo r -tipo τ se cumple una de las siguientes condiciones:

- M tiene mas elementos de r -tipo τ que N .
- Tanto M como N tienen mas de $k \cdot d$ elementos de r -tipo τ .

Se sigue que $M \equiv_{FO^+,k} N$

Demostración. La idea es análoga a la del teorema de Hanf, usar las hipótesis para mostrar que el Duplicator tiene una estrategia ganadora en el juego E-F positivo entre M y N de k turnos. El argumento es casi idéntico, con la única diferencia que como el Spoiler está restringido a elegir elementos de M , solo importa que para todo r -tipo en M haya siempre al menos uno en N idéntico con el que pueda responder, pero no importa si hay más. \square

1.4.3. Juego E-F Biyectivo

Definición 1.67. Sea \mathcal{L} un lenguaje relacional y M y N dos \mathcal{L} -estructuras finitas de la misma cardinalidad. En el juego de Ehrenfeucht-Fraisse biyectivo, el Spoiler debe elegir en cada ronda una biyección $f : M \rightarrow N$. Luego, el Spoiler elige un elemento de una de las estructuras y el Duplicator está forzado a responder usando f , es decir, si el Spoiler elige $m \in M$ el Duplicator deberá elegir $f(m) \in N$ y si el Spoiler elige $n \in N$ el Duplicator deberá elegir $f^{-1}(n)$. La condición de victoria es la misma que en el juego E-F

Observación 1.68. Si el Duplicator comienza el turno eligiendo $f : M \rightarrow N$ y $f(m) = n$, es equivalente que el Spoiler elija m o n , dado que el Duplicator está obligado a elegir el otro. Luego, podemos suponer sin pérdida de generalidad que el Spoiler elige siempre elementos de M .

Definición 1.69. La lógica de primer orden con cuantificadores de conteo, denotada $FO(C)$ extiende a la lógica de primer orden con cuantificadores $\exists_{>k}$ para todo $k > 0$, con la siguiente semántica: $M \models \exists_{>k} x : \phi(x)$ si y solo si el conjunto $\{m : M \models \phi(m)\}$ tiene más de k elementos.

La lógica $FO(C)$ **no** es una extensión propiamente dicha de la lógica de primer orden, ya que toda oración en $FO(C)$ es equivalente a una oración de primer orden: $\exists_{>k} x : \phi(x)$ es equivalente a la fórmula de primer orden $\exists x_1 \dots \exists x_k : (\bigwedge_i \phi(x_i)) \wedge (\bigwedge_{i \neq j} x_i \neq x_j)$. Sin embargo, es más *expresiva* en el sentido que puede expresar ciertas propiedades con fórmulas mucho menos complejas que las requeridas en la lógica de primer orden. Además, es una extensión natural desde el punto de vista computacional, ya que verificar si hay algún elemento de M que cumple $\phi(m)$ requiere en el peor de los casos el mismo tiempo de cómputo que contar cuantos elementos de M cumplen $\phi(m)$.

Lema 1.70. *Sea A y B dos conjuntos de igual cardinalidad equipados con relaciones unarias P_1, \dots, P_n . Hay una biyección $f : A \rightarrow B$ que preserva los P_i si y solo si para todo subconjunto $S \subset \{1, \dots, n\}$ los conjuntos $\{a \in A : P_i(a) \forall i \in S\}$ y $\{b \in B : P_i(b) \forall i \in S\}$ tienen igual cardinalidad.*

Demostración. \implies) De existir una $f : A \rightarrow B$ que preserve los P_i , induce una biyección entre los conjuntos $\{a \in A : P_i(a) \forall i \in S\}$ y $\{b \in B : P_i(b) \forall i \in S\}$, luego tienen la misma cardinalidad.

\impliedby) Por inducción en n . El caso $n = 1$ es trivial. Para $n = m + 1$, por hipótesis los conjuntos $\{a \in A : P_n(a)\}$ y $\{b \in B : P_n(b)\}$ tienen la misma cardinalidad. Por hipótesis inductiva, hay una biyección f_1 entre esos dos conjuntos que preserva P_1, \dots, P_m . Por otro lado, como A y B tienen la misma cardinalidad podemos concluir que los conjuntos $\{a \in A : \neg P_n(a)\}$ y $\{b \in B : \neg P_n(b)\}$ tienen la misma cardinalidad y de nuevo, por hipótesis inductiva hay un f_2 entre esos dos conjuntos que preserva P_1, \dots, P_m . Luego, podemos definir $f : A \rightarrow B$ como $f(x) = f_1(x)$ si $A \models P_n(x)$ y $f(x) = f_2(x)$ si $A \models \neg P_n(x)$ y es fácil corroborar que f preserva P_1, \dots, P_n . \square

Teorema 1.71. (Teorema fundamental de los juegos E-F biyectivos) *Sea \mathcal{L} lenguaje relacional y sean M y N dos \mathcal{L} -estructuras de la misma cardinalidad. Luego, son equivalentes:*

- *El Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïssé biyectivo entre M y N de k turnos.*
- $M \equiv_{FO(C), k} N$

Demostración. \Leftarrow) Dada fórmula ϕ en $FO(C)$ con $qr(\phi) = k$ que distingue a M de N debemos hallar una estrategia ganadora para el Spoiler en el juego E-F revelado de k turnos. La demostración es análoga a la del juego E-F clásico y procede por inducción en k . La única diferencia es el caso $\phi = \exists_{>s} x : \psi(x)$. Si $M \models \phi$ pero $N \not\models \phi$ se sigue que el conjunto $\{m : M \models \psi(m)\}$ tiene mayor cardinal que $\{n : N \models \psi(n)\}$. Luego, para cualquier biyección $f : M \rightarrow N$ que el Duplicator elija al principio del primer turno, habrá un $m_0 \in M$ tal que $M \models \psi(m_0)$ pero $N \not\models \psi(f(m_0))$. Si el Spoiler responde en el primer turno eligiendo ese m_0 , por hipótesis inductiva tendrá una estrategia ganadora.

\implies) Dada estrategia ganadora del Spoiler en el juego E-F biyectivo de k turnos entre M y N , debemos hallar $\phi \in FO(C)$ con $qr(\phi) = k$ que distinga a M de N . Por inducción en k : El caso $k = 0$ es fácil de ver. Para $k > 0$, la estrategia ganadora del Spoiler empezara, dada biyección $f : M \rightarrow N$, eligiendo un $m_f \in M$. Luego, el Spoiler tiene una estrategia ganadora entre $M[x = m_f]$ y $N[x = f(m_f)]$, luego por hipótesis inductiva existe fórmula $\phi_f(x) \in FO(C)$ con $qr(\phi_f) = k - 1$ tal que $M[x = m_f] \models \phi_f(x)$ pero $N[x = f(m_f)] \not\models \phi_f(x)$. Llamemos T al conjunto de toda las ϕ_f . No hay ninguna biyección entre M y N que preserve T , luego por el lema anterior hay un $S \subset T$ tal que si establecemos $\psi(x) = (\bigwedge_{f \in S} \phi_f(x))$ los conjuntos $\{m : M \models \psi(m)\}$ y $\{n : N \models \psi(n)\}$ tienen distinta cantidad de elementos, luego una fórmula de la forma $\exists_{>s} x \psi(x)$ que distingue a M de N y como $qr(\psi) = k - 1$, $qr(\phi) = k$. \square

Con los juegos E-F biyectivos, podemos demostrar un análogo del teorema 1.26 para la lógica $FO(C)$

Teorema 1.72. Sea \mathcal{L} un lenguaje, y sean M_1, M_2, N_1, N_2 cuatro \mathcal{L} -estructuras finitas tal que $M_1 \equiv_{FO(C),k} N_1$ y $M_2 \equiv_{FO(C),k} N_2$. Luego:

- $M_1 \times M_2 \equiv_{FO(C),k} N_1 \times N_2$
- $M_1 \amalg M_2 \equiv_{FO(C),k} N_1 \amalg N_2$

Demostración. Por el teorema fundamental de los juegos E-F biyectivos, se sigue de las hipótesis que el Duplicator tiene una estrategia ganadora τ_1 en el juego E-F biyectivo de k turnos entre M_1 y N_1 y una estrategia ganadora τ_2 en el juego E-F biyectivo entre N_1 y N_2 . Veamos que se sigue que tiene una estrategia ganadora en el juego E-F biyectivo de k turnos entre $M_1 \times N_1$ y $N_1 \times N_2$ y entre $M_1 \amalg N_1$ y $N_1 \amalg N_2$.

- Como notamos en la demostración de 1.26, una partida del juego E-F entre $M_1 \times N_1$ y $M_2 \times N_2$ es equivalente a dos partidas simultaneas del juego E-F, una entre M_1 y N_1 y otra entre M_2 y N_2 . Luego la estrategia ganadora del Duplicator consistirá en aplicar simultáneamente las estrategias τ_1 y τ_2 : Si en el turno i la estrategia dada por τ_1 consiste en elegir el isomorfismo $f : M_1 \rightarrow M_2$ y la estrategia dada por τ_2 consiste en elegir el isomorfismo $g : N_1 \rightarrow N_2$, elegirá el isomorfismo $f \times g : M_1 \times N_1 \rightarrow M_2 \times N_2$
- La estrategia del Duplicator consiste en aplicar **paralelamente** la estrategia ganadora que tiene en cada subconjunto: Si en el turno i la estrategia dada por τ_1 consiste en elegir el isomorfismo $f : M_1 \rightarrow M_2$ y la estrategia dada por τ_2 consiste en elegir el isomorfismo $g : N_1 \rightarrow N_2$, elegirá el isomorfismo $f + g : M_1 \amalg N_1 \rightarrow M_2 \amalg N_2$

□

1.4.4. Juegos E-F de segundo orden

Definición 1.73. En los juegos de Ehrenfeucht-Fraisse de segundo orden, en cada el turno el Spoiler puede, en lugar de elegir un elemento de una estructura, elegir una relación de aridad s en una de las estructuras, es decir, un subconjunto de M^s o de N^s . El Duplicator deberá responder con una relación de aridad s en la otra estructura. Tras los k turnos, el Duplicator gana si hay un isomorfismo parcial entre los elementos elegidos teniendo en cuenta las relaciones elegidas durante el juego además de las de \mathcal{L} .

Teorema 1.74. Sea \mathcal{L} lenguaje relacional y M y N dos \mathcal{L} -estructuras. Luego son equivalentes:

- El duplicator tiene una estrategia ganadora en el juego E-F de segundo orden de k turnos entre M y N
- $M \equiv_{k,SO} N$

Demostración. Análoga al teorema fundamental de juegos E-F. □

Casi sin excepción, la lógica de segunda de orden completa es demasiado expresiva para obtener resultados interesantes e intentar jugar juegos E-F de segundo orden es una tarea condenada al fracaso. Por eso se suelen estudiar fragmentos de la lógica de segundo orden: en la **lógica de segundo orden monádica (MSO)** se permite cuantificar solo sobre conjuntos, en vez de relaciones de cualquier aridad, y en la **lógica de segundo orden existencial (ESO)** solo se permite cuantificar sobre de forma existencial sobre relaciones. Dichos fragmentos

vienen con sus respectivas versiones de los juegos E-F de segundo orden, que consisten en restringir las posibles jugadas del Spoiler a seleccionar solo conjuntos y a seleccionar solo relaciones sobre una de las estructuras, respectivamente. Por supuesto, valen los siguientes análogos del teorema anterior:

Teorema 1.75. *Sea \mathcal{L} lenguaje relacional y M y N dos \mathcal{L} -estructuras. Luego son equivalentes:*

- *El Duplicator tiene una estrategia ganadora en el juego E-F de segundo orden monádico de k turnos entre M y N*
- $M \equiv_{k,MSO} N$

Demostración. Análoga al teorema fundamental de los juegos de Ehrenfeuch-Fraisse. \square

Esta versión de los juegos es un poco mas manejable y nos permite probar ciertos resultados sobre la lógica de segundo orden monádica, como el siguiente análogo del teorema 1.26:

Teorema 1.76. *Sea \mathcal{L} un lenguaje relacional y sean M_1, M_2, N_1, N_2 cuatro \mathcal{L} -estructuras tal que $M_1 \equiv_{MSO,k} M_2$ Y $N_1 \equiv_{MSO,k} N_2$. Luego se sigue que $M_1 \amalg N_1 \equiv_{MSO,k} M_2 \amalg N_2$*

Demostración. Similar a 1.26. Por hipótesis, el Duplicator tiene una estrategia ganadora τ_1 en el juego E-F de segundo orden monadico de k turnos entre M_1 y M_1 y una estrategia ganadora τ_2 en el juego de segundo orden monádico de k turnos entre M_2 Y N_2 . Veamos que se sigue que el Duplicator tiene una estrategia ganadora en el juego E-F de segundo orden monádico de k turnos entre $M_1 \amalg M_2$ y $N_1 \amalg N_2$.

Como observamos anteriormente, un juego entre $M_1 \amalg N_1$ y $M_2 \amalg N_2$ se puede ver como dos juegos simultaneos. Si en el turno i el Spoiler elige un elemento en una de las estructuras, sin perdida de generalidad $m \in M$, el Duplicator respondera en M_2 usando τ_1 . Por otro lado, supongamos que el Spoiler elige un conjunto en una de las estructuras, $S \subset M_1 \amalg N_1$ sin perdida de generalidad. Podemos escribir a S como $S_1 \cup S_2$ con $S_1 \subset M_1$ y $S_2 \subset N_1$. El Duplicator respondera con $Z_1 \cup Z_2$ donde $Z_1 \subset M_2$ es la respuesta dada por τ_1 a S_1 y $Z_2 \subset N_2$. \square

Observación 1.77. No es cierto, sin embargo, que si $M_1 \equiv_{k,MSO} N_1$ y $M_2 \equiv_{k,MSO} N_2$ se sigue que $M_1 \times M_2 \equiv_{k,MSO} N_1 \times N_2$. Shelah nota en[32] que para conjuntos ordenados se puede construir una oración ϕ en la lógica MSO que es cierta en un producto $A \times B$ si y solo A y B son del mismo cardinal. Por otro lado, es bien sabido que existen conjuntos ordenados de diferente cardinal que son k -equivalentes. Esto se puede ver a partir del teorema de Büchi-Elgot-Trakhtenbrot [5] o mas directamente, por argumentos mediante juegos E-F de segundo orden monádicos.

A su vez, el teorema anterior nos permite demostrar algunos resultados de in-expresabilidad para lógica de segundo orden monádica:

Proposición 1.78. *No existe ninguna oración ϕ en lógica de segundo orden monádica que exprese la propiedad "la estructura tiene una cantidad par de elementos"*

Demostración. Sea M el conjunto de 2^k elementos y N el conjunto de $2^k + 1$ elementos. Veamos que el Duplicator gana el juego E-F de segundo orden monádico de k turnos entre M y N , de lo que se sigue por el lema 1.19 que no hay ninguna $\phi \in MSO$ con $qr(\phi) = k$ que exprese la propiedad "la estructura tiene una cantidad par de elementos".

La demostración es por inducción en k . El caso $k = 0$ es trivial. Para $k > 0$, supongamos que el Spoiler empieza la partida eligiendo un conjunto $S \subset M$ (La elección de un elemento se puede ver como la elección de un singleton). Sin pérdida de generalidad, podemos suponer que S tiene menos de 2^{k-1} elementos, ya que sino podemos considerar que el Spoiler elige S^c . El Duplicator responderá eligiendo un $S' \subset N$ con la misma cantidad de elementos que S . Como ahora M y N tienen a S y S' destacados respectivamente, los podemos ver como $S \amalg (M - S)$ y $S' \amalg (N - S')$ respectivamente. Como S y S' tienen la misma cardinalidad son isomorfos, en particular, $S \equiv_{k-1, MSO} S'$. Por otro lado, $M - S$ y $N - S'$ tienen 2^{k-1} y $2^{k-1} + 1$ respectivamente, luego por hipótesis inductiva $M - S \equiv_{k-1, MSO} N - S'$.

Luego, por el teorema 1.76, se sigue que $S \amalg (M - S) \equiv_{MSO, k-1} S' \amalg (N - S')$, es decir, el Duplicator tiene una estrategia ganadora en el juego E-F de segundo orden monádico de $k - 1$ turnos entre M con S destacado y N con S' destacado. Como eso vale independientemente de que subconjunto S empiece eligiendo el Spoiler, se sigue que el Duplicator tiene una estrategia ganadora en el juego E-F de segundo orden monádico de k turnos entre M y N . \square

Observación 1.79. De este ultimo resultado se sigue que la lógica de segundo orden monádica es estrictamente mas débil que la lógica de segundo orden completa, ya que la propiedad "la estructura tiene una cantidad par de elementos" se puede expresar en la siguiente oración de lógica de segundo orden existencial:

$$[\exists f \text{ funcion} : (\forall x (f(f(x)) = x) \wedge (f(x) \neq x))]$$

Es fácil ver que dada oración ϕ fija en lógica de segundo existencial el problema de comprobar si $M \models \phi$ esta en NP: Se puede usar el no determinismo para "adivinar" las relaciones sobre las que se cuantifica y despues queda ver si M cumple una oración fija de primer orden, que se comprobar en tiempo polinómico respecto al tamaño de M . El teorema de Fagin afirma, sorprendentemente, que vale el reciproco:

Teorema 1.80. (Fagin)[1] *Sea \mathcal{L} un lenguaje relacional y C una clase de \mathcal{L} -estructuras finitas. C es decidible por un algoritmo en tiempo polinómico no determinístico si y solo si existe $\phi \in ESO$ tal que $C = \{M : M \models \phi\}$*

Dicha teorema dio a comienzo a la **teoría de complejidad descriptiva**, que analiza clases de complejidad computacional segun que lógica es necesaria para describirlas. Y así como antes usamos los juegos E-F para probar resultados de inexpressabilidad para lógica de primer orden, en principio seria posible usar versiones generalizadas de los juegos E-F para probar resultados de inexpressabilidad en estas lógicas, resolviendo famosísimos problemas como P vs NP o NP vs co-NP. Desafortunadamente, ningún resultado así ha logrado demostrarse: Como dijimos antes, jugar los juegos E-F de segundo orden es extremadamente difícil y todavía no se han descubierto (si es que existen) resultados auxiliares como el teorema de Hanf o el de Gaifman que faciliten dicha tarea.

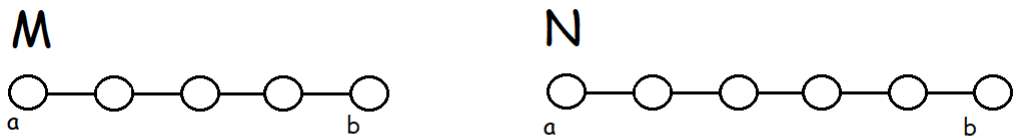
1.4.5. Juegos con fichas

Una variante de particular interes es el denominado juego de las fichas, conocida en ingles (*pebble game*).

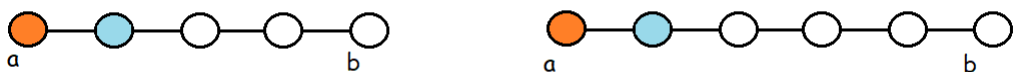
Definición 1.81. El **juego con fichas** es una generalización de carácter dinámico del juego de Ehrenfeucht-Fraisse. Sea L lenguaje relacional y M y N dos L -estructuras. El juego con k fichas y s turnos ($k \leq s$) entre M y N se juega de la siguiente manera: Hay k pares de fichas numeradas. En cada turno hay pares que están asociados a elementos de M y N respectivamente y pares que están libres. Al comienzo del juego, están todas libres. El Spoiler puede hacer dos tipos de jugadas: el primer tipo consiste en seleccionar un par de fichas libres y colocar una sobre un elemento de una de las estructuras, a lo que el Duplicator tiene que responder colocando su contraparte en la otra estructura. El segundo seleccionar una ya colocada en cierta estructura y cambiarla de posición, a lo que el Duplicator debe responder seleccionando su contraparte en la otra estructura .

En cada turno, el Duplicator debe mantener un isomorfismo parcial, de no poder hacerlo gana el Spoiler. Por otro lado, el Duplicator gana si mantiene un isomorfismo parcial por los s turnos del juego.

Ejemplo 1.82. Veamos un ejemplo del juego con fichas en grafos. Consideremos los grafos M y N , que tienen los vértices a y b destacados. Demostraremos que el Spoiler puede ganar el juego con fichas entre ellos con 2 fichas y 5 turnos



El Spoiler empieza colocando una ficha en a y la otra a su derecha. Para mantener el isomorfismo parcial, el Duplicator debe responder haciendo lo mismo.

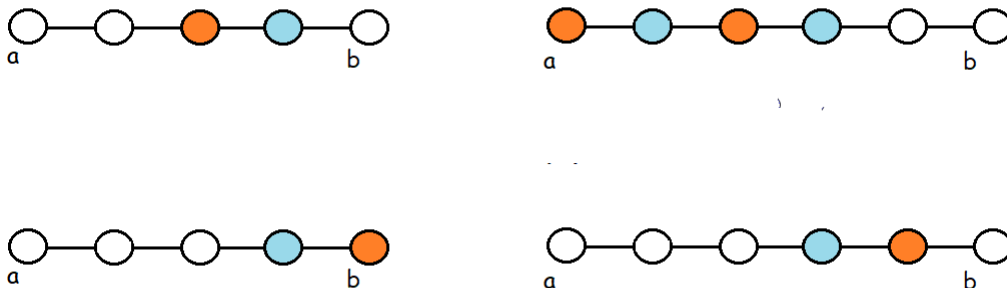


En el siguiente turno, El Spoiler mueve la ficha naranja a la derecha de la celeste en M . El Duplicator debe responder haciendo lo mismo.



El Spoiler sigue moviendo la fichas hacia la derecha en M por dos turnos mas, hasta que llega a la siguiente posición. Como la ficha naranja en M esta

en el vértice b pero su contraparte en N no, luego el isomorfismo parcial se rompe y el Spoiler gana.



Proposición 1.83. Sean M y N dos \mathcal{L} -estructuras y $s \geq k$. Tenemos la siguiente cadena de implicaciones: Spoiler gana el juego E-F de k turnos entre M y $N \implies$ Spoiler gana el juego de las fichas con k fichas y s turnos entre M y $N \implies$ Spoiler gana el juego E-F de s turnos.

Demostración. Si el Spoiler gana el juego EF de k turnos, el Spoiler gana el juego de las fichas imitando su estrategia en los primeros k turnos, es decir, en cada turno colocando una ficha nueva como lo haría en el juego E-F. Por otro lado, si el Spoiler gana el juego de las fichas con k fichas y s turnos, puede ganar el juego E-F de s turnos imitando dicha estrategia: si en el turno i el Spoiler pone una ficha en cierto elemento, en el turno i lo selecciona, y si en el turno i mueve una ficha a un nuevo elemento, selecciona este nuevo elemento. Como el Duplicator no puede ganar contra la anterior estrategia, tampoco podrá ahora ya que hay estrictamente más fichas en juego. \square

La lógica que este juego captura es un fragmento sutil de la lógica de primer orden, cuya diferencia radica en la posibilidad de re-usar cuantificadores. Veamos que significa esto: en la semántica tarskiana usual, $M \models \exists x \phi(x)$ si existe $m_0 \in M$ tal que $M[x = m_0] \models \phi$. De esta forma $M \models \exists x \exists y : \phi(x, y)$ si existen m_0 y m_1 en M tal que $M[x = m_0, y = m_1] \models \phi$. ¿Pero que pasa con una oración de la forma $\exists x \exists x : \phi(x)$? Siguiendo un razonamiento análogo al anterior, debería cumplirse si existen m_0 y m_1 tal que $M[x = m_0, x = m_1] \models \phi$. ¿Pero que interpretación le damos a $M[x = m_0, x = m_1]$? Hay tres posibilidades:

1. La primera es simplemente determinar que expresiones como $\exists x \exists x : \phi(x)$ están mal formadas, es decir, solo es válido cuantificar sobre variables libres.
2. La segunda es interpretar $M[x = m_0, x = m_1]$ como $M[x = m_0]$, en cuyo caso expresiones como la anterior están bien formadas pero la lógica no adquiere más expresividad.
3. La tercera y más interesante es interpretar $M[x = m_0, x = m_1]$ como $M[x = m_1]$ es decir, se reutiliza la variable x . Esto es análogo a la capacidad de las computadoras para reutilizar su memoria RAM (o cinta, en el formalismo de las máquinas de Turing).

Aunque en la practica la mayoría de los matemáticos se inclinan por la opción 1, casi todos los libros de lógica[36][5] definen la lógica de primer orden de acuerdo a la opción 3, es decir, permiten la reutilización de variables.

Definición 1.84. Denotamos con FO^s al fragmento de la lógica de primer orden con s variables.

La posibilidad de re-usar variables hace que las lógicas FO^n sean bastante expresivas, incluso con n pequeño. Por ejemplo:

Proposición 1.85. Para todo n , existe una fórmula $\phi_n(a, b)$ en la lógica FO^2 de grafos que expresa la propiedad Hay un camino de a a b de n pasos.

Demostración. Podemos definir las fórmulas ϕ_n recursivamente como $\phi_{n+1}(a, b) = \exists x : (a \sim x \wedge \phi_n(x, b))$. Si en cada paso introducimos una variable nueva, podemos definir a ϕ_{n+1} usando n variables. Pero es innecesario, ya que se puede alternar entre solo dos variables, x e y . Así, por ejemplo, nos queda la siguiente expresión casi caleidoscópica:

$$\phi_4(a, b) = \exists x : (a \sim x \wedge (\exists y : x \sim y \wedge (\exists x : y \sim x \wedge (\exists y : x \sim y \wedge y \sim b))))$$

□

Observación 1.86. Mas en general, se ha demostrado que toda oración en lógica modal es equivalente a una oración en FO^2 [6], lo que muestra que solo dos variables bastan para un comportamiento extremadamente rico.

Teorema 1.87. (Teorema fundamental de los juegos con fichas) Sean M y N L -estructuras finitas. Son equivalentes:

- El Duplicator tiene una estrategia ganadora en el juego de k turnos y s fichas entre M y N .
- $M \equiv_{FO^s, k} N$

Demostración. Análogo al teorema fundamental de los juegos E-F. □

Ejemplo 1.88. Vimos antes un ejemplo de dos grafos M y N con vértices destacados tal que el Spoiler ganaba el juego con 2 fichas y 5 turnos entre ellos. Por el teorema anterior tiene que haber una fórmula en FO^2 que los distinga: Es fácil ver que la fórmula $\phi_4 =$ “Hay un camino de a a b de 4 pasos” los distingue y como mostramos anteriormente $\phi_4 \in FO^2$.

Juegos y lógicas infinitarios

En esta sección presentaremos una generalización de carácter infinitario del juego de Ehrenfeucht-Fraisse, introducida originalmente por Karp[23].

Definición 1.89. El **juego con fichas infinito** es una variación del juego con fichas. Se juega entre dos L -estructuras M y N con k fichas. La dinámica del juego es igual, pero no hay un limite de turnos: el Spoiler gana si puede forzar a que el juego termine tras un numero finito de pasos, el Duplicator gana si es capaz de continuar el juego indefinidamente.

A diferencia de los juegos anteriores, que terminaban siempre tras una cantidad finita de pasos, no es inmediatamente obvio que uno de los jugadores tiene una estrategia ganadora y demostrarlo requiere el axioma de elección.

Proposición 1.90. *Sea \mathcal{L} un lenguaje relacional y sean M y N dos \mathcal{L} -estructuras. Asumiendo el axioma de elección, uno de los dos jugadores tiene una estrategia ganadora en el juego de las fichas infinitas entre M y N .*

Demostración. Sea Ω el conjunto de todas las posiciones del juego, y sea $A \subset \Omega$ el conjunto de posiciones desde las cuales el Spoiler tiene una estrategia ganadora. Observemos que si $s \notin A$, para toda jugada que haga el Spoiler desde s debe existir una respuesta del Duplicator que lleve a una posición $s' \notin A$, ya que sino existiría una estrategia ganadora del Spoiler empezando desde s .

Luego, usando el axioma de elección, podemos elegir para todo $s \notin A$ y toda jugada del Duplicator una respuesta que nos lleve a $s' \notin A$, lo que constituye una estrategia ganadora para el Duplicator empezando desde s . \square

Observación 1.91. Como detalle técnico, la demostración anterior se puede hacer con una versión mas débil que el axioma de elección general, el **axioma de elección dependiente**.

Definición 1.92. La lógica L_∞ extiende la lógica de primer orden admitiendo conjunciones y disyunciones infinitas, con la semántica intuitiva: dado conjunto S de formulas, $M \models \bigwedge_{\phi \in S} \phi$ si para todo $\phi \in S$ $M \models \phi$ y $M \models \bigvee_{\phi \in S} \phi$ si para algún $\phi \in S$ $M \models \phi$. La lógica $L_{\infty,w}$ es como la lógica L_∞ pero permite solo conjunciones y disyunciones **contables**. La lógica $L_{\infty,w}^k$ permite solo conjunciones y disyunciones contables y solo con k variables, aunque las variables se pueden «reusar» en el sentido mencionado anteriormente.

Las lógicas \mathcal{L}_∞ y $\mathcal{L}_{\infty,w}$ son de crucial importancia para teoría de modelos, pero el siguiente resultado muestra que, sobre modelos finitos, esas lógicas son tan expresivas que se vuelven completamente carentes de interés.

Proposición 1.93. *Sea \mathcal{L} un lenguaje relacional y S cualquier colección de \mathcal{L} -estructuras finitas. Existe una fórmula $\phi \in \mathcal{L}$ tal que para toda L -estructura M vale que $M \models \phi$ si y solo si $M \in S$.*

Demostración. Sabemos que para toda L -estructura finita M hay una fórmula ϕ_M que la define hasta isomorfismo. Basta entonces con tomar $\phi = \bigvee_{M \in S} \phi_M$. \square

Teniendo en cuenta eso, en teoría de modelos finitos nos interesaremos por las lógicas $L_{\infty,w}^k$, que como muestra la siguiente proposición son capaces de expresar muchas propiedades interesantes que la lógica de primer orden no puede sin ser "demasiado" expresivas.

Proposición 1.94. *Las siguientes propiedades son expresable en la lógica $L_{\infty,w}^4$ de grafos.*

- *El grafo es conexo.*
- *El grafo es bipartito.*

- El grafo es un árbol.

Demostración. Sea $\phi_n(a, b)$ la fórmula FO^2 que expresa la existencia de un camino de n pasos entre a y b , que definimos anteriormente. Luego:

- Que el grafo sea conexo es equivalente a $\forall a \forall b \bigvee_n \phi_n(a, b)$
- Que el grafo sea bipartito es equivalente a que no tenga ciclos impares, es decir, a $\forall a \bigwedge_{n \text{ impar}} \neg \phi_n(a, a)$
- Dados $c, d \in G$, podemos definir una fórmula $\phi_n(a, b, c, d) \in FO^3$ que exprese la existencia de un camino de n pasos entre a y b que no pase por el vértice cd , usando la siguiente definición recursiva: $\psi_{n+2}(a, b, c, d) = \exists x \exists y : (a \sim x \wedge x \sim y \wedge \neg(x = c \wedge y = d) \wedge \psi_n(y, b))$. Luego, como ser un árbol es equivalente a ser conexo y pero dejar de ser conexo si se remueve alguna arista, esto se puede expresar como:

$$\forall a \forall b \left(\bigvee_n \phi_n(a, b) \right) \wedge (\forall c \forall d (c \sim d) \Rightarrow (\exists a \exists b : \bigwedge_n \neg \psi_n(a, b, c, d)))$$

□

Definición 1.95. Podemos definir el rango de cuantificador para formulas en $L_{\infty, w}$ (y por restricción, en $L_{\infty, w}^k$) pero en lugar de un numero natural sera un **ordinal contable**. Se define de forma recursiva:

- $qr(\phi) = 0$ si ϕ es una fórmula atómica
- $qr(\neg\phi) = qr(\phi)$
- $qr(\exists x \phi) = qr(\phi) + 1$
- $qr(\bigwedge_n \phi_n) = \sup \{qr(\phi_n)\}$

Ejemplo 1.96. Sea $\phi_n(a, b) = \llcorner$ Hay un camino de a a b de a lo sumo n pasos \gg y consideremos las siguientes formulas en $L_{\infty, w}$:

- $\psi = \exists x_1 \exists x_2 : (\bigwedge_n \neg \phi_n(x_1, x_2)) = \llcorner$ El grafo no es conexo \gg
- $\theta = \bigwedge_k (\exists x_1 \dots \exists x_k : \bigwedge_{i, j < k; i \neq j} (\bigwedge_n \neg \phi_n(x_i, x_j))) = \llcorner$ El grafo tiene infinitas compentes conexas \gg

Luego, $qr(\psi) = \omega + 2$ y $qr(\theta) = \omega + \omega = 2 \cdot \omega$.

Teorema 1.97. (Teorema fundamental de los juegos con fichas infinitos)[5][23]
Sea \mathcal{L} lenguaje relacional y M y N dos \mathcal{L} -estructuras contables. Son equivalentes:

- El Duplicator gana el juego con k fichas infinito entre M y N
- $M \equiv_{L_{\infty, w}^k} N$

Demostración. ■ i) \implies ii) Supongamos que existe una $\phi \in L_{\infty, w}^k$ tal que $M \models \phi$ y $N \not\models \phi$ y veamos que se sigue que el Spoiler tiene una estrategia ganadora en el juego con fichas infinito entre M y N .

Por inducción **transfinita** en $qr(\phi)$. Si $qr(\phi) < \omega$ se sigue por el teorema fundamental de los juegos con fichas. Si $qr(\phi) = \alpha + 1$ podemos asumir sin pérdida de generalidad que ϕ es de la forma $\exists x \psi$ con $qr(\psi) = \alpha$, luego existe un $m_0 \in M$ tal que para todo $n \in N$ $M[x = m_0] \models \psi$ pero $N[x = n] \not\models \psi$. Por hipótesis inductiva el Spoiler tiene una estrategia ganadora entre $M[x = m_0]$ y $N[x = n]$. Luego, la estrategia ganadora entre M y N sera, si la ficha x no esta colocada, colocarla en m_0 y continuar con la otra estrategia o bien, si la ficha x ya esta colocada, moverla a m_0 y continuar con la otra estrategia.

Si $qr(\phi)$ es un ordinal limite λ , podemos asumir sin pérdida de generalidad que ϕ es de la forma $\bigwedge_n \psi_n$ con $qr(\psi_n) < \lambda$ para todo n . Si $M \models \phi$ pero $N \not\models \phi$, existe un n tal que $M \models \psi_n$ pero $N \not\models \psi_n$. Luego, por hipótesis inductiva, el Spoiler tiene una estrategia ganadora.

- ii) \implies i) Supongamos que el Spoiler tiene una cierta estrategia ganadora y demostraremos que hay una fórmula ϕ en $L_{\infty, w}^k$ que distingue a M y N . Para hacer eso, tenemos que introducir una noción de la "complejidad" de la posición, sobre la que haremos inducción transfinita.

Sea Ω el conjunto de todas las posiciones posibles del juego con k fichas infinito entre M y N . Definimos una relación binaria R en Ω de la siguiente manera: $s_1 R s_2$ si es posible que el juego evolucione de la posición s_1 a la posición s_2 **si el Spoiler juega con la estrategia ganadora**. La relación R es transitiva, antisimétrica y bien fundada: No existe una sucesión infinita s_1, s_2, \dots tal que $s_1 R s_2 R s_3 \dots$. Esto ultimo se debe a que como el Spoiler esta jugando con una estrategia ganadora, el juego siempre terminara tras un numero finito de pasos. Luego, la relación R es un **buen orden parcial**. Recordemos que en un buen orden parcial se le puede asignar a todo elemento s un tipo ordinal: un ordinal $[s]$ tal que si $s R t$ entonces $[s] > [t]$ y tal que si s es minimal $[s] = 0$.

Demostraremos ahora el teorema por inducción transfinita en el tipo ordinal de la posición s . Si $[s] = 0$ significa que el Spoiler ya gano, pues el Duplicator no puede hacer ninguna jugada valida. Luego hay una oración de primer orden sin cuantificadores que distingue a M y N . Supongamos que $[s] = \alpha > 0$ y que la estrategia ganadora del Spoiler comienza moviendo (o colocando si todavía no estaba en juego) la ficha x a m_0 . El Duplicator debe responder moviendo (o colocando) x' a algún n . Denotemos a la nueva posición s_n . Como el Spoiler esta jugando una estrategia ganadora, la posición del tablero ahora tiene un tipo ordinal $\beta < \alpha$. Luego, por hipótesis inductiva hay una $\phi_n(x)$ que distingue a $M[s_n]$ de $N[s_n]$. Por tanto, $\psi = \exists x : \bigwedge_n \phi_n(x)$ distingue a $M[s_n]$ de $N[s_n]$.

□

En el caso finito se suman otras dos equivalencias, mostrando que cuando tenemos una cota en el tamaño de las estructuras la lógica infinitaria **no** es mas expresiva que la lógica de primer orden:

Teorema 1.98. (Teorema fundamental de los juegos con fichas infinitos, versión finita)[5] Sea \mathcal{L} lenguaje relacional y M y N dos L -estructuras con menos de n elementos. Sea $m = k \cdot n^{2k}$. Son equivalentes:

1. Duplicator gana el juego con k fichas infinito entre M y N .
2. Duplicator gana el juego con k fichas y m turnos entre M y N .
3. $M \equiv_{FO^k, m} N$
4. $M \equiv_{L_{\infty, w}^k} N$

Demostración. ■ iv) \implies iii) La lógica FO^k es un fragmento de $L_{\infty, w}^k$.

- iii) \implies ii) Por el Teorema fundamental de los juegos con fichas.
- ii) \implies i) Veamos que si el Spoiler gana el juego con fichas infinito entre M y N , es capaz de hacerlo en menos de m turnos. En un turno dado, la posición del juego está definida por que fichas están en juego y donde están colocados. Si hay t pares de fichas en juego, hay como mucho n^{2t} formas distintas de colocarlas entre M y N . Luego, la cantidad de posiciones posibles del juego está acotada por $n^2 + n^4 + \dots + n^{2k} < k \cdot n^{2k} = m$. Luego, cualquier partida que dure más de m turnos deberá pasar dos veces por la misma posición. Pero si el Spoiler juega de forma óptima nunca pasarán dos veces por la misma posición. Por tanto, si el Spoiler es capaz de ganar el juego de k fichas infinito entre M y N es capaz de hacerlo en menos de m turnos.
- i) \implies iv) Por el teorema fundamental de los juegos con fichas infinitos. □

Capítulo 2

Ley 0-1 de Fagin para grafos aleatorios y otras estructuras

Este capítulo está dedicado a la ley 0-1 de Fagin. Empezaremos demostrándolo para un caso particular pero de mucha importancia para la combinatoria, los grafos aleatorios, antes de exponer una generalización de Oberschelp del resultado de Fagin a una clase mucho más amplia de estructuras.

2.1. Grafos Aleatorios

2.1.1. La ley de Fagin

Desde su introducción por Paul Erdős y Alfred Rényi en la década del 50, el estudio de los grafos aleatorios se ha convertido en una de las áreas más activas y profundas de tanto la teoría de Probabilidad como de la Combinatoria, con numerosas aplicaciones a otros campos como la física estadística, los algoritmos aleatorios y la teoría de redes. Hay muchos modelos distintos de grafos aleatorios, pero en este capítulo trabajaremos exclusivamente con el llamado grafo aleatorio de Erdős-Rényi, también conocido como grafo aleatorio uniforme:

Definición 2.1. Dado $p \in (0, 1)$, el **grafo aleatorio de Erdős-Rényi**, denotado $G(n, p)$ es una distribución de probabilidad sobre el conjunto de grafos en $\{1, 2, \dots, n\}$ tal que para todo par $i < j$ los eventos $i \sim j$ tienen probabilidad p y son independientes entre sí.

Observación 2.2. $G(n, \frac{1}{2})$ tiene una distribución uniforme sobre el conjunto de todos los grafos definidos sobre $\{1, 2, \dots, n\}$. Para abreviar, denotaremos a esa distribución simplemente como $G(n)$.

Nuestro objetivo en este capítulo será estudiar las propiedades **asintóticas** de los grafos aleatorios $G(n, p)$, es decir, el límite de las probabilidades cuando $n \rightarrow \infty$. Empecemos con algunos ejemplos sencillos:

Proposición 2.3. $\lim_{n \rightarrow \infty} Pr(G(n, p) \text{ contiene un triángulo}) = 1$

Demostración. Vamos a ver que la probabilidad de que no haya ningún triángulo tiende a 0. Dados vértices $i, j, k \in G$ denotemos con $S_{i,j,k}$ el evento "Los vértices i, j, k forman un triángulo". La probabilidad de $S_{i,j,k}$ es p^3 , pues cada uno de

los tres vértices tiene probabilidad p de aparecer y son independientes entre si. Luego:

$$Pr(\text{No hay ningún triángulo en el grafo}) = Pr\left(\bigcap S_{i,j,k}^C\right) \leq Pr(S_{1,2,3}^C \cap S_{4,5,6}^C \cap \dots)$$

Pero todos esos eventos son independientes, luego esto es igual a:

$$Pr(S_{1,2,3}^C) \cdot Pr(S_{4,5,6}^C) \cdot (\dots) = (1 - p^3)^{n/3}$$

Luego, como $(1 - p^3) < 1$, la probabilidad de que no exista ningún triángulo tiende a 0. □

Proposición 2.4. $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \text{ tiene un vértice aislado}) = 0$

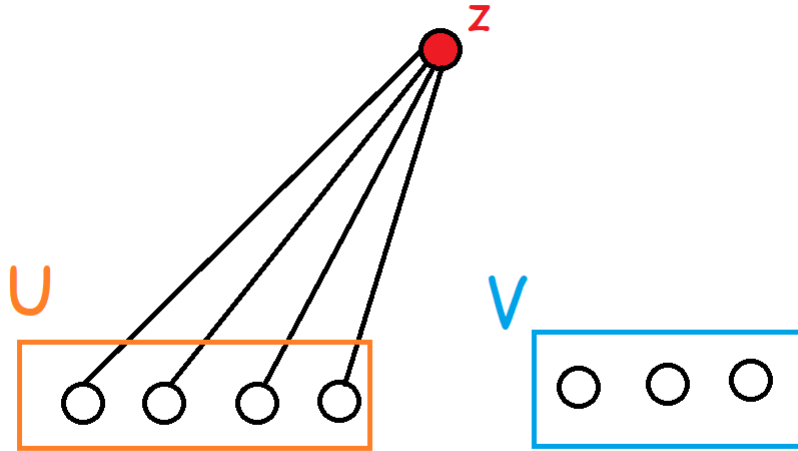
Demostración. Sea $x \in \{1, 2, \dots, n\}$. $Pr(x \text{ esta aislado}) = \prod_{y \neq x} Pr(x \not\sim y)$. Pero como todos estos eventos son independientes, es igual a $\prod_{y \neq x} Pr(x \not\sim y) = (1 - p)^{n-1}$. Luego, la probabilidad de que haya algún vértice aislado en el grafo es $Pr(\bigcup x \text{ esta aislado}) \leq \sum Pr(x \text{ esta aislado}) = n \cdot (1 - p)^{n-1}$, que tiende a 0 cuando n tiende a infinito. □

Luego, una de las propiedades tiene probabilidad asintótica 1 (pasa en casi todos los grafos) y la otra tiene propiedad asintótica 0 (no pasa en casi ningún grafo). Notemos, además, que ambas propiedades pueden ser expresadas en lógica de primer orden de grafos, por las oraciones $\phi_1 = \exists x \exists y \exists z : (x \sim y \wedge x \sim z \wedge y \sim z)$ y $\phi_2 = \exists x : \forall y \neg(x \sim y)$ respectivamente. Fagin demuestra en [12] que lo mismo ocurre para **toda** oración en lógica de primer orden de grafos:

Teorema 2.5. (Ley 0-1 de Fagin para grafos)[12] *Sea $p \in (0, 1)$ y ϕ fórmula en la lógica de primer orden de grafos. Luego $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) \in \{0, 1\}$ y mas aun, el limite no depende de la elección de p .*

Necesitamos algunas definiciones y resultados preliminares antes de demostrarlo.

Definición 2.6. Sean $s, t > 0$. Decimos que un grafo G cumple la **propiedad de (s, t) -extensión** si para toda colección de vértices distintos x_1, \dots, x_s y y_1, \dots, y_t hay un vértice z tal que $z \sim x_i$ para todo i y $z \not\sim y_i$ para todo i , en cuyo caso se dice que z extiende a $\{x_1, \dots, y_1, \dots\}$.



Por razones que serán obvias en un momentos, las propiedades de extensión son muchas veces referidas como los "axiomas de extensión" o en algunas ocasiones[37], como "la propiedad del restaurante de Alice", en referencia a la canción Alice's Restaurant Massacre de Arlo Guthrie:

Just a half a mile from the railroad track

You can get anything you want

At Alice's Restaurant

Observación 2.7. Si G cumple la propiedad de (s, t) -extensión, cumple la propiedad de (s', t') -extensión para todo $s' \leq s$ y $t' \leq t$

Definición 2.8. Decimos que un grafo cumple la propiedad de la k -extensión si cumple toda propiedad de (s, t) -extensión para todo par s, t tal que $s + t = k$

Proposición 2.9. Sean $s, t > 0$. Luego, casi todo grafo cumple la propiedad de (s, t) -extensión

Demostración. Sean x_1, \dots, x_k y y_1, \dots, y_l vértices distintos. Sea $z \notin \{x_1, \dots, y_1, \dots\}$. La probabilidad de que z extienda a $\{x_1, \dots, y_1, \dots\}$ es $p^k \cdot (1-p)^l$. Además, dicha probabilidad es independiente de si otro w extiende a $\{x_1, \dots, y_1, \dots\}$. Luego, la probabilidad de que no haya ningún elemento que extienda a $\{x_1, \dots, y_1, \dots\}$ es $(1 - p^k \cdot (1-p)^l)^{n-k-l}$. Denotemos a dicho evento $S_{\{x_1, \dots, y_1, \dots\}}$

Calculemos ahora la probabilidad de que $G(n, p)$ no cumple la propiedad de (k, l) -extensión: Es $Pr(\bigcup S_{\{x_1, \dots, y_1, \dots\}})$ donde la union es sobre todas las posibles elecciones de $\{x_1, \dots, y_1, \dots\}$. Pero esta probabilidad es menor a

$$\sum Pr(S_{\{x_1, \dots, y_1, \dots\}}) = \binom{n}{k, l} \cdot (1 - p^k \cdot (1-p)^l)^{n-k-l} \leq n^{k+l} \cdot (1 - p^k \cdot (1-p)^l)^{n-k-l}$$

Y este último término tiende a 0 cuando n tiende a infinito .

□

Corolario 2.10. Sea $k > 0$. Luego casi todo grafo cumple la propiedad de la k -extensión.

El hecho de que las propiedades de k -extensión se cumplen en casi todos los grafos nos permite demostrar el siguiente resultado, que generaliza la proposición 2.3:

Proposición 2.11. Sea H un grafo dado. Luego:

$$\lim_{n \rightarrow \infty} Pr(G(n, p) \text{ contiene } H \text{ como subgrafo inducido}) = 1$$

Demostración. Sea $G \sim G(n, p)$. Por inducción en la cantidad de vértices de H . Es trivialmente cierto si $|H| = 1$. Supongamos que $|H| = k+1$. Podemos escribir a H como $H' \cup \{v\}$ donde H' es un grafo de k vértices. Por hipótesis inductiva, H' aparece casi seguro como subgrafo inducido de $G(n, p)$ cuando $n \rightarrow \infty$, es decir, casi seguro existe una inclusión $i : H' \rightarrow G$. Sean U los elementos de H' conectados a v y sean V los elementos de H' no conectados a v . Sabemos que casi seguro el grafo G cumple la propiedad de k -extensión, luego existe un $x \in G$ tal que x está conectado a todo elemento de $i(U)$ y a ningún elemento de $i(V)$. Se sigue que i se puede extender a una inclusión $j : H \rightarrow G$ definiendo $j(v) = x$. \square

El siguiente resultado es el que conecta el resto de la argumentación, que es totalmente probabilística, con la lógica:

Teorema 2.12. Sean G y H dos grafos con la propiedad de k -extensión. Luego, el Duplicator tiene una estrategia ganadora en el juego con k fichas infinito entre G y H . En particular, tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïssé de k turnos entre G y H .

Demostración. Vamos a ver que las propiedades de extensión permiten siempre al Duplicator responder a un jugada del Spoiler manteniendo el isomorfismo parcial. Supongamos que las fichas en juego son $x_1, \dots, x_k \in G$ y $x'_1, \dots, x'_k \in H$ y que el Duplicator logra mantener un isomorfismo parcial hasta entonces. Supongamos que el Spoiler coloca una nueva ficha $x_{k+1} \in G$. Sin pérdida de generalidad, supongamos que hay un número s tal que $x_{k+1} \sim x_i$ para todo $i \leq s$ y que $x_{k+1} \not\sim x_i$ para todo i con $s < i \leq n$. Ahora, como H cumple la propiedad de $(s, k-s)$ -extensión, debe haber un elemento z de H tal que $z \sim x'_i$ para todo $i \leq s$ y que $z \not\sim x'_i$ para todo i con $s < i \leq n$. Luego es claro que si el Duplicator coloca la ficha x'_{k+1} en z sigue manteniendo el isomorfismo parcial. El mismo argumento sirve si el Duplicator elige mover una ficha ya existente en lugar de colocar una nueva. \square

Proposición 2.13. Sea $k > 0$, θ_k la oración que expresa la propiedad de k -extensión y $\gamma = (\forall x \neg(x \sim x)) \wedge (\forall x \forall y(x \sim y) \Rightarrow (y \sim x))$ la fórmula que expresa la definición de un grafo. Luego, para cualquier oración ψ perteneciente a $\mathcal{L}_{\infty, w}^k$, o bien $\theta_k \wedge \gamma$ implica ψ o bien $\theta_k \wedge \gamma$ implica $\neg\psi$.

Demostración. Por el absurdo. Si no implicara ni ψ ni $\neg\psi$, existirían grafos G y H tal que ambos cumplen θ_k pero $G \models \psi$ y $H \not\models \psi$. Por el teorema anterior, sabemos que el Duplicator tiene una estrategia ganadora en el juego con k fichas entre G y H , pues ambos cumplen suficientes propiedades de extensión. Por otro lado, por el teorema fundamental de los juegos con fichas infinitas, el Spoiler debe tener una estrategia ganadora en el juego con k fichas, pues G y H pueden ser distinguidos por una fórmula en $\mathcal{L}_{\infty, w}^k$, absurdo. \square

Corolario 2.14. Sea θ_k la oración de primer orden que expresa la propiedad de k -extensión y $\gamma = (\forall x \neg(x \sim x)) \wedge (\forall x \forall y(x \sim y) \Rightarrow (y \sim x))$ la oración que expresa la definición de un grafo. Luego el conjunto $T = \{\gamma, \theta_1, \theta_2, \dots\}$ es una teoría completa.

Ahora ya tenemos todo lo necesario para demostrar rápidamente la ley 0-1 de Fagin para grafos, mas aun, demostraremos una generalización demostrada originalmente por [26] en la que se consideran también oraciones infinitarias.

Teorema 2.15. (Ley 0-1 de Fagin, versión infinitaria)[26][5] Sea $p \in (0, 1)$ y ϕ oración en $\mathcal{L}_{\infty, w}^k$. Luego $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) \in \{0, 1\}$ y más aún, el límite no depende de la elección de p

Demostración. Sea θ_k la fórmula que expresa la propiedad de k extensión. Por el teorema anterior, o bien θ_k implica ϕ o bien θ_k implica $\neg\phi$. Supongamos que estamos en el primer caso. Luego:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) \geq \lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi \wedge \theta_k) = \lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \theta_k)$$

y como sabemos que este ultimo tiende a 1, se sigue que $Pr(\mathbf{G}(n, p) \models \phi)$ debe tender a 1 también. Si por el contrario, θ_k implica $\neg\phi$, por el mismo argumento $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \neg\phi) = 1$ y por tanto $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) = 0$. \square

Ejemplo 2.16. Vimos en el capítulo anterior que la propiedad de que un grafo sea conexo era expresable por una fórmula ϕ en $\mathcal{L}_{\infty, w}^4$. Es fácil ver que para todo p vale que $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) = 1$.

Como mencionamos en la introducción, la ley 0-1 de Fagin se puede ver como un resultado de inexpresabilidad. En el capítulo anterior, usamos los juegos de Ehrenfeucht-Fraïsse para demostrar que ciertas propiedades, como "El grafo tiene una cantidad par de aristas" no eran expresables en lógica de primer orden ni en las lógicas infinitarias $\mathcal{L}_{\infty, w}^k$, es decir, no hay ninguna oración en tales lógicas que capture exactamente esa propiedad. La ley 0-1 de Fagin nos dice que vale algo mucho mas fuerte: no existe ninguna oración en tales lógicas que capture esas propiedad en la mayoría de los casos.

Definición 2.17. Sea P una propiedad de grafos. Decimos que una oración ϕ en la lógica de primer orden de grafos **expresa asintóticamente** P si vale que:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi \mid \mathbf{G}(n) \text{ cumple } P) = 1$$

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi \mid \mathbf{G}(n) \text{ no cumple } P) = 0$$

Proposición 2.18. Para todo k , no hay ninguna oración $\phi \in \mathcal{L}_{\infty, w}^k$ sobre el lenguaje de grafos que exprese la propiedad "Hay una cantidad par de aristas" asintóticamente.

Demostración. Supongamos que existiera tal oración ϕ . Luego:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi) = \lim_{n \rightarrow \infty} Pr((n) \text{ tiene una cantidad par de aristas}) = \frac{1}{2}$$

Pero eso contradice la ley 0-1 de Fagin infinitaria. \square

Proposición 2.19. Para todo k , no hay ninguna oración $\phi \in \mathcal{L}_{\infty,w}^k$ sobre el lenguaje de grafos que exprese la propiedad "Hay una cantidad par de vértices" asintóticamente.

Demostración. Supongamos que existiera tal oración ϕ . Luego:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(2n) \models \phi) = \lim_{n \rightarrow \infty} Pr(\mathbf{G}(2n) \text{ tiene una cantidad par de vértices}) = 1$$

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}(2n+1) \models \phi) = \lim_{n \rightarrow \infty} Pr(\mathbf{G}(2n+1) \text{ tiene una cantidad par de vértices}) = 0$$

Luego se sigue que el la probabilidad asintótica de ϕ no existe, pero eso contradice la ley 0-1 de Fagin infinitaria. \square

Observación 2.20. Las propiedad "Hay una cantidad par de vértices" es expresables en lógica de segundo orden (SO), por la oración $\phi = \exists f : \forall x (f(x) \neq x) \wedge (f(f(x)) = x)$, por lo que la ley 0-1 de Fagin **no** se puede a extender a dicha lógica. De hecho, demostraremos en el capítulo 4 que ni siquiera se puede extender a la lógica de segundo orden monádica (MSO), aunque eso requerirá mas trabajo.

Proposición 2.21. (Decidibilidad de la ley 0-1 de Fagin) Hay un algoritmo que, dada ϕ en lógica de primer orden de grafos, determina el valor de $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi)$

Demostración. Sea γ la oración que expresa que una \mathcal{L} -estructura es un grafo y sea ϕ_k la oración que expresa la propiedad de k -extensión. Sea $T = \{\gamma, \theta_1, \theta_2, \dots\}$. Sabemos por la demostración de la ley 0-1 de Fagin que vale que $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) = 1$ si y solo si $T \models \phi$. Como la teoría T es completa por la proposición 2.13, si escribimos un algoritmo que enumerare todas las demostraciones posibles partiendo de T eventualmente llegara a una demostración de ϕ , en cuyo caso $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) = 1$, o a una demostración de $\neg\phi$, en cuyo caso $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n, p) \models \phi) = 0$ \square

2.1.2. Cotas para la convergencia de la ley de Fagin

El algoritmo anterior es muy poco satisfactorio, sin embargo. El tener que enumerar **todas** las demostraciones posibles implica una cota inferior exponencial en la longitud de ϕ , y peor aun, ni siquiera parece posible demostrar una cota superior, porque no sabemos que longitud tendrá la menor demostración de ϕ o $\neg\phi$ a partir de T . ¿Existirá un algoritmo mas eficiente para determinar que oraciones se cumplen en casi todos los grafos? Uno podría estar tentado de proponer un algoritmo de naturaleza probabilística: Dada oración ϕ , generar una cierta cantidad de grafos G_i del tamaño apropiado, probar cuantos de esos grafos cumplen ϕ y en base a eso determinar la probabilidad limite de ϕ . Un algoritmo así puede funcionar, pero ¿que significa el tamaño apropiado?

Para responder esa pregunta debemos encontrar una versión **efectiva** de la ley de Fagin, es decir, determinar cuan rápido converge el limite $\lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi)$, tópico que recibió escasa atención en la literatura. En su paper original [12],

Fagin observa que **fijada** una oración ϕ , el limite $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi)$ converge exponencialmente rápido, observación que se sigue inmediatamente del hecho de que fijado un k , la probabilidad de que $G(n, p)$ cumpla la propiedad de la k -extensión converge exponencialmente a 1, que vimos como parte de la demostración de la proposición 2.9.

Sin embargo, esa observación no es suficiente para nuestros propósitos. Lo que necesitamos es una cota uniforme en la velocidad de convergencia, que dependa solo del tamaño y complejidad de la oración ϕ . Vamos a empezar nuestra investigación recordando un resultado clásico de Erdős:

Proposición 2.22. [11] Sea $p \in (0, 1)$ y sea $n = p^{-\frac{k}{2}}$. Luego la probabilidad de que un grafo aleatorio $G(n, p)$ contenga un k -clique tiende a 0 cuando $k \rightarrow \infty$

Demostración. Sea $G \sim G(p^{-\frac{k}{2}}, p)$ y sea X la variable aleatoria que cuenta el número de k -cliques en G . Vamos a calcular la esperanza de X y ver que tiende a 0, y por la desigualdad de Markov $Pr(X \geq 1) \leq \mathbb{E}(X)$, lo que implica que la probabilidad de que haya un k -clique tiende a 0. Tenemos que $X = \sum_S 1_{clique(S)}$, donde la suma es sobre todo subconjunto $S \subset \{1, 2, \dots, n\}$ de k elementos. Luego:

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_S 1_{clique(S)}\right) = \sum_S \mathbb{E}(1_{clique(S)}) = \sum_S Pr(clique(S))$$

Como un k -clique tiene $\frac{k \cdot (k-1)}{2}$ aristas, la probabilidad de que un S dado sea un k -clique es $p^{\frac{k \cdot (k-1)}{2}}$. Por otro lado, hay $\binom{n}{k}$ subconjuntos de k elementos. Luego:

$$\mathbb{E}(X) = \binom{n}{k} \cdot p^{\frac{k \cdot (k-1)}{2}} < \frac{n^k}{k!} \cdot p^{\frac{k \cdot (k-1)}{2}}$$

Como $n = p^{-\frac{k}{2}}$ tenemos entonces:

$$\mathbb{E}(X) < \frac{p^{-\frac{k^2}{2}}}{k!} \cdot p^{\frac{k \cdot (k-1)}{2}} = \frac{p^{-\frac{k}{2}}}{k!} \rightarrow 0$$

□

Erdős uso el resultado anterior para probar el siguiente celebrado teorema:

Teorema 2.23. Para todo $k \geq 3$, existe un grafo de $2^{\frac{k}{2}}$ vértices que no contiene ningún k -clique ni ningún conjunto independiente de tamaño k .

Demostración. En la demostración de la proposición 2.22 mostramos que la probabilidad de que un grafo aleatorio $G(2^{\frac{k}{2}}, \frac{1}{2})$ contenga un k -clique es menor a $\frac{2^{\frac{k}{2}}}{k!}$. Por el principio de dualidad, se sigue que la probabilidad de que $G(2^{\frac{k}{2}}, \frac{1}{2})$ contenga un conjunto independiente de tamaño k también es $\frac{2^{\frac{k}{2}}}{k!}$, luego, la probabilidad de que $G(2^{\frac{k}{2}}, \frac{1}{2})$ contenga un k -clique o un k -conjunto independiente es $2 \cdot \frac{2^{\frac{k}{2}}}{k!}$, que es menor a 1 para $k \geq 3$, luego debe haber un grafo sin un k -clique ni un k -conjunto independiente. □

Observación 2.24. La demostración anterior es el primer ejemplo de lo que mas tarde se conocería como el **método probabilístico**, el uso de argumentos probabilísticos para demostrar la existencia de un objeto con ciertas propiedades, que desde su introducción por Erdős se convirtió en una herramienta estándar en combinatoria. Lo interesante de demostraciones de dicha naturaleza es que aunque prueban la existencia de un objeto, no dan **ningún** método para construirlo. A día de la fecha, seguimos sin saber de ninguna construcción explícita de un grafo de $2^{\frac{k}{2}}$ vértices sin k -cliques ni k -conjuntos independientes. De hecho, ni siquiera sabemos de una construcción de un grafo de dichas características y $(1 + \varepsilon)^k$ vértices para ningún $\varepsilon > 0$.

Pero volviendo a nuestro objetivo, la importancia de la proposición anterior es que nos da una cota inferior exponencial en la convergencia de la ley de Fagin. Sabemos, por el teorema 2.11, que la oración $\phi_k = \exists x_1 \dots \exists x_k \bigwedge_{i,j} (x_i \sim x_j)$, de rango de cuantificador k , se cumple en casi todos los grafos, pero la proposición 2.22 el numero de vértices necesario para ϕ_k que se acerque a su probabilidad limite crece al menos exponencialmente con k . La siguiente proposición muestra que una cota exponencial es suficiente para oraciones de esa naturaleza:

Proposición 2.25. Sea $p \in (0, 1)$ y $n = p^{-k}$. Luego la probabilidad de que el grafo aleatorio $G(n, p)$ tenga un k -clique tiende a 1 cuando $k \rightarrow \infty$.

Demostración. Sea X la variable aleatoria que cuenta la cantidad de k -cliques que aparecen en G . Sabemos por la demostración de la proposición anterior que:

$$\mathbb{E}(X) = \binom{n}{k} \cdot p^{\frac{k(k-1)}{2}}$$

Vamos a calcular ahora la varianza, para poder aplicar la desigualdad de Chebyshev y demostrar el teorema.

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum 1_{\text{clique}(S_i)} 1_{\text{clique}(S_j)}\right)$$

Vamos a dividir la suma según cuantos elementos tienen S_i y S_j en común, que llamaremos r . Notemos que la elección de dos conjuntos de k elementos con r elementos en común es equivalente a la elección de un conjunto de r elementos y dos conjuntos de $k - r$ elementos. Además, la probabilidad de que S_i como S_j sean cliques si tienen r elementos en común es $p^{2 \cdot \binom{k}{2} - \binom{r}{2}}$

$$\mathbb{E}(X^2) = \sum_{r=0}^k \mathbb{E}\left(\sum 1_{\text{clique}(S_i)} 1_{\text{clique}(S_j)}\right) = \sum_{r=0}^k \binom{n}{r} \cdot \binom{n-r}{k-r} \cdot \binom{n-2r}{k-r} \cdot p^{2 \cdot \binom{k}{2} - \binom{r}{2}}$$

$$\mathbb{E}(X^2) = \sum_{r=0}^k \frac{n \cdot (\dots) \cdot (n-2k+r-1)}{r! \cdot (k-r)!^2} \cdot p^{2 \cdot \binom{k}{2} - \binom{r}{2}} < \sum_{r=0}^k \frac{n^{2k-r}}{r! \cdot (k-r)!^2} \cdot p^{2 \cdot \binom{k}{2} - \binom{r}{2}}$$

Tomando $n = p^{-k}$ nos queda:

$$\mathbb{E}(X^2) < \sum_{r=0}^k \frac{(p^{-k})^{2k-r}}{r! \cdot (k-r)!^2} \cdot p^{2 \cdot \binom{k}{2} - \binom{r}{2}} = \sum_{r=0}^k \frac{1}{r! \cdot (k-r)!^2} \cdot p^{-k(2k-r) + 2 \cdot \binom{k}{2} - \binom{r}{2}}$$

Juntando los términos del denominador llegamos a:

$$\mathbb{E}(X^2) < \sum_{r=0}^k \frac{1}{r! \cdot (k-r)!^2} \cdot p^{-k^2+(r-1)k-\frac{r^2}{2}+r}$$

Notemos que para $r = 0$ la expresión nos queda igual a $\frac{1}{k!^2} \cdot p^{-k^2-k} = (\frac{1}{k!} \cdot p^{-\frac{k^2-k}{2}})^2 = \mathbb{E}(X)^2$. Luego:

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 < \sum_{r=1}^k \frac{1}{r! \cdot (k-r)!^2} \cdot p^{-k^2+(r-1)k-\frac{r^2}{2}+r}$$

Se puede ver que el exponente aumenta cuando r aumenta, luego podemos acotar la sumatoria con $r = 1$ y nos queda:

$$Var(X) < \frac{1}{k!} \cdot p^{-k^2}$$

$$\sigma(X) < \frac{1}{\sqrt{k!}} \cdot p^{-\frac{k^2}{2}}$$

Finalmente, por Chebyshev:

$$Pr(X = 0) \leq Pr(|X - p^{-\frac{k^2-k}{2}}| < p^{-\frac{k}{2}} \cdot p^{-\frac{k^2}{2}}) \leq Pr(|X - \mu| - p^{-\frac{k}{2}} \cdot \sigma < p^k \rightarrow 0$$

□

Observación 2.26. Realizando la demostración anterior con cuidado se puede ver que de hecho, basta con tomar $n = p^{-(\frac{1}{2}+\varepsilon)k}$ para cualquier $\varepsilon > 0$.

Sin embargo, las oraciones ϕ_k con las que estábamos trabajando eran oraciones existenciales, que como vimos en el capítulo anterior tiene un poder expresivo muy reducido en comparación a la lógica de primer orden completa. La siguiente proposición muestra que para oraciones mas generales la cota inferior para que la convergencia es un poco mayor, del orden de $k \cdot p^{-k}$:

Teorema 2.27. (Cota inferior para la convergencia de la ley de Fagin) Sea $p \in (0, 1)$. Para todo k existe un $c > 0$ y una oración ϕ_k en el lenguaje de los grafos con $qr(\phi_k) = k + 1$ y longitud $O(k)$ tal que ϕ_k es valida en casi todos los grafos pero

$$\lim_{k \rightarrow \infty} Pr(\mathbf{G}(c \cdot k \cdot p^{-k}, p) \models \phi_k) = 0$$

Demostración. Sea ϕ_k la oración que expresa la propiedad de $(k, 0)$ -extensión, es decir, la oración $\phi = \forall x_1 \dots \forall x_k \exists z (\bigwedge_i z \sim x_i)$. Sabemos si G es un grafo aleatorio de n elementos, fijados elementos $\{x_1, \dots, x_k\}$ de G la probabilidad de que exista un z que este conectado a todos ellos es $1 - (1 - p^k)^{n-k}$. Para acotar la probabilidad de que todo conjunto de k elementos tiene un z que lo extiende realizaremos la siguiente construcción: Particionemos G en $\frac{n}{k}$ conjuntos de k elementos llamados $A_1, \dots, A_{\frac{n}{k}}$. La probabilidad de que todo conjunto de k elementos tenga extensión es claramente menor a la probabilidad de que todo A_i tenga extensión. Además, como los A_i son disjuntos la probabilidad de que uno tenga extensión es independiente al resto. Luego, nos queda:

$$Pr(G \models \phi) < Pr(\bigcap_i A_i \text{ tiene extensión}) = \prod_i Pr(A_i \text{ tiene extensión})$$

$$Pr(G \models \phi) < (1 - (1 - p^k)^{n-k})^{\frac{n}{k}} \sim (1 - (1 - p^k)^n)^{\frac{n}{k}}$$

Este ultimo termino es igual a $(1 - (1 - p^k)^n \cdot \frac{n}{k})^{\frac{n}{k}}$ y como $(1 - p^k)^n \cdot n$ tiende a 0, podemos pasar al limite y por tanto:

$$Pr(G \models \phi) < exp(-(1 - p^k)^n \cdot \frac{n}{k})$$

Veamos que el si $n = c \cdot k \cdot p^{-k}$, el termino $(1 - p^k)^n \cdot \frac{n}{k}$ tiende a ∞ cuando k tiende a ∞ , de lo que se sigue que $exp(-(1 - p^k)^n \cdot \frac{n}{k})$ tiende a 0 cuando k tiende a ∞ . Tenemos;

$$= (1 - p^k)^{c \cdot k \cdot p^{-k}} \cdot \frac{c \cdot k \cdot p^{-k}}{k} = (1 - p^k)^{c \cdot k \cdot p^{-k}} \cdot c \cdot p^{-k} \longrightarrow e^{c \cdot k} \cdot c \cdot p^{-k}$$

Y es fácil ver que si $c > \ln(\frac{1}{p})$ el último término tiende a infinito. \square

Observación 2.28. El teorema anterior se puede mejorar para valores de $p \geq \frac{1}{2}$ observando que por simetría $G(n, p)$ y $G(n, 1 - p)$ tienen la misma velocidad de convergencia, luego para esos valores podemos usar la cota $n = c \cdot k \cdot (1 - p)^{-k}$.

Conversamente, podemos demostrar una cota superior para la convergencia de la ley de Fagin, con un orden de crecimiento ligeramente más rápido que la inferior.:

Teorema 2.29. (Cota superior para la convergencia de la ley de Fagin) Sea $p \leq \frac{1}{2}$. Para todo $\varepsilon > 0$, existe un $c > 0$ tal que para todo k y toda oración ϕ en la lógica de primer orden de grafos con $qr(\phi) = k$ vale que:

- Si ϕ es valida en casi todos los grafos, $Pr(G(c \cdot k^2 \cdot p^{-k}, p) \models \phi) > 1 - \varepsilon$
- Si ϕ no es valida en casi ningún grafo, $Pr(G(c \cdot k^2 \cdot p^{-k}, p) \models \phi) < \varepsilon$

Demostración. Demostraremos que para todo $c > \ln(\frac{1}{p})$, la probabilidad cuando $k \rightarrow \infty$ de que $G(c \cdot k^2 \cdot p^{-k}, p)$ cumpla todas las propiedades de (s, t) -extensión con $s + t = k$ tiende a 1. Luego, eligiendo un c suficientemente alto, podemos asegurarnos que para **todo** k la probabilidad de que $G(c \cdot k^2 \cdot p^{-k}, p)$ cumpla todas las propiedades de (s, t) -extensión con $s + t > k$ sea mayor a $1 - \varepsilon$, de lo que se sigue por que para todo ϕ cierta en casi todos los grafos.

La propiedad de (s, t) -extensión asegura que para todo par de subconjuntos disjuntos U, V con $|U| = s, |V| = t$, existe un $x \in G$ que extiende a (U, V) , es decir, que este conectado a todo elemento de U y ninguno de V . Luego:

$$Pr(G \text{ no cumple } (s, t)\text{-extensión}) = Pr(\bigcup_{U, V} (U, V) \text{ no tienen extensión})$$

$$Pr(G \text{ no cumple } (s, t)\text{-extensión}) \leq \sum_{U, V} Pr((U, V) \text{ no tiene extensión})$$

Vimos anteriormente que la probabilidad de que un par (U, V) dado no tuviera extensión es $(1 - p^s \cdot (1 - p)^t)^{n-k}$ luego tenemos:

$$Pr(G \text{ no cumple } (s, t)\text{-extensión}) \leq \sum_{U, V} (1-p^s \cdot (1-p)^t) = \binom{n}{s, t} \cdot (1-p^s \cdot (1-p)^t)^{n-k}$$

$$Pr(G \text{ no cumple } (s, t)\text{-extensión}) < n^k \cdot (1-p^k)^{n-k}$$

$$Pr\left(\bigcup_{s+t=k} G \text{ no cumple } (s, t)\text{-extensión}\right) \leq \sum_{s+t=k} Pr(G \text{ no cumple } (s, t)\text{-extensión}) < k \cdot n^k \cdot (1-p^k)^{n-k}$$

Queremos ver que si $n = c \cdot k^2 \cdot p^{-k}$, este ultimo termino tiende a 0 cuando $k \rightarrow \infty$. Veamos, equivalentemente, que su logaritmo natural tiende a $-\infty$ cuando $k \rightarrow \infty$. Tomando logaritmo nos queda:

$$\ln(k \cdot n^k \cdot (1-p^k)^{n-k}) = k \cdot \ln(n) + \ln(k) + (n-k) \cdot \ln(1-p^k) < k \cdot \ln(n) + \ln(k) - p^k \cdot (n-k)$$

Remplazando n por $c \cdot k^2 \cdot p^{-k}$ nos queda:

$$k \cdot \ln(c \cdot k^2 \cdot p^{-k}) + \ln(k) - p^k \cdot (p^{-k} \cdot c \cdot k^2 - k)$$

Reordenando:

$$= \left(\ln\left(\frac{1}{p}\right) - c\right) \cdot k^2 + 2 \cdot k \cdot \ln(k) + \ln(c) \cdot k + \ln(k) + p^k \cdot k$$

Luego, si $c > \ln\left(\frac{1}{p}\right)$, el termino k^2 tiene coeficiente negativo, y como el resto de los términos tienen crecimiento sub-cuadrático, se sigue que la expresión tiende a $-\infty$ cuando k tiende a infinito. \square

Observación 2.30. La cota anterior es mas sorprendente cuando recordamos que estamos acotando el tiempo de convergencia para **todas** las oraciones de rango de cuantificador k , que como vimos en el capitulo anterior pueden tener una longitud que crece **no elementariamente** con k .

Por si sola, la ley 0-1 de Fagin para grafos no es de demasiado interés probabilístico: la razón es que para la mayoría de oraciones naturales de primer orden, es intuitivo determinar si se cumplen en casi todos los grafos o en casi ninguno. Sin embargo, la cota anterior lo convierte en un resultado mucho mas útil, porque nos permite dar rápidamente cotas no triviales para la velocidad de convergencia de ciertas propiedades. Por ejemplo:

Teorema 2.31. Sea $p \leq \frac{1}{2}$. Para todo $\varepsilon > 0$, existe un $c > 0$ tal que para todo k vale que si $n = c \cdot k^2 \cdot p^{-k}$, entonces con probabilidad mayor a $1 - \varepsilon$ todo vértice en $G(n, p)$ esta contenido en un k -clique.

Demostración. La propiedad que estamos considerando esta codificada por la oración de primer orden $\phi = \forall x \exists x_2 (\dots) \exists x_k (\bigwedge x_i \sim x_j)$, que tiene rango de cuantificador k . Esa fácil ver por las propiedades de extensión que se cumple en casi todos los grafos, luego se aplica nuestra cota superior demostrado anteriormente. \square

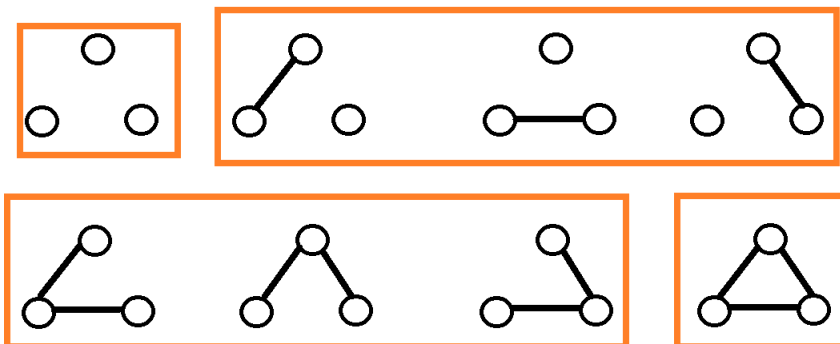
Para concluir esta sección, notemos que nuestra cota superior para la convergencia de la ley de Fagin es de orden $O(k^2 \cdot 2^k)$ mientras que nuestra cota inferior es de orden $O(k \cdot 2^k)$. Se puede ver que nuestra demostración de la cota inferior es particularmente burda, ya que acotamos la existencia de una extensión para **todo** conjunto de k elementos por la existencia de una extensión para **algunos** conjuntos de k elementos, para poder explotar la independencia. Eso nos lleva a conjeturar lo siguiente:

Conjetura: Existe un $c > 0$ tal que para todo k hay una ϕ_k con $qr(\phi_k) = k$ tal que es válida en casi todos los grafos pero $\lim_{k \rightarrow \infty} Pr(G(c \cdot k^2 \cdot 2^k, \frac{1}{2}) \models \phi_k) = 0$

2.1.3. Ley de Fagin no etiquetada

Cuando definimos los grafos aleatorios de Erdős-Renyi, notamos que la distribución $G(n, \frac{1}{2})$ era equivalente a elegir de forma uniforme entre todos los grafos definidos en el conjunto $\{1, \dots, n\}$. Sin embargo, es distribución es uniforme sobre grafos **etiquetados** y en muchos contextos estamos mas interesados en elegir sobre grafos no etiquetados, es decir, elegir uniformemente entre todas las clases de isomorfismo de grafos definidos en $\{1, \dots, n\}$.

Ejemplo 2.32. Hay 8 grafos etiquetados de 3 vértices, pero solo 4 grafos no etiquetados de 3 vértices.



Vamos a ver que con esa nueva distribución sigue valiendo la ley 0 – 1 de Fagin. La herramienta principal sera un resultado de carácter combinatorio, descubierto originalmente por Erdős y Renyi[10] (aunque las técnicas utilizadas fueron introducidas por Polya) que clasifica asintoticamente la cantidad de grafos no etiquetados de cierto tamaño. El desarrollo presentado a continuación es estándar [5][4].

Definición 2.33. Dado $n \in \mathbb{N}$, sea L_n el conjunto de todos los grafos etiquetados de n elementos y sea U_n el conjunto de todos los grafos no etiquetados de n elementos.

Definición 2.34. Dado $n \in \mathbb{N}$ y una propiedad P preservada por isomorfismos, denotaremos L_n^P al conjunto de todos los grafos etiquetados de n elementos que cumplan P y U_n^P al conjunto de grafos no etiquetados que cumplan P

Proposición 2.35. (Identidades de Pólya) Valen las siguientes igualdades que relacionan $|U_n|$ y $|L_n|$:

$$|U_n| = \sum_{G \in L_n} \frac{|Aut(G)|}{n!}$$

$$|L_n| = \sum_{G \in U_n} \frac{n!}{|Aut(G)|}$$

Demostración. U_n es un conjunto de clases de equivalencia de L_n . La clase de equivalencia de $G \in L_n$ tiene $\frac{n!}{|Aut(G)|}$ elementos, luego se sigue que $|L_n| = \sum_{G \in U_n} \frac{n!}{|Aut(G)|}$. Por otro lado, como una manera de contar las clases de equivalencia es sumar por cada elemento el inverso de cuantos elementos perteneces a su clase de equivalencia, tenemos que $|U_n| = \sum_{G \in L_n} \frac{|Aut(G)|}{n!}$ \square

Corolario 2.36. Vale la desigualdad $|U_n| > \frac{|L_n|}{n!}$, o equivalentemente, $|L_n| < |U_n| \cdot n!$

Demostración. $|U_n| = \sum_{G \in L_n} \frac{|Aut(G)|}{n!} < \sum_{G \in L_n} \frac{1}{n!} = \frac{|L_n|}{n!}$ \square

Definición 2.37. Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación. Denotamos con $Fix(\sigma)$ el conjunto de todos los grafos etiquetados G tal que $\sigma \in Aut(G)$.

Proposición 2.38. Vale la siguiente igualdad:

$$|U_n| = \sum_{\sigma \in S_n} \frac{|Fix(\sigma)|}{n!}$$

Demostración. Por la proposición anterior tenemos que:

$$|U_n| = \sum_{G \in L_n} \frac{|Aut(G)|}{n!} = \sum_{G \in L_n} \sum_{\sigma \in Aut(G)} \frac{1}{n!} = \sum_{\sigma \in S_n} \frac{|Fix(\sigma)|}{n!}$$

\square

Definición 2.39. Dada permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$, definimos permutación $\tilde{\sigma} : A \rightarrow A$, donde A es el conjunto de pares (no ordenados) de elementos distintos de $\{1, \dots, n\}$ y $\tilde{\sigma}$ esta dada por $\tilde{\sigma}((i, j)) = (\sigma(i), \sigma(j))$

Proposición 2.40. Para toda permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ vale que:

$$|Fix \sigma| = 2^{\# \text{ ciclos } \tilde{\sigma}}$$

Demostración. El conjunto A de pares no ordenados de elementos distintos de $\{1, \dots, n\}$ se puede ver como el conjunto de potenciales aristas de un grafo sobre $\{1, \dots, n\}$. Para que σ sea un automorfismo de G , para todo $a \in A$ se tiene que cumplir que:

$$a \in E \text{ sii } \tilde{\sigma}(a) \in E \text{ sii } \tilde{\sigma}(\tilde{\sigma}(a)) \in E$$

Luego, si σ es un automorfismo de G , para cada ciclo de $\tilde{\sigma}$ o bien todos los elementos son aristas de G o ninguno lo es. Se sigue que la cantidad de grafos etiquetados que tienen a σ como automorfismo es $2^{\# \text{ ciclos } \tilde{\sigma}}$. \square

Lema 2.41. Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación con k elementos no fijos. Luego, vale que:

$$\#\text{ciclos } \tilde{\sigma} \leq \binom{n}{2} - \frac{kn}{2} + \frac{k^2}{4} + \frac{k}{4}$$

Demostración. Vamos a acotar la cantidad de elementos fijos de $\tilde{\sigma}$. Notemos que si i no está fijo por σ para todo $j \neq i$ vale que (i, j) no está fijo por $\tilde{\sigma}$. La cantidad de pares de esa naturaleza es:

$$k \cdot (n-1) - \binom{k}{2} = k \cdot (n-1) - \frac{k \cdot (k-1)}{2} = kn - \frac{k^2}{2} - \frac{k}{2}$$

Ya que hay k elementos i no fijos por σ , $n-1$ elementos j distintos de i y el término $\binom{k}{2}$ está restando porque los pares (i, j) donde ninguno está fijo por σ están contados dos veces en la multiplicación anterior. Ahora, sabemos que hay al menos $kn - \frac{k^2}{2} - \frac{k}{2}$ elementos no fijos por $\tilde{\sigma}$, que deben estar en orbitas de al menos 2 elementos. Luego tenemos que:

$$\#\text{ciclos } \tilde{\sigma} \leq \binom{n}{2} - (kn - \frac{k^2}{2} - \frac{k}{2}) + \frac{kn - \frac{k^2}{2} - \frac{k}{2}}{2} = \binom{n}{2} - \frac{kn}{2} + \frac{k^2}{2} + \frac{k}{2}$$

□

Corolario 2.42. Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación con k puntos no fijos. Luego, vale que:

$$|\text{Fix } \sigma| \leq 2^{\binom{n}{2} - \frac{kn}{2} + \frac{k^2}{4} + \frac{k}{4}}$$

Demostración. Consecuencia inmediata de 2.40 y 2.41. □

Teorema 2.43. (Erdős-Renyi)[10] Vale que:

$$\lim_{n \rightarrow \infty} \frac{|U_n| \cdot n!}{|L_n|} = 1$$

Demostración. Vimos anteriormente que $|U_n| > \frac{|L_n|}{n!}$, luego se sigue que para todo n vale que $\frac{|U_n| \cdot n!}{|L_n|} > 1$ y por tanto $\lim_{n \rightarrow \infty} \frac{|U_n| \cdot n!}{|L_n|} \geq 1$. Para ver que es efectivamente uno, empezamos con la igualdad establecida en la proposición 2.38:

$$|U_n| \cdot n! \cdot 2^{-\binom{n}{2}} = \left(\sum_{\sigma \in S_n} \frac{|\text{Fix}(\sigma)|}{n!} \right) \cdot n! \cdot 2^{-\binom{n}{2}} = \left(\sum_{\sigma \in S_n} |\text{Fix}(\sigma)| \right) \cdot 2^{-\binom{n}{2}}$$

Como todo grafo es invariante por la permutación trivial id nos queda:

$$2^{-\binom{n}{2}} \cdot 2^{-\binom{n}{2}} + \left(\sum_{\sigma \in S_n - Id} |\text{Fix}(\sigma)| \right) \cdot 2^{-\binom{n}{2}} = 1 + \left(\sum_{\sigma \in S_n - Id} |\text{Fix}(\sigma)| \right) \cdot 2^{-\binom{n}{2}}$$

Veamos que este último término a la derecha tiende a 0, de lo que se sigue el enunciado. Denotemos con $s(\sigma)$ la cantidad de puntos no fijos de σ y observemos que para todo k , hay menos de n^k permutaciones con k puntos no fijos, ya que

para especificar la permutación basta especificar la imagen de esos k elementos. Luego, usando el lema 2.41 nos queda:

$$\begin{aligned} \left(\sum_{\sigma \in S_n - Id} |Fix(\sigma)| \right) \cdot 2^{-\binom{n}{2}} &= \left(\sum_{k=1}^n \sum_{s(\sigma)=k} |Fix(\sigma)| \right) < \sum_{k=1}^n n^k \cdot 2^{\binom{n}{2} - \frac{kn}{2} + \frac{k^2}{4} + \frac{k}{4}} \cdot 2^{-\binom{n}{2}} \\ &= \sum_{k=1}^n n^k \cdot 2^{-\frac{kn}{2} + \frac{k^2}{4} + \frac{k}{4}} = \sum_{k=1}^n 2^{-\frac{kn}{2} + \frac{k^2}{4} + \frac{k}{4} + k \cdot \log_2(n)} = \sum_{k=1}^n 2^{k \cdot \left(-\frac{n}{2} + \frac{k}{4} + \frac{1}{4} + \log_2(n)\right)} \\ &\leq \sum_{k=1}^n 2^{k \cdot \left(-\frac{n}{4} + \frac{1}{4} + \log_2(n)\right)} = \sum_{k=1}^n 2^{\left(-\frac{n}{4} + \frac{1}{4} + \log_2(n)\right)k} \end{aligned}$$

Por ultimo, como $2^{\left(-\frac{n}{4} + \frac{1}{4} + \log_2(n)\right)k}$ tiende a 0 cuando n tiende a infinito, se sigue que la sumatoria también tiende a 0 cuando n tiende a infinito, como queríamos. \square

Observación 2.44. La importancia del teorema anterior, y la razón por la que lo investigaba Polya, es que nos da una estimación asintótica de la cantidad de grafos no etiquetados, específicamente:

$$|U_n| = \frac{|L_n|}{n!} \cdot (1 + o(1)) = \frac{2^{\binom{n}{2}}}{n!} \cdot (1 + o(1))$$

Dicha aproximación es bastante burda, pero usando métodos similares se pueden obtener estimaciones mucho mas precisas. En [19] se demuestra que:

$$|U_n| = \frac{2^{\binom{n}{2}}}{n!} \cdot \left(1 + \frac{n^2}{2^{n-1}} - \frac{n}{2^{n-1}} + o\left(\frac{1}{2^n}\right)\right)$$

Definición 2.45. Denotamos con $G_{un}(n)$ la variable aleatoria con distribución uniforme sobre U_n , es decir, la variable aleatoria que toma valores en grafos no etiquetados de n vértices de manera uniforme.

Observación 2.46. Equivalentemente, podemos definir a $G_{un}(n)$ como la variable que toma valores en L_n de manera tal que $G \in L_n$ es elegido con peso $|Aut(G)|$.

Proposición 2.47. Cuando n tiende a infinito, casi seguro $G_{un}(n)$ es rígido, es decir, su único automorfismo es trivial.

Demostración. Denotemos con U_n^r el conjunto de los grafos etiquetados de n vértices rígidos y con U_n^c el conjunto de los grafos no etiquetados de n vértices no rígidos. Vamos a ver que cuando n tiende a infinito, la proporción $|U_n^r|/|U_n|$ tiende a 1. Por las fórmulas de Polya, tenemos:

$$|L_n| = \sum_{G \in U_n} \frac{n!}{|Aut(G)|} = \sum_{G \in U_n^r} n! + \sum_{G \in U_n^c} \frac{n!}{|Aut(G)|} \geq \sum_{G \in U_n^r} n! + \sum_{G \in U_n^c} \frac{n!}{2}$$

Dividiendo lados por $|U_n| \cdot n!$ nos queda:

$$\frac{|L_n|}{|U_n| \cdot n!} \geq \frac{|U_n^r|}{|U_n|} + \frac{|U_n^c|}{2 \cdot |U_n|} = \frac{1}{2} + \frac{|U_n^r|}{2 \cdot |U_n|}$$

Ahora, por el lema 2.41 sabemos que el termino de la izquierda tiende a 1 cuando n tiende a infinito, luego para que se cumpla la desigualdad anterior $|U_n^r|/|U_n|$ debe tender a 1. \square

Proposición 2.48. *Cuando n tiende a infinito, casi seguro $\mathbf{G}(n)$ es rígido, es decir, su único automorfismo es trivial.*

Demostración. Denotemos con L_n^r el conjunto de los grafos etiquetados de n vértices rígidos. Sabemos que $|L_n^r| < |U_n| \cdot n!$. Por la identidad de Polya tenemos que:

$$|L_n^r| = \sum_{G \in U_n^r} \frac{n!}{|Aut(G)|} = \sum_{G \in U_n^r} n! = |U_n^r| \cdot n!$$

Luego, como vimos en la proposición anterior que casi todos los grafos no etiquetados son rígidos:

$$\frac{|L_n^r|}{|L_n|} > \frac{|U_n^r| \cdot n!}{|U_n| \cdot n!} = \frac{|U_n^r|}{|U_n|} \rightarrow 1 \quad (2.1)$$

\square

Ahora finalmente podemos demostrar que la ley 0-1 vale para grafos no etiquetados:

Teorema 2.49. (Ley 0-1 de Fagin para grafos no etiquetados)[5] *Sea ϕ una oración en la lógica de primer orden de grafos. Luego, se cumple que:*

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}_{un}(n) \models \phi) = \lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi)$$

En particular, se sigue que:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}_{un}(n) \models \phi) \in \{0, 1\}$$

Demostración. Demostramos anteriormente que cuando n tiende a infinito, tanto $\mathbf{G}_{un}(n)$ como $\mathbf{G}(n)$ son casi seguro rígidos. Luego, basta demostrar que:

$$\lim_{n \rightarrow \infty} Pr(\mathbf{G}_{un}(n) \models \phi | \mathbf{G}_{un}(n) \text{ es rígido}) = \lim_{n \rightarrow \infty} Pr(\mathbf{G}(n) \models \phi | \mathbf{G}(n) \text{ es rígido})$$

Ahora, $\mathbf{G}(n)$ tiene una distribución uniforme sobre L_n mientras que $\mathbf{G}_{un}(n)$ tiene una distribución sobre L_n con peso proporcional a $|Aut(G)|$. Pero si nos restringimos al caso en que son rígidos tienen la misma distribución, ya que por definición todo grafo rígido cumple que $|Aut(G)| = 1$, luego se sigue que $Pr(\mathbf{G}_{un}(n) \models \phi | \mathbf{G}_{un}(n) \text{ es rígido}) = Pr(\mathbf{G}(n) \models \phi | \mathbf{G}(n) \text{ es rígido})$. \square

La demostración anterior muestra de hecho algo mas fuerte, que cuando n tiende a infinito $\mathbf{G}(n)$ y $\mathbf{G}_{un}(n)$ tienen una distribución idéntica. En particular, eso nos permite trasladar las cotas sobre la convergencia de la ley de Fagin que demostramos anteriormente al caso no etiquetado:

Teorema 2.50. (Cota inferior para la convergencia de la ley de Fagin, versión no etiquetada) Para todo k hay una oración ϕ_k en el lenguaje de los grafos con $qr(\phi_k) = k + 1$ tal que ϕ_k es válida en casi todos los grafos no etiquetados pero

$$\lim_{k \rightarrow \infty} Pr(\mathbf{G}_{un}(k \cdot 2^k) \models \phi_k) = 0$$

Teorema 2.51. (Cota superior para la convergencia de la ley de Fagin, versión no etiquetada) Para todo $\varepsilon > 0$, existe un $c > 0$ tal que para todo k y toda oración ϕ en la lógica de primer orden de grafos con $qr(\phi) = k$ vale que:

- Si ϕ es válida en casi todos los grafos, $Pr(\mathbf{G}_{un}(c \cdot k^2 \cdot 2^k) \models \phi) > 1 - \varepsilon$
- Si ϕ no es válida en casi ningún grafo, $Pr(\mathbf{G}_{un}(c \cdot k^2 \cdot 2^k) \models \phi) < \varepsilon$

2.2. Una generalización: Las estructuras paramétricas

2.2.1. Introducción

En la sección anterior presentamos la demostración de Fagin de la ley 0-1 para grafos. Notemos que dicho resultado no guarda en principio relación alguna con el teorema principal del paper, que mencionamos en la introducción y es una ley 0-1 para \mathcal{L} -estructuras elegidas de manera uniforme, ya que los grafos no son una clase de \mathcal{L} -estructuras: Si \mathcal{L} es el lenguaje con una relación binaria R , las \mathcal{L} -estructuras son grafos dirigidos y no simples. Esa observación lleva a Fagin a preguntar por una generalización de su teorema que englobe ambos casos.

Una generalización de esa naturaleza es demostrada por Oberschelp en [30]. Oberschelp introduce el concepto de clase paramétrica, que incluye a todas las clases de \mathcal{L} -estructuras, a los grafos y a una gran variedad de otras estructuras, y demuestra que todas cumplen un análogo de la ley 0-1 de Fagin, que llamaremos ley 0-1 de Oberschelp. Empecemos con algunas definiciones:

Definición 2.52. Sea \mathcal{L} un lenguaje relacional. Una fórmula ϕ sobre \mathcal{L} se dice **paramétrica** si es de la forma $\forall x_1 \dots \forall x_n \text{ distintos } \psi(x_1, \dots, x_n)$, donde ψ es una combinación booleana de formulas atómicas que usan **todas** las variables x_i .

Ejemplo 2.53. Sea \mathcal{L} el lenguaje relacional que consiste en una única relación binaria R . Consideremos las formulas:

- $\phi_1 = \forall x \neg(xRx) = \text{"La relación } R \text{ es antireflexiva"}$
- $\phi_2 = \forall x \forall y \text{ distintos } (xRy \leftrightarrow yRx) = \text{"La relación } R \text{ es simétrica"}$
- $\phi_3 = \forall x \forall y \text{ distintos } (xRy \leftrightarrow \neg yRx) = \text{"La relación } R \text{ es antisimétrica"}$
- $\phi_4 = \forall x \forall y \forall z \text{ distintos } ((xRy \wedge yRz) \rightarrow xRz) = \text{"La relación } R \text{ es transitiva"}$

Las formulas ϕ_1 , ϕ_2 y ϕ_3 son paramétricas, pero ϕ_4 no, pues contiene formulas atómicas que no usan todas las variables, como xRy .

Observación 2.54. En general, si \mathcal{L} tiene relaciones de aridad hasta k , toda fórmula paramétrica sobre \mathcal{L} tendrá como mucho k variables

Definición 2.55. Sea \mathcal{L} lenguaje relacional. Una clase \mathcal{C} de \mathcal{L} -estructuras se llama **paramétrica** si esta axiomatizada por un conjunto finito de formulas paramétricas.

Observación 2.56. Si \mathcal{L} es un lenguaje relacional, la clase de todas las \mathcal{L} -estructuras es trivialmente una clase paramétrica, ya que esta axiomatizada por el conjunto \emptyset .

Veamos una serie de ejemplos de clases paramétricas:

Ejemplo 2.57. Consideremos de vuelta el lenguaje \mathcal{L} de una única relación binaria R y las formulas ϕ_i del ejemplo anterior. La clase de **grafos dirigido** (axiomatizada por ϕ_1) es paramétrica, así como la de **grafos** (axiomatizada por ϕ_1 y ϕ_2) y la de **torneos** (axiomatizada por ϕ_1 y ϕ_3).

Como nota histórica, aunque hoy en día el estudio de los grafos aleatorios es mucho más importante que el de otras estructuras aleatorias, el estudio de los torneos aleatorios es de hecho más antiguo, ya que fueron introducidos por Ronald Fisher y Frank Yates en la década del 20' para formalizar ciertos problemas de estadística.

Ejemplo 2.58. Sea \mathcal{L} el lenguaje con una relación binaria R y relaciones unarias P_1, P_2, \dots, P_k . La clase de \mathcal{L} -estructuras axiomatizada por ϕ_1, ϕ_2 y $\forall x (\bigvee_i P_i(x) \wedge (\bigwedge_{i \neq j} P_i(x) \wedge P_j(x)))$ es paramétrica y se corresponde a la clase de **grafos con vértices coloreados**, en la que P_1, \dots, P_k representan los colores posibles. De manera análoga podemos definir la clase de grafos dirigidos con vértices coloreados y torneos con vértices coloreados, que también son paramétricas.

Ejemplo 2.59. Sea \mathcal{L} el lenguaje con relaciones R_1, \dots, R_k . La clase de \mathcal{L} -estructuras axiomatizada por $\forall x (\bigwedge \neg R_i(x, x))$ y $\forall x \forall y \bigwedge_{i \neq j} \neg (R_i(x, y) \wedge R_j(x, y))$, que se corresponde con la clase de grafos con aristas coloreadas, es paramétrica. De manera análoga podemos definir la clase de torneos con aristas coloreadas y la clase de grafos dirigidos con aristas coloreadas, que también son paramétricas.

Proposición 2.60. Sea \mathcal{L} lenguaje relacional con relaciones de aridad como mucho k y sea \mathcal{C} una clase paramétrica de \mathcal{L} -estructuras. Luego M pertenece a \mathcal{C} si y solo si toda subestructura $N \subset M$ de k elementos pertenece a \mathcal{C} .

Demostración. \implies) Los axiomas de una clase paramétrica son todas formulas universales, luego si $M \in \mathcal{C}$ y $N \subset M$ se sigue que $N \in \mathcal{C}$. En particular eso vale cuando N tiene k elementos.

\impliedby) Quiero ver que si toda subestructura de k elementos de M esta en \mathcal{C} se sigue que $M \in \mathcal{C}$. Sea ϕ una de las formulas que parametrizan \mathcal{C} , de la forma $\forall x_1 \dots \forall x_s$ distintos $\psi(x_1, \dots, x_s)$. $M \models \phi$ si y solo si para todo conjunto $m_1, \dots, m_s \in M$ vale que $M[x_1 = m_1, \dots, x_s = m_s] \models \psi(x_1, \dots, x_s)$. Pero eso es equivalente a que $N[x_1 = m_1, \dots, x_s = m_s] \models \psi(x_1, \dots, x_s)$ donde N es la subestructura $\{m_1, \dots, m_s\}$ y sabemos que eso vale porque por hipótesis $N \in \mathcal{C}$ \square

Corolario 2.61. *La clase de ordenes totales no es una clase paramétrica.*

Demostración. Consideremos la \mathcal{L} -estructura sobre $\{a, b, c\}$ dada por $R = \{(a, b), (b, c), (c, a)\}$. Aunque ciertamente no es un orden total, ya que R no es transitiva, toda subestructura de dos elementos si lo es, luego por la proposición anterior la clase de ordenes totales no es paramétrica. \square

Definición 2.62. Sea \mathcal{L} un lenguaje relacional con relaciones de aridad menor a k y sea \mathcal{C} una clase paramétrica de \mathcal{L} -estructuras. \mathcal{C} se dice no trivial si contiene estructuras de k elementos.

Ejemplo 2.63. Sea \mathcal{L} un lenguaje relacional. La clase de \mathcal{L} -estructuras axiomatizada por $\forall x \forall y \text{ distintos } x = y$ es paramétrica pero es trivial, ya que solo contiene a las estructuras de cardinalidad 1.

Vamos a introducir una nueva terminología para clasificar mejor el comportamiento de las clases paramétricas.

Definición 2.64. Sea \mathcal{L} un lenguaje relacional, M una \mathcal{L} -estructura y $\{x_1, \dots, x_k\} \subset M$. El k -esqueleto de $\{x_1, \dots, x_k\}$ queda determinado por que relaciones que involucran a todos los x_i cumple $\{x_1, \dots, x_k\}$.

Ejemplo 2.65. Sea \mathcal{L} el lenguaje relacional consistente en una única relación binaria R . Hay dos tipos posibles de 1-esqueleto para x : $\{R(x, x)\}$ y $\{\neg R(x, x)\}$. Hay 4-tipos posibles de 2-esqueleto para $\{x, y\}$: $\{R(x, y), R(y, x)\}$, $\{\neg R(x, y), R(y, x)\}$, $\{R(x, y), \neg R(y, x)\}$ y $\{\neg R(x, y), \neg R(y, x)\}$. Por ultimo, para todo $k \geq 3$ hay un solo tipo 3-esqueleto, porque \mathcal{L} no incluye relaciones que puedan involucrar a mas de dos variables.

Con esa nueva terminología, podemos dar una nueva definición de clase paramétrica. Una clase paramétrica sobre un lenguaje \mathcal{L} es una clase de \mathcal{L} -estructuras en la que solo pueden aparecer ciertos tipo de esqueleto.

Ejemplo 2.66. Sea \mathcal{L} el lenguaje relacional consistente en una única relación binaria R . La clase de grafos solo permite el 1-esqueleto $\{\neg R(x, x)\}$ y los 2-esqueletos $\{R(x, y), R(y, x)\}$ y $\{\neg R(x, y), \neg R(y, x)\}$. La clase de Torneos solo permite el 1-esqueleto $\{\neg R(x, x)\}$ y los 2-esqueletos $\{R(x, y), \neg R(y, x)\}$ y $\{\neg R(x, y), R(y, x)\}$

Proposición 2.67. *Sea \mathcal{L} un lenguaje relacional, \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras. Luego podemos generar una \mathcal{L} -estructura sobre $\{1, \dots, n\}$ perteneciente de la siguiente manera aleatoria: para todo subconjunto $S \subset \{1, \dots, n\}$ de tamaño menor a k elegimos un esqueleto de forma uniforme entre todos los permitidos en \mathcal{C} . Luego, la estructura resultante pertenece a \mathcal{C} y más aun, la distribución resultante es uniforme entre todas las formas de elegir una \mathcal{L} -estructura en \mathcal{C} sobre $\{1, \dots, n\}$.*

Demostración. Es evidente que la estructura resultante esta en \mathcal{C} , ya que todo k -esqueleto es por definición compatible con los axiomas de \mathcal{C} y como toda estructura puede ser generada de una sola forma, se sigue que la distribución resultante es uniforme. \square

Aunque por su generalidad el procedimiento anterior para generar una \mathcal{L} -estructura en \mathcal{C} de manera aleatoria parece muy abstracto, en todos los casos que nos interesan es equivalente a la forma mas natural. Por ejemplo:

Ejemplo 2.68. Para generar un torneo aleatorio de forma uniforme sobre $\{1, \dots, n\}$ basta con, para cada par $i > j$, elegir con probabilidad $\frac{1}{2}$ y de forma independiente si $i \sim j$ y $j \not\sim i$ o si $i \not\sim j$ y $j \sim i$

Ejemplo 2.69. Para generar un grafo con vértices de k colores de forma uniforme sobre $\{1, \dots, n\}$ basta con, para cada par $i > j$, elegir con probabilidad $\frac{1}{2}$ y de forma independiente si $i \sim j$ y $j \sim i$ o si $i \not\sim j$ y $j \not\sim i$, y además para cada i elegir de manera uniforme e independiente cual de los k colores tendrá.

Definición 2.70. Sea \mathcal{L} un lenguaje relacional y sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} estructuras. Denotamos con $S_{\mathcal{C}}(n)$ la variable aleatoria con distribución uniforme entre las \mathcal{L} -estructuras en \mathcal{C} sobre $\{1, \dots, n\}$.

2.2.2. La ley de Oberschelp

Ahora podemos anunciar el siguiente teorema, descubierto por Walter Oberschelp en 1982, que generaliza la ley 0-1 de Fagin a toda clase paramétrica de \mathcal{L} -estructuras:

Teorema 2.71. (Ley 0-1 de Oberschelp)[30][5] Sea \mathcal{L} lenguaje relacional, \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras y ϕ una oración en lenguaje \mathcal{L} . Luego, o bien ϕ es válida en casi todas las estructuras de \mathcal{C} o no es válida en casi ninguna estructura de \mathcal{C} .

Definición 2.72. Sea \mathcal{L} un lenguaje relacional, sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras con relaciones de aridad menor a r y sea M una \mathcal{L} -estructura en \mathcal{C} . Decimos que M tiene la propiedad de la r -**extensión** si se cumple lo siguiente: Para todo subconjunto $A \subset M$ de tamaño k y para toda elección, dado subconjunto $S \subset A$ de tamaño s , de un tipo de $(s+1)$ -esqueleto en \mathcal{C} , existe un z tal que para todo S se cumple que $S \cup \{z\}$ tiene esqueleto τ_S .

Aunque la definición anterior es bastante complicada en general, para los casos que nos interesan adopta formas bastantes simples:

Ejemplo 2.73. Un torneo G cumple la propiedad de la k -extensión si para todo conjunto $U \subset G$ de k elementos y partición de U en conjuntos disjuntos U, V existe un z tal que $x \sim z$ para todo $x \in U$ y $z \not\sim x$ para todo $x \in V$

Ejemplo 2.74. Un grafo dirigido G cumple la propiedad de k -extensión si para todo conjunto $U \subset G$ de k elementos y toda partición de U en subconjuntos disjuntos U_1, U_2, U_3, U_4 existe un $z \in G$ tal que dado $x \in U$ vale que $z \sim x$ si y solo si $x \in U_1 \cup U_2$ y vale que $x \not\sim z$ si y solo $x \in U_3 \cup U_4$.

Ejemplo 2.75. Un hipergrafo 3-regular (H, R) cumple la propiedad de la r -extensión si para todo conjunto U de r elementos y todo conjunto V de pares no ordenados de elementos de U existe un $z \in U$ tal que para todo par $x, y \in V$ vale que $R(x, y, z)$ si y solo si $(x, y) \in V$

Proposición 2.76. Sea \mathcal{L} un lenguaje relacional y \mathcal{C} una clase paramétrica de \mathcal{L} -estructuras. Luego la propiedad de k extensión es expresable por una oración ϕ_k en lenguaje \mathcal{L} con rango de cuantificador $k + 1$.

Demostración. Es expresable por una oración de la forma $\phi_r = \forall x_1 \forall x_2 (\dots) \forall x_k (\bigwedge_i \exists z : \psi_i(x_1, \dots, x_k, z))$ donde las ψ_i expresan las distintas extensiones posibles, es decir, las distintas elecciones de esqueleto para todo subconjunto $S \subset \{x_1, \dots, x_k\}$. \square

Proposición 2.77. Sea \mathcal{L} un lenguaje relacional, $k > 0$ y \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras. Luego casi toda estructura en \mathcal{C} tiene la propiedad de k -extensión.

Demostración. Sea $k > 0$. Supongamos que el lenguaje \mathcal{L} tiene relaciones de aridad hasta r y que la clase \mathcal{C} admite C_1 1-esqueletos distintos, C_2 2-esqueletos distintos, \dots , C_r r -esqueletos distintos.

Sea $M \sim \mathcal{S}_{\mathcal{C}}(n)$, una estructura de n elementos en \mathcal{C} elegida de manera uniforme. Fijemos $m_1, \dots, m_k \in M$ y fijemos, para cada $S \subset \{m_1, \dots, m_k\}$ de tamaño menor a $r - 1$ una elección τ_S de esqueleto en \mathcal{C} .

Dado un subconjunto $S \subset \{m_1, \dots, m_k\}$ de tamaño t , la probabilidad de que un $x \in M - \{m_1, \dots, m_k\}$ tenga un tipo de $(t + 1)$ -esqueleto dado es C_{t+1}^{-1} . Como dicha probabilidad es independiente para todos los subconjuntos S , se sigue que la probabilidad de que la probabilidad de que un $x \in M - \{m_1, \dots, m_k\}$ dado cumpla la extensión dada por los τ_S es:

$$C_1 \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}}$$

Mas aun, como la elección es independiente para distintos $x \in M - \{m_1, \dots, m_k\}$, la probabilidad de que **ninguno** cumpla la extensión dada por τ_S es:

$$(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}})^{n-k}$$

Ahora, nosotros queremos acotar la probabilidad de que para **todo** subconjunto S y para **toda** elección de extensión dada por los τ_S existe un $z \in M - \{m_1, \dots, m_r\}$ que la realice. Como hay $\binom{n}{k}$ subconjuntos de tamaño k y $C_1^{\binom{k}{1}} \cdot C_2^{\binom{k}{2}} \cdot (\dots) \cdot C_k^{\binom{k}{k}}$ elecciones de extensión, por la cota de la union la probabilidad de que todas ellas tengan extensión es menor a :

$$\binom{n}{k} \cdot (C_1^{\binom{k}{1}} \cdot C_2^{\binom{k}{2}} \cdot (\dots) \cdot C_r^{\binom{k}{r}}) \cdot (1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r}})^{n-k}$$

Ahora, notemos que el termino de la derecha crece como un polinomio en función de n , mientras que el termino de la izquierda decae exponencialmente en función de n . Se sigue que cuando n tiende a infinito el producto anterior tiende a 0, es decir, la probabilidad de que haya un subconjunto de tamaño k sin todas las extensiones tiende a 0. \square

Proposición 2.78. Sea \mathcal{L} lenguaje relacional, \mathcal{C} una clase de \mathcal{L} -estructuras paramétricas. Sean M y N dos \mathcal{L} -estructuras en \mathcal{C} con la propiedad de k -extensión. Luego, el Duplicator tiene una estrategia ganadora en el juego con m fichas infinito entre M y N . En particular, tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïssé de m turnos entre M y N .

Demostración. análoga a la proposición 2.12: En cada turno $i < k$ la propiedad de la k -extensión permite al Duplicator hallar una respuesta que preserve el isomorfismo parcial y por tanto ganar el juego de k turnos. \square

Proposición 2.79. *Sea \mathcal{L} un lenguaje relacional, \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras axiomatizada por ϕ y sea θ_k la oración que expresa la propiedad de k -extensión para \mathcal{C} . Luego para toda fórmula ψ en $\mathcal{L}_{w,\infty}^k$ o bien $\phi \wedge \theta_k$ implica ψ o bien $\phi \wedge \theta_k$ implica $\neg\psi$.*

Demostración. Si ninguna de las dos implicaciones fuera válida, existirían \mathcal{L} -estructuras M y N pertenecientes a \mathcal{C} con la propiedad de k -extensión tal que $M \models \psi$ y $N \not\models \psi$. Se sigue que el Spoiler tiene una estrategia ganadora en el juego con k fichas entre M y N , pero eso contradice la proposición anterior, absurdo. \square

Corolario 2.80. *Sea \mathcal{L} un lenguaje relacional, \mathcal{C} una clase paramétrica de \mathcal{L} -estructuras axiomatizada por ϕ_1, \dots, ϕ_n y θ_k las fórmulas de primer orden que expresan las propiedades de k -extensión para \mathcal{C} . Luego $\{\phi_1, \dots, \phi_n, \theta_1, \theta_2, \dots\}$ es una teoría completa.*

Ahora podemos demostrar una versión infinitaria de la ley 0-1 de Oberschelp:

Teorema 2.81. *(Ley 0-1 de Oberschelp, versión infinitaria)[5] Sea \mathcal{L} lenguaje relacional, \mathcal{C} una clase paramétrica de \mathcal{L} -estructuras y ϕ una fórmula en $\mathcal{L}_{\infty,w}^k$ sobre lenguaje L . Luego, o bien ϕ es válida en casi todas las estructuras de \mathcal{C} o no es válida en casi ninguna estructura de \mathcal{C} .*

Demostración. Sea $\mathcal{S}_{\mathcal{C}}(n)$ una \mathcal{C} -estructura de n elementos elegida de forma uniforme, sea $k = qr(\phi)$ y sea θ_k la oración que expresa el axioma de k -extensión. Por la proposición anterior, o bien θ_k implica a ϕ o bien θ_k implica a $\neg\phi$.

En el primer caso, $Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi) \geq Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi \wedge \theta_k) = Pr(\mathcal{S}_{\mathcal{C}}(n) \models \theta_k)$ y como sabemos que $\lim_{n \rightarrow \infty} Pr(\mathcal{S}_{\mathcal{C}}(n) \models \theta_k) = 1$, se sigue que $\lim_{n \rightarrow \infty} Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi) = 1$. En el segundo caso, $Pr(\mathcal{S}_{\mathcal{C}}(n) \models \neg\phi) \geq Pr(\mathcal{S}_{\mathcal{C}}(n) \models \neg\phi \wedge \theta_k) = Pr(\mathcal{S}_{\mathcal{C}}(n) \models \theta_k)$ y como sabemos que $\lim_{n \rightarrow \infty} Pr(\mathcal{S}_{\mathcal{C}}(n) \models \theta_k) = 1$, se sigue que $\lim_{n \rightarrow \infty} Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi) = 0$ \square

2.2.3. Cotas para la convergencia de la ley de Oberschelp

En esta sección generalizaremos las cotas anteriores sobre la velocidad de convergencia de la ley de Fagin a resultados sobre la velocidad de convergencia de la ley de Oberschelp, usando ideas muy similares. Empezaremos con una cota inferior

Teorema 2.82. (Cota inferior para la convergencia de la ley de Oberschelp) *Sea \mathcal{L} un lenguaje relacional con relaciones de aridad menor a r y sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} estructuras, con C_1 tipos de 1-esqueleto, C_2 tipos de 2-esqueleto, ..., C_r tipos de r -esqueleto.*

Luego, para todo k existe una oración ϕ_k en lenguaje \mathcal{L} tal que es válida en casi todas las estructuras de \mathcal{C} , pero si $n = c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}$ cuando k tiende a infinito la probabilidad de que una estructura en \mathcal{C} de n elementos tienda a 0.

Demostración. Sea ϕ_k la oración que postula que todo conjunto de k elementos tiene una extensión dada. Vimos anteriormente que dicha oración tiene rango de cuantificador $k+1$. Vamos a partir el conjunto $\{1, \dots, n\}$ en $\frac{n}{k}$ conjuntos disjuntos de tamaño k . Notemos que como los conjuntos son disjuntos, la probabilidad de que tengan una extensión es independiente entre si, luego si denotamos con p la probabilidad de que $\mathcal{S}_{\mathcal{C}} \models \phi$, nos queda:

$$p \leq Pr\left(\bigcap A_i \text{ tiene extensión}\right) = \prod_i Pr(A_i \text{ tiene extensión})$$

Vimos en la demostración de 2.77 que la probabilidad de dado un conjunto de k elementos no tenga una extensión dada es igual a $(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}}) \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}}$. Luego, nos queda:

$$p \leq (1 - (1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}}) \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}})^{\frac{n}{k}}$$

Usando la desigualdad $(1 - x) \leq e^{-x}$ nos queda:

$$p \leq e^{-(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}}) \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}} \cdot \frac{n}{k}}$$

Concentrémonos ahora en el exponente. Veamos que si $n = c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}$, cuando k tiende a infinito el termino en el exponente tiende a $-\infty$ y por tanto p tiende a 0. Reemplazando y usando de nuevo la desigualdad anterior nos queda:

$$\begin{aligned} & -(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}}) \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}} \cdot [c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}]^{-k} \cdot [c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}] \\ & \leq -e^{-c \cdot k^{r-1}} \cdot [c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}] \end{aligned}$$

Analicemos la ultima expresión. El único término que tiende a 0 es $e^{-c \cdot k^{r-1}}$, pero si elegimos un c suficientemente pequeño (específicamente $c < \frac{\ln(C_r)}{(r-1)!}$) queda claro que $\lim_{k \rightarrow \infty} e^{-c \cdot k^{r-1}} \cdot C_r^{\binom{k}{r-1}} = +\infty$. Como $C_r^{\binom{k}{r-1}}$ es el termino de mas rápido crecimiento entre todos los dos

luego $\lim_{k \rightarrow \infty} -e^{-c \cdot k^{r-1}} \cdot [c \cdot k^{r-1} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}] = -\infty$, como queríamos. \square

Ahora, la cota superior:

Teorema 2.83. (Cota superior para la convergencia de la ley de Oberschelp)

Sea \mathcal{L} un lenguaje relacional con relaciones de aridad menor a r y sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} estructuras, con C_1 tipos de 1-esqueleto, C_2 tipos de 2-esqueleto, ..., C_r tipos de r -esqueleto. Luego, para todo $\varepsilon > 0$, existe un $c > 0$ tal que si

$$n = c \cdot k^r \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}$$

Se cumple que:

- Si ϕ es valida en casi todas las estructuras de \mathcal{C} y $k = qr(\phi)$, vale que $Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi) > 1 - \varepsilon$

- Si ϕ no es válida en casi ninguna estructura de \mathcal{C} y $k = qr(\phi)$, vale que $Pr(\mathcal{S}_{\mathcal{C}}(n) \models \phi) < \varepsilon$

Demostración. Vamos a ver que cuando k tiende a infinito, la probabilidad de que $\mathcal{S}_{\mathcal{C}}(n)$ cumpla la propiedad de k -extensión tiende a 1. Como sabemos que si ϕ es válida en casi todas las estructuras de \mathcal{C} vale que $\theta_k \Rightarrow \phi$ y si ϕ no es válida en casi ninguna estructura de \mathcal{C} , se sigue el enunciado.

En la demostración de la proposición 2.77, demostramos que la probabilidad de que $S(n)$ no cumpliera el axioma de k -elección esta acotada por:

$$\begin{aligned} & \binom{n}{k} \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}} \cdot (1 - C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}})^{n-k} \\ & < n^k \cdot C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}} \cdot e^{-C_1 \cdot C_2^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}} \cdot (n-k)} \end{aligned}$$

Tomemos $n = c \cdot k^r \cdot C_1^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}$. Nos queda:

$$\begin{aligned} & = [c \cdot k^r \cdot C_1^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}}]^k \cdot C_1^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}} \cdot e^{-c \cdot k^r} \\ & = c^k \cdot k^{rk} \cdot C_1^{\binom{k}{1} \cdot k} \cdot (\dots) \cdot C_r^{\binom{k}{r-1} \cdot k} \cdot C_1^{\binom{k}{1}} \cdot (\dots) \cdot C_r^{\binom{k}{r-1}} \cdot e^{-c \cdot k^r} \end{aligned}$$

Notemos que todos los términos salvo el $e^{-c \cdot k^r}$ son crecientes, pero el único término que tiene un crecimiento del orden es de $e^{-c \cdot k^r}$ es $C_r^{\binom{k}{r-1} \cdot k} \sim C_r^{k^r / (r-1)!} = e^{\frac{\ln(C_r)}{(r-1)!} \cdot k^r}$. Luego, se sigue que si $c \geq \frac{\ln(C_r)}{(r-1)!}$, cuando k tiende a infinito la probabilidad de que $\mathcal{S}_{\mathcal{C}}(n)$ no cumpla la propiedad de la k -extensión tiende a 0. \square

A modo de ejemplo, enunciaremos ambos resultados para grafos dirigidos. Denotemos con $\mathbf{G}_d(n)$ la variable aleatoria con distribución uniforme entre todos los grafos dirigidos de n vértices. Luego, notando que la clase de grafos dirigidos tiene 1 tipo 1-esqueleto y 4 tipos de 2-esqueleto, y usando los dos teoremas anteriores nos queda que:

Teorema 2.84. (Cota inferior para la convergencia de la ley de Oberschelp para grafos dirigidos) Para todo k hay una oración ϕ_k en el lenguaje de los grafos dirigidos con $qr(\phi_k) = k + 1$ y longitud $O(k)$ tal que ϕ_k es válida en casi todos los grafos dirigidos pero

$$\lim_{k \rightarrow \infty} Pr(\mathbf{G}_d(k \cdot 4^k) \models \phi_k) = 0$$

Teorema 2.85. (Cota superior para la convergencia de la ley de Oberschelp en grafos dirigidos) Para todo $\varepsilon > 0$, existe un $c > 0$ tal que para toda oración ϕ en el lenguaje de los grafos dirigidos, si $k = qr(\phi)$ se cumple que:

- Si ϕ es válida en casi todos los grafos dirigidos, $Pr(\mathbf{G}_d(c \cdot k^2 \cdot 4^k) \models \phi) > 1 - \varepsilon$
- Si ϕ no es válida en casi ninguna grafo dirigido, $Pr(\mathbf{G}_d(c \cdot k^2 \cdot 4^k) \models \phi) < \varepsilon$

2.3. Estructuras universales

2.3.1. El grafo de Rado

Hasta ahora trabajamos con grafos aleatorios de Erdős-Renyi definidos sobre conjuntos finitos, pero la misma definicion se puede usar para generar una distribución de probabilidad de grafos sobre cualquier conjunto. Estudiaremos ahora las propiedades de un grafo aleatorio sobre un conjunto contable, que denotaremos $G(\omega, p)$, y como se relaciona con las propiedades asintóticas de grafos finitos.

Proposición 2.86. *Sea G un grafo contable elegido con la distribución de Erdős-Renyi $G(\omega, p)$. Luego casi seguramente G cumple la propiedad de k -extensión para todo k .*

Demostración. Vamos a ver que para todo par de conjuntos finitos disjuntos $U, V \subset G$ con probabilidad 1 existe un $z \in G$ tal que z esta conectado a todo elemento de U y ningún elemento de V . Como hay una cantidad contable de subconjuntos finitos de G y la intersección de contables eventos de probabilidad 1 tiene probabilidad 1 se sigue que con probabilidad 1 G cumple todos los axiomas de elección.

Sea $k > 0$ y sean x_1, \dots, x_k y y_1, \dots, y_k dos conjuntos disjuntos de vértices. Dado $z \notin \{x_1, \dots, x_k, y_1, \dots, y_k\}$ la probabilidad de que z este conectado a todos los x_i y a ningún y_j es $(p \cdot (1-p))^k$. Como dicho es evento es independiente para todo par de vértices diferente z_1 y z_2 , la probabilidad de que no haya ningún z que cumpla la propiedad de extensión deseada es $\lim_{n \rightarrow \infty} (1 - (p \cdot (1-p))^k)^n = 0$, ya que $1 - (p \cdot (1-p))^k < 1$. \square

Teorema 2.87. *Sean G y H dos grafos contables que para todo k cumplen la propiedad de k -extensión. Luego G Y H son isomorfos.*

Demostración. La demostración es mediante el método **back and forth**, introducido por Cantor para demostrar la categoricidad de los ordenes lineales densos. Vamos a construir el isomorfismo de forma inductiva, como el limite de una serie de isomorfismos parciales. Escribamos los elementos de G como $\{g_1, g_2, \dots\}$ y los de H como $\{h_1, h_2, \dots\}$. Empezamos estableciendo $f(g_1) = h_1$. Para el paso inductivo, supongamos que tenemos un isomorfismo parcial entre G y H dado por $f : U \subset G \rightarrow V \subset H$. Sea g_n el menor elemento de G no perteneciente a U . Sea U_+ el conjunto de los elementos de U conectados a g_n y U_- el conjunto de elementos de U no conectados a g_n . Defino V_+ como $f(U_+)$ y V_- como $f(U_-)$. Como H tiene todas las propiedades de extensión, debe haber un $h \in H$ conectado a todo elementado de V_+ y a ningún elemento de V_- . Luego podemos extender el isomorfismo parcial $f : U \rightarrow V$ al isomorfismo parcial $\tilde{f} : U \cup \{g_n\} \rightarrow V \cup \{h\}$ estableciendo $\tilde{f}(g_n) = h$.

La misma construcción se puede hacer en el sentido inverso: Si h_n es el menor elemento de H no contenido en V se puede hallar un $g \in G$ tal que el isomorfismo parcial $f : U \rightarrow V$ se extiende a $\tilde{f} : U \cup \{g\} \rightarrow V \cup \{h_n\}$ estableciendo $\tilde{f}(g) = h_n$. Alternando entre ambas construcciones (de donde viene el nombre *back and forth*), el isomorfismo parcial $f : \{h_1\} \rightarrow \{g_1\}$ se puede extender a un isomorfismo $f : G \rightarrow H$. \square

Observación 2.88. A primera vista parece que la demostración anterior, de construir el isomorfismo como limite de isomorfismos parciales, se puede generalizar a pares de grafos G y H de cardinal infinito κ con todas las propiedades

de extensión, pero como las propiedades de extensión solo nos permiten extender conjuntos finitos, el argumento no funciona. De hecho, el teorema falla de manera espectacular para cardinales mas grandes: Para todo cardinal infinito no contable κ , hay 2^κ grafos de cardinal κ no isomorfos que cumplen todas las propiedades de extensión. Esto se relaciona a temas muy profundos de teoría de modelos, específicamente, a que la teoría de grafos con todas las propiedades de extensión es **no estable**. [20]

Definición 2.89. Al único grafo contable hasta isomorfismo con todas las propiedades de extensión se lo denomina el **grafo de Rado** y se lo denota \mathcal{R} .

Corolario 2.90. Sea G un grafo elegido con la distribución $\mathbf{G}(\omega, p)$. Luego, casi seguro $G \cong \mathcal{R}$.

No es difícil dar construcciones explícitas del grafo de Rado. Presentamos tres a modo de ejemplo, la primera fue la construcción original de Rado y Erdős, la segunda es estándar en la literatura y la tercera es original.

Proposición 2.91. Consideremos los siguientes grafos definidos sobre \mathbb{N}

- G_1 es el grafo definido por la siguiente relación: dados $n, m \in \mathbb{N}$ tal que $n > m$, $n \sim m$ si y solo si el m -ésimo dígito en la representación binaria de n es 1.
- G_2 es el grafo definido por la siguiente relación: dados $n, m \in \mathbb{N}$, $n \sim m$ si y solo si $U_n \in U_m$ o $U_m \in U_n$, donde U_i es una enumeración de los conjuntos hereditariamente finitos.
- G_3 es el grafo definido por la siguiente relación: dados $n, m \in \mathbb{N}$ tal que $n > m$, $n \sim m$ si y solo si el $f_n(m) = 1$, donde f_i es una enumeración de las funciones computables.

Los tres grafos son isomorfos al grafo de Rado.

Demostración. Ambos grafos son contables, luego para ver que son isomorfos a \mathcal{R} basta ver que cumplen todas las propiedades de extensión.

- Empecemos con G_1 . Sean U y V conjuntos disjuntos de naturales. Sea c un numero mayor a todos los elementos de $U \cup V$. Consideremos ahora $n = 2^c + \sum_{u \in U} 2^u$. Es claro que n es mayor a todo u_i y v_i , luego por definición $n \sim u_i$ para todo i pero $n \not\sim v_j$ para ningún j .
- Para G_2 , sean U y V conjuntos disjuntos de naturales. Consideremos el conjunto $\{A_i : i \in U\}$. Es hereditariamente finito, luego es igual a A_n para algun $n \in \mathbb{N}$. Contiene a todo A_i para $i \in U$ y a ningún A_i para $i \in V$, además, por el axioma de fundación, no esta contenido en A_i para ningún $i \in U$ o $i \in V$. Luego, se sigue por la definición de G_2 que n esta conectado a todo elemento de U y a ninguno de V .
- Para G_3 , sean U y V conjuntos disjuntos de naturales. Como U es finito, sabemos que hay una funcion computable f tal que $f(x) = 1 \iff x \in U$. Es un resultado básico de teoría de la computabilidad que hay infinitos indices n tal que $f_n = f$. En particular, debe haber un n mayor a todo elemento de $U \cup V$ tal que $f_n = f$. Se sigue que ese n esta conectado a todo elemento de U y a ninguno de V .

□

Observación 2.92. Se puede ver explícitamente que las dos primeras construcciones anteriores son isomorfas por la llamada **codificación de Ackermann**, una biyección $f : HFS \rightarrow \mathbb{N}$ definida de forma recursiva como $f(\emptyset) = 0$ y $f(\{x_1, \dots, x_n\}) = \sum_{i=1}^n 2^{f(x_i)}$.

Veamos otras propiedades interesantes del grafo de Rado:

Proposición 2.93. Para todo par de subconjuntos disjuntos U y V de \mathcal{R} existen infinitos elementos de \mathcal{R} que están conectados a todo elemento de U y a ninguno de V .

Demostración. Como \mathcal{R} tiene todas las propiedades de extensión, sabemos que existe un elemento z_1 que esta conectado a todo elemento de U y a ningún elemento de V . Por otro lado, como $U \cup \{z_1\}$ y V son disjuntos debe existir un z_2 conectado a todo elemento de $U \cup \{z_1\}$ y a ninguno de V . Repitiendo el proceso llegamos aun conjunto z_1, z_2, \dots de vértices conectados a todo elemento de U y a ningún elemento de V . □

Proposición 2.94. Todo grafo H de cardinal numerable es isomorfo a un subgrafo del grafo de Rado \mathcal{R} .

Demostración. Sea $H = \{h_1, h_2, \dots\}$. Vamos a construir una inclusión $i : H \rightarrow \mathcal{R}$ como limite de inclusiones parciales. Empecemos estableciendo definiendo $i_1 : \{h_1\} \rightarrow \mathcal{R}$ de forma arbitrara. Para el paso inductivo, supongamos que tenemos definida una inclusión de $i_k : \{h_1, \dots, h_k\} \rightarrow \mathcal{R}$. Sea U el subconjunto de $\{h_1, \dots, h_k\}$ formado por los vértices de que están conectados a h_{k+1} y V el subconjunto de $\{h_1, \dots, h_k\}$ que no están conectados a h_{k+1} . Como $i_k(U)$ y $i_k(V)$ son conjuntos disjuntos y \mathcal{R} cumple todas las propiedades de extensión, existe $x \in \mathcal{R}$ tal que x esta conectado a todo elemento de $i_k(U)$ y a ninguno de $i_k(V)$. Luego podemos extender a i_k a $i_{k+1} : \{h_1, \dots, h_{k+1}\} \rightarrow \mathcal{R}$ estableciendo $i_{k+1}(h_{k+1}) = x$. □

Proposición 2.95. El grafo de Rado es autocomplementario, es decir, isomorfo a su complemento.

Demostración. Basta ver que el complemento de \mathcal{R} tiene todas las propiedades de extensión. Sean $U, V \subset \mathcal{R}$. Existe un $x \in \mathcal{R}$ conectado a todos los elementos de V y ninguno de U , luego existe un $x \in \mathcal{R}^c$ conectado a todos los elementos de U y ninguno de V . □

Definición 2.96. Un grafo G se dice **regular por partición** si para todo par de subconjuntos disjuntos U_1, U_2 tal que $G = U_1 \cup U_2$ o bien $G \cong U_1$ o bien $G \cong U_2$.

Observación 2.97. Si G es regular por partición se cumple que para toda colección de subconjuntos disjuntos U_1, \dots, U_k tal que $G = \bigcup U_i$ existe un i tal que G es isomorfo a U_i .

Proposición 2.98. El grafo \mathcal{R} es regular por partición y mas aun, es el único grafo contable regular por partición ademas del grafo vacío y el grafo completo.

Demostración. Empecemos viendo que \mathcal{R} es regular por partición. Sean U_1, U_2 disjuntos tal que $\mathcal{R} = U_1 \cup U_2$. Supongamos que ningún U_i es isomorfo al grafo de Rado. Se sigue que ningún U_i cumple todas las propiedades de extensión, es decir, para todo i existen $A_i, B_i \subset U_i$ disjuntos tal que no existe ningún $z \in U_i$ que este conectado a todo elemento de A_i y a ninguno de B_i . Consideremos los conjuntos $A_1 \cup A_2$ y $B_1 \cup B_2$. Son disjuntos y como \mathcal{R} tiene todas las propiedades de extensión, se sigue que debe haber un $z \in \mathcal{R}$ tal que z esta conectado con todo elemento de $A_1 \cup A_2$ y con ninguno de $B_1 \cup B_2$. Por otro lado, como $\mathcal{R} = U_1 \cup U_2$, debe existir un i tal que $z \in U_i$, pero entonces hay un elemento de U_i conectado a todo elemento de A_i y a ninguno de B_i , lo que es absurdo por definición de A_i y B_i . Luego, al menos un U_i debe ser isomorfo a \mathcal{R}

Demostremos ahora que todo grafo regular por partición que no sea ni el vacío ni el completo debe cumplir todas las propiedades de extensión, y por tanto, los únicos grafos contables regulares por partición son el grafo vacío, el grafo completo y el grafo de Rado. Sea G un grafo regular por partición ni vacío ni completo. G no puede tener vértices aislados: Si los tuviera sea U el conjunto de vértices aislados y $V = G - U$. $G = U \cup V$, pero G no es isomorfo a U (ya que U es vacío) ni a V (ya que V no tiene vértices aislados) lo que contradice la regularidad por partición de G . Por un argumento dual, tampoco puede tener vértices conectados a todos los otros vértices.

Supongamos que G es un grafo ni vacío ni completo con todas las propiedades de extensión y que G no es isomorfo a \mathcal{R} . Luego, deben existir un par de subconjuntos finitos disjuntos $U, V \subset G$ tal que (U, V) no tiene extensión en G de **tamaño minimal con dicha propiedad**. Además, como demostramos anteriormente que G no puede tener vértices aislados ni conectados a todos los vecinos, se sigue que $|U \cup V| = |U| + |V| > 1$. Luego existen subconjuntos disjuntos y no vacíos $A, B \subset U \cup V$. Definamos los subconjuntos X, Y de G de la siguiente manera:

$$X = A \cup \{x : x \text{ no realiza extensión de } (U \cap A, V \cap A), x \notin B\}$$

$$Y = A \cup \{x : x \text{ no realiza extensión de } (U \cap B, V \cap B) x \notin X\}$$

Empecemos notando que por definición, X e Y son disjuntos y mas aun, $G = X \cup Y$, ya que si hubiera un $x \notin X \cup Y$ se seguiría que x realiza extensión de (U, V) , lo que contradice nuestras hipótesis. Como G es regular por partición, se sigue que $G \cong X$ o $G \cong Y$. Ahora, en X el par $(U \cap A, V \cap A)$ no tiene extensión, pero es de tamaño menor a (U, V) y en G todo par de tamaño menor a G tiene extensión, luego G no puede ser isomorfo a X . De manera análoga, G no puede isomorfo a Y porque en Y el par $(U \cap B, V \cap B)$ no tiene extensión. Esto contradice la hipótesis de que G es regular por partición, luego no puede existir tal G . \square

Definición 2.99. Decimos que un grafo G es **robusto frente a cambios finitos** si cualquier grafo G' que se obtiene agregando o borrando finitas aristas es isomorfo a G .

Proposición 2.100. *El grafo de Rado \mathcal{R} es robusto frente a cambios finitos.*

Demostración. Sea G un grafo resultante de remover o agregar finitas aristas a \mathcal{R} . Veamos que G tiene todas las propiedades de extensión y por tanto, es

isomorfo a \mathcal{R} . Sean U y V subconjuntos finitos y disjuntos de vértices. Sabemos por la proposición 2.93 que hay infinitos $z \in \mathcal{R}$ que están conectados a todo elemento de U y ninguno de V . Como G solo difiere de \mathcal{R} en finitas aristas, debe haber uno de esos z no afectado por los cambios, es decir, tal que en G este conectado a todo elemento de U y a ningún elemento de V . \square

Teorema 2.101. (Ultrahomogeneidad del grafo de Rado) Sean M y N dos subconjuntos de \mathcal{R} y $f : M \rightarrow N$ un isomorfismo. Luego existe un automorfismo $\tilde{f} : \mathcal{R} \rightarrow \mathcal{R}$ tal que $\tilde{f}|_M = f$. Es decir, todo isomorfismo local se extiende a un isomorfismo global.

Demostración. Por el método back and forth, de manera casi idéntica a la demostración del teorema 2.87. Vamos a construir a \tilde{f} como el límite de una serie de isomorfismos parciales $f_i : U_i \subset \mathcal{R} \rightarrow V_i \subset \mathcal{R}$, empezando con $f_0 = f$ y en cada paso usando las propiedades de extensión de \mathcal{R} para extender $f_i : U_i \rightarrow V_i$ a $f_{i+1} : U_i \cup \{a\} \rightarrow V_i \cup \{b\}$ o $f_{i+1} : U_i \rightarrow V_i \cup \{b\}$. \square

Aunque el grafo de Rado es un objeto de mucho interés en si mismo, el siguiente teorema muestra que tiene una íntima conexión con las propiedades asintóticas de grafos finitos.

Teorema 2.102. Sea ϕ una fórmula en la lógica de primer orden de grafos. Luego ϕ se cumple en casi todos los grafos si y solo si $\mathcal{R} \models \phi$.

Demostración. Sea ϕ una fórmula con $qr(\phi) = k$ y sea θ_k la fórmula que expresa la propiedad de k -extensión. Vimos que ϕ se cumple en casi todos los grafos si y solo si θ_k implica ϕ . Por otro lado, sabemos que $\mathcal{R} \models \theta_k$, luego se sigue que ϕ se cumple en casi todos los grafos si y solo si $\mathcal{R} \models \phi$. \square

Esto nos permite demostrar algunos resultados interesantes, como el siguiente:

Proposición 2.103. Sea $\phi(x)$ una fórmula en la lógica de primer orden de grafos. Luego son equivalentes:

- $\exists x \phi(x)$ se cumple en casi todos los grafos
- $\forall x \phi(x)$ se cumple en casi todos los grafos

Demostración. Por el teorema anterior, es equivalente ver que para toda fórmula $\phi(x)$ se cumple que $\mathcal{R} \models \forall x \phi(x)$ si y solo si $\mathcal{R} \models \exists x : \phi(x)$. Veamos que $\mathcal{R} \models \exists x \phi(x)$ implica $\mathcal{R} \models \forall x \phi(x)$, la otra implicación es obvia.

Por definición, si $\mathcal{R} \models \exists x : \phi(x)$ debe existir un $a \in \mathcal{R}$ tal que $\mathcal{R} \models \phi(a)$. Sea $b \in \mathcal{R}$. Por la ultrahomogeneidad del grafo de Rado sabemos que hay un isomorfismo $f : \mathcal{R} \rightarrow \mathcal{R}$ tal que $f(a) = b$. Como la validez de fórmulas de primer orden se preserva por isomorfismos se sigue que $\mathcal{R} \models \phi(b)$ y como lo demostramos para cualquier $b \in \mathcal{R}$ se sigue que $\mathcal{R} \models \forall x : \phi(x)$ \square

Pero la consecuencia mas importante es el siguiente teorema, que muestra que hay un algoritmo para determinar que oraciones tienen probabilidad limite 1, y a diferencia del algoritmo estocástico propuesto anteriormente, podemos establecer una cota explicita de su complejidad computacional.

Teorema 2.104. Hay un algoritmo en PSPACE que determina, dada oración ϕ en la lógica de primer orden de grafos, si ϕ se cumple en casi todos los grafos.

Demostración. Daremos un algoritmo para determinar, dada oración ϕ en la lógica de primer orden de grafos si vale que $\mathcal{R} \models \phi$, lo que por el teorema anterior determina si ϕ se cumple en casi todos los grafos.

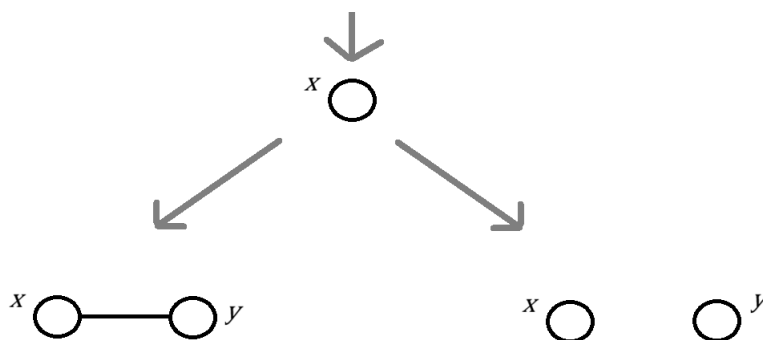
Para empezar, notemos que podemos convertir en tiempo polinómico toda fórmula ϕ en una fórmula equivalente en fórmula normal prenex, es decir, de la forma $\tilde{\phi} = Q_1 x_1 \dots Q_n x_n \psi(x_1, \dots, x_n)$ donde cada Q_i es \exists o \forall y ψ es libre de cuantificadores. Luego basta dar un algoritmo para ese tipo de formulas en particular.

El algoritmo que daremos funcionara por **eliminación de cuantificadores**. Supongamos que queremos determinar si $\mathcal{R}[x_1 = a_1, \dots, x_k = a_k] \models \exists x \psi$. Por definición, esa fórmula es cierta si existe algún $a \in \mathcal{R}$ tal que $\mathcal{R}[x_1 = a_1, \dots, x_k = a_k, x = a] \models \psi$. A priori, sería necesario probar con cada $a \in \mathcal{R}$, tarea imposible porque la infinitud de \mathcal{R} , pero gracias a la propiedad de ultrahomogeneidad del grafo de Rado podemos reducirlo a un caso finito.

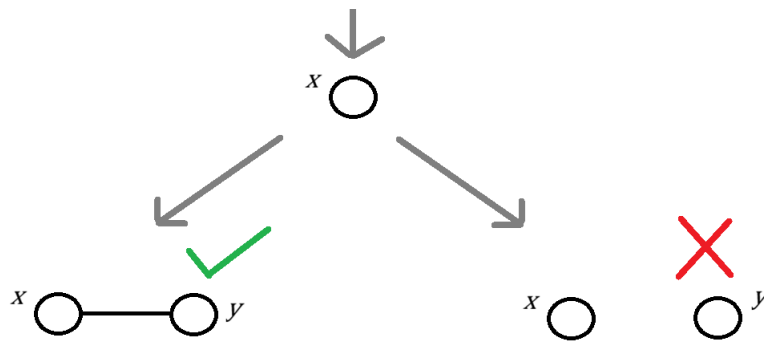
Supongamos que $f : \mathcal{R} \rightarrow \mathcal{R}$ es un automorfismo que preserva los a_i y tal que $f(a) = b$. Como los isomorfismos preservan la verdad de las formulas de primer orden se sigue que $\mathcal{R}[x_1 = a_1, \dots, x_k = a_k, x = a] \models \psi$ si y solo si $\mathcal{R}[x_1 = a_1, \dots, x_k = a_k, x = b] \models \psi$. Por la ultrahomogeneidad del grafo de Rado, existe un automorfismo de tales características si y solo si $\{a_1, \dots, a_k, a\}$ y $\{a_1, \dots, a_k, b\}$ son isomorfos, que es equivalente a que el conjunto de a_i conectados a a sea el mismo que de a_i conectados a b . Luego, no hace falta probar con todo $a \in \mathcal{R}$ sino solo con los de la forma a_U con $U \subset \{1, \dots, k\}$ definido por $a_U \sim a_i$ si y solo si $i \in U$.

Aplicando el procedimiento anterior de forma recursiva llegamos al problema de evaluar $\mathcal{R} \models \psi$ donde ψ no tiene cuantificadores, que se puede hacer directamente. En cada momento de la recursión el algoritmo solo debe guardar la elección de a_i , que ocupa un espacio en la memoria de orden k^2 donde k es la cantidad de variables. Luego, el algoritmo esta en PSPACE \square \square

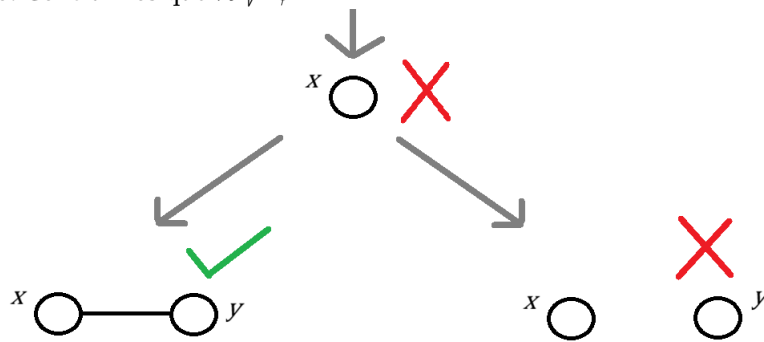
Ejemplo 2.105. Sea $\phi = \exists x \forall y : x \sim y$ y usemos el algoritmo anterior para determinar si $\mathcal{R} \models \phi$. Notemos que ϕ ya esta en forma normal prenex. Empezamos considerando un árbol que incluya todos los grafos sobre el conjunto $\{x, y\}$, que son solo dos:



Para cada uno de esos grafos comprobamos si cumplen $x \sim y$:



Ahora subimos un nivel. Como el cuantificador que acompañaba a y es un \forall , y en al menos un caso no se cumple la oración, se sigue en el nivel de arriba tampoco. Concluimos que $\mathcal{R} \not\models \phi$



Aunque el algoritmo anterior requiere solo de espacio polinómico, esta muy lejos de ser eficiente, ya que al necesitar explorar todas las elecciones posibles de a_i requiere $O(2^{k^2})$ operaciones. ¿Se podrá hallar un algoritmo que logre lo mismo en tiempo polinómico? A día de la fecha no tenemos la respuesta, pero la siguiente proposición muestra que de ser el caso se seguiría que $P = PSPACE$, proposición que casi unánimemente se cree falsa.

Proposición 2.106. *El problema de determinar si una oración ϕ dada en la lógica de primer orden de grafos se cumple en casi todos los grafos es PSPACE-hard.*

Demostración. Vamos a ver que se puede reducir el problema de decidir si una fórmula booleana cuantificada es cierta (TQBF), que es bien sabido que es PSPACE-hard[1], a decidir si una oración ϕ es valida en casi todos los grafos.

Recordemos que una fórmula booleana cuantificada es de la forma $\phi = Q_1 p_1 \dots Q_n p_n \phi(p_1, \dots, p_n)$ donde las p_i son variables proposicionales y cada Q_i es \exists o \forall . Definamos la fórmula $\tilde{\phi}$ de la siguiente formula: por cada p_i definimos dos variables x_i y y_i cuantificadas de igual manera que p_i y reemplazamos cada ocurrencia de la proposición p_i en ψ por $x_i \sim y_i$. Por ejemplo, $\exists p \forall q (p \vee q)$ se convierte en $\exists x_1 \exists y_1 \forall x_2 \forall y_2 ((x_1 \sim y_1) \vee (x_2 \sim y_2))$. Es fácil ver que para todo grafo no vacío ni completo G vale que ϕ es cierta si y solo si $G \models \tilde{\phi}$, y como casi todos los grafos no son vacíos ni completos se sigue que ϕ es cierta si y solo si $\tilde{\phi}$ es valida en casi todos los grafos.

□

Corolario 2.107. *El problema de determinar una oración ϕ en la lógica de primer orden de grafos se cumple en casi todos los grafos es PSPACE-completo.*

2.3.2. El caso general

En esta sección generalizaremos los resultados anteriores a cualquier clase paramétrica \mathcal{C} de \mathcal{L} -estructuras. Empezaremos definiendo un análogo de el grafo de Rado para cualquier clase paramétrica \mathcal{C} , que llamaremos la \mathcal{C} -estructura universal, y demostraremos que cumple propiedades análogas y terminaremos dando un resultado que clasifica la complejidad algorítmica de decidir si una oración es válida en casi todas las estructuras de una clase paramétrica \mathcal{C} .

Definición 2.108. Sea \mathcal{L} un lenguaje relacional, \mathcal{C} una clase paramétrica de \mathcal{L} -estructura. Denotamos $\mathcal{S}_{\mathcal{C}}(\omega)$ a la \mathcal{L} -estructura aleatoria sobre $\{1, 2, \dots\}$ generada según la proposición 2.67.

Proposición 2.109. *Casi seguro que $\mathcal{S}_{\mathcal{C}}(\omega)$ cumple todas las propiedades de k -extensión de estructuras en \mathcal{C} .*

Demostración. Vamos a ver que para todo subconjunto finito $A \subset \{1, \dots\}$ y toda extensión τ , la probabilidad de que existe un z tal que z realice la extensión τ tiene probabilidad 1. Como la intersección de contables eventos de probabilidad 1 tiene probabilidad 1, se sigue que la probabilidad de que todo subconjunto finito A tenga todas sus extensiones es U .

Supongamos que la clase \mathcal{C} tiene C_1 tipos de 1-esqueleto, C_2 tipos de 2-esqueleto, ..., C_r tipos de r -esqueleto. Vimos en la proposición 2.77 que la probabilidad de que un $z \notin A$ realice una cierta extensión es $C_1^{-1} \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}}$ y que dicho evento es independiente para distintos z . Luego, la probabilidad de que ningún z realice dicha extensión es menor a $(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}})^n$ para todo n . Como $(1 - C_1^{-1} \cdot C_2^{-\binom{k}{1}} \cdot C_3^{-\binom{k}{2}} \cdot (\dots) \cdot C_r^{-\binom{k}{r-1}}) < 1$, se sigue que la probabilidad de que ningún z realice dicha extensión es igual a 1. □

Proposición 2.110. *Sean M y N dos \mathcal{L} -estructuras en \mathcal{C} que cumplen todas las propiedades de extensión. Luego, M y N son isomorfas.*

Demostración. Por el método back and forth, de forma análoga a 2.87. Sea $M = \{m_1, m_2, \dots\}$ y sea $N = \{n_1, \dots, n_2\}$. Vamos a construir un isomorfismo $f : M \rightarrow N$ como el límite de una serie de isomorfismos parciales $f_i : U \subset M \rightarrow V \subset N$. Empezamos con $f_1(m_1) = n_k$, donde n_k es un elemento con el mismo 1-esqueleto que m_1 , que debe existir porque N cumple todas las propiedades de extensión. Para el caso inductivo, sea $f_i : U \rightarrow V$ y sea m el menor elemento de M no perteneciente a U . Para cada subconjunto $S \subset U$, sea τ_S el tipo de esqueleto de $S \cup \{m\}$. Como N tiene todas las propiedades de extensión, existe un $n \in N$ tal que para todo $S \subset U$, $f_i(S) \cup \{n\}$ tiene esqueleto de tipo τ_S . Se sigue que $f_{i+1} : U \cup \{m\} \rightarrow V \cup \{n\}$ es un isomorfismo que extiende a f_i .

Por un argumento dual, si $n \in N$ es el menor elemento que no pertenece a V , podemos hallar un $m \in M$ tal que $f_{i+1} : U \cup \{m\} \rightarrow V \cup \{n\}$ es un isomorfismo que extiende a f_i . Alternando entre estos dos pasos nos da la construcción deseada. □

Definición 2.111. Se sigue que hay una única \mathcal{L} -estructura en \mathcal{C} contable que cumpla todas las propiedades de \mathcal{C} -extensión hasta isomorfismo. Se la denomina la **\mathcal{C} -estructura universal** y se la denota $U_{\mathcal{C}}$.

Corolario 2.112. *Casi seguro que $S_{\mathcal{C}}(\omega)$ es isomorfo a $U_{\mathcal{C}}$.*

Al igual que con el grafo de Rado, no es difícil dar construcciones explícitas de \mathcal{C} -estructuras universales. A modo de ejemplo, presentamos una variación de la construcción dada en 2.91 para grafos dirigidos y torneos.

Proposición 2.113. *Consideremos los siguientes grafos dirigidos definidos sobre \mathbb{N}*

- G_1 es el grafo dirigido definido por la siguiente relación: dados $n, m \in \mathbb{N}$ tal que $n > m$, $n \sim m$ si y solo si el m -ésimo dígito en la representación cuaternaria de n es 1 o 2, y $m \sim n$ si y solo si el m -ésimo dígito en la representación cuaternaria de n es 1 o 3.
- G_2 es el grafo dirigido definido por la siguiente relación: dados $n, m \in \mathbb{N}$ tal que $n > m$, $n \sim m$ si y solo si $f_n(m) \geq 0$ y $m \sim n$ si y solo si $f_n(m) \leq 0$ donde f_i es una enumeración de las funciones computables.

Ambos grafos son isomorfos al grafo dirigido universal.

Demostración. Basta con verificar que ambos grafos cumplen todas las propiedades de extensión para grafos dirigidos, que explicitamos en el ejemplo 2.74.

1. Sean $U_1, U_2, U_3, U_4 \subset \mathbb{N}$ cuatro subconjuntos finitos disjuntos. Consideremos el número n definido como $n = \sum_{i \in U_1} 4^i + \sum_{i \in U_2} 2 \cdot 4^i + \sum_{i \in U_3} 3 \cdot 4^i$. Es fácil ver que n es mayor a todo elemento de los subconjuntos anteriores, y por definición vale $n \sim i$ si y solo si $i \in U_1 \cup U_2$ y $i \sim n$ si y solo si $i \in U_1 \cup U_3$.
2. Sean $U_1, U_2, U_3, U_4 \subset \mathbb{N}$ cuatro subconjuntos finitos disjuntos. Como son finitos, hay una función computable tal que $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que:

$$f(i) = \begin{cases} 0 & \text{si } i \in U_1 \\ 1 & \text{si } i \in U_2 \\ -1 & \text{si } i \in U_3 \\ \uparrow & \text{si } i \notin U_1 \cup U_2 \cup U_3 \end{cases}$$

Además, como sabemos que para toda función computable hay infinitos índices que la realizan, existe un n tal que $f_n = f$ y tal que n es mayor a todo elemento de $U = U_1 \cup U_2 \cup U_3 \cup U_4$. Para ese n , para todo $i \in U$ se cumple por definición que $n \sim i$ si y solo si $i \in U_1 \cup U_2$ y que $i \sim n$ si y solo si $i \in U_1 \cup U_3$.

□

El siguiente teorema muestra que, de manera análoga al grafo de Rado, la estructura universal $U_{\mathcal{C}}$ cumple solo las oraciones que son válidas en casi todas las estructuras de \mathcal{C} .

Teorema 2.114. *Sea ϕ una oración en la lógica de primer orden de \mathcal{L} . Luego, ϕ es válida en casi todas las estructuras de \mathcal{C} si y solo si $U_{\mathcal{C}} \models \phi$.*

Demostración. Sea $k = qr(\phi)$ y sea θ_k la oración que expresa la propiedad de k -extensión para \mathcal{L} -estructuras. Sabemos que ϕ es válida en casi todas las estructuras de \mathcal{C} si y solo si θ_k implica ϕ y que $U_{\mathcal{C}}$ cumple θ_k . Luego, se sigue que ϕ es válida en casi todas las estructuras de \mathcal{C} si y solo si $U_{\mathcal{C}} \models \phi$. \square

El siguiente teorema generaliza el teorema 2.102.

Teorema 2.115. *Sea \mathcal{L} un lenguaje relacional y \mathcal{C} una clase paramétrica no trivial con un solo tipo de r -esqueleto para todo $r \leq k$. Sea $\phi(x_1, \dots, x_r)$ una fórmula el lenguaje \mathcal{L} . Luego, son equivalentes:*

- $\exists x_1 (\dots) \exists x_r \phi(x_1, \dots, x_r)$ es válida en casi todas las estructuras de \mathcal{C}
- $\forall x_1 (\dots) \forall x_r \phi(x_1, \dots, x_r)$ es válida en casi todas las estructuras de \mathcal{C}

Demostración. La única dirección que necesita demostración es (i) \Rightarrow (ii). Supongamos que $U_{\mathcal{C}} \models \phi$, es decir, existen $a_1, \dots, a_r \in U_{\mathcal{C}}$ tal que $U_{\mathcal{C}} \models \psi(a_1, \dots, a_r)$. Sea $b_1, \dots, b_r \in \mathcal{U}$. Como existe un único s -esqueleto en \mathcal{C} para todo $s \leq r$, se sigue que hay un isomorfismo $f : \{a_1, \dots, a_r\} \rightarrow \{b_1, \dots, b_r\}$ y como $U_{\mathcal{C}}$ es ultrahomogeneo, esto se extiende a un isomorfismo $\hat{f} : U_{\mathcal{C}} \rightarrow U_{\mathcal{C}}$ y como los isomorfismos preservan la validez de fórmulas, se sigue que $U_{\mathcal{C}} \models \phi(b_1, \dots, b_r)$. Como esto vale para cualquier tupla b , se sigue que $U_{\mathcal{C}} \models \forall x_1 (\dots) \forall x_r \phi(x_1, \dots, x_r)$ y por tanto $\forall x_1 (\dots) \forall x_r \phi(x_1, \dots, x_r)$ es válida en casi todas las estructuras de \mathcal{C} . \square

Por ultimo, nos queda clasificar la complejidad algorítmica de decidir que oraciones se cumplen en casi todas estructuras de \mathcal{C} .

Teorema 2.116. *Sea \mathcal{L} un lenguaje relacional y sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras. Existe un algoritmo en PSPACE que determina, dada oración ϕ en la lógica de primer orden en lenguaje \mathcal{L} , si ϕ se cumple en casi todas las estructuras de \mathcal{C} .*

Demostración. Análogo a 2.104. Vamos a dar un algoritmo que comprueba si, dada oración ϕ en forma normal prenex, $U_{\mathcal{C}} \models \phi$ por eliminación de cuantificadores. Como $U_{\mathcal{C}}$ es ultrahomogeneo, para verificar si $U_{\mathcal{C}} \models \exists x : \phi(a_1, \dots, a_n, x)$ basta corroborar, para cada tipo posible de extensión τ de $\{a_1, \dots, a_n\}$, si $U_{\mathcal{C}} \models \phi(a_1, \dots, a_n, a)$, donde a_{τ} es un elemento que realiza la extensión τ . Como en cada paso hay solo finitos tipos de extensión posibles, el algoritmo termina. Mas aun, en cada paso el algoritmo solo necesita guardar las relaciones en el conjunto $\{a_1, \dots, a_n\}$ y si el lenguaje \mathcal{L} tiene relaciones de aridad hasta k eso requiere $O(n^k)$ espacio. \square

Teorema 2.117. *Sea \mathcal{L} un lenguaje relacional y sea \mathcal{C} una clase paramétrica no trivial de \mathcal{L} -estructuras. El problema de determinar, dada oración ϕ en la lógica de primer orden en lenguaje \mathcal{L} , si ϕ se cumple en casi todas las estructuras de \mathcal{C} es PSPACE-hard.*

Demostración. Vamos a mostrar que el problema de decidir si una oración booleana cuantificada es cierta (TQBF), que es PSPACE-hard, es reducible al problema de decidir si una oración en lógica de primer orden solo con igualdad (que en particular, es una oración en el lenguaje \mathcal{L}) se cumple en casi todas las estructuras de \mathcal{C} . Sea $\phi = Q_1 p_1 (\dots) Q_n p_n \psi(p_1, \dots, p_n)$ una oración booleana cuantificada. Vamos a construir una oración $\hat{\phi}$ en lenguaje \mathcal{L} de la siguiente

manera: $\bar{\phi} = \forall x Q_1 x_1 (\dots) Q_n x_n \psi(x = x_1, \dots, x_n)$, donde Q_i son los mismos cuantificadores que en ϕ .

En cualquier estructura de mas de dos elementos, donde las formulas átomicas $x_i = y$ pueden adoptar ambos valores booleanos, $\bar{\phi}$ es valida si y solo si ϕ es valida, en particular, ϕ es valida si y solo si $U_C \models \hat{\phi}$ y sabemos que eso pasa si y solo si ϕ se cumple en casi todas las estructuras de C . \square

Juntando ambos resultados:

Teorema 2.118. *Sea \mathcal{L} un lenguaje relacional y sea C una clase paramétrica no trivial de \mathcal{L} -estructuras. El problema de determinar, dada oración ϕ en la lógica de primer orden en lenguaje \mathcal{L} , si ϕ se cumple en casi todas las estructuras de C es PSPACE-completo.*

Capítulo 3

Ley de convergencia para grafos muy sparse

3.1. Introducción

En el capítulo anterior, estudiamos las propiedades asintóticas de los grafos aleatorios $G(n, p)$, es decir, la probabilidad de que $G(n, p)$ tenga una propiedad dada cuando n tiende a infinito. Pero eso no es más que arañar la superficie de la teoría de grafos aleatorios, ya que los problemas más interesantes consisten en estudiar las propiedades asintóticas de $G(n, p(n))$ donde $p(n)$ es una función de los naturales en $(0, 1)$ tal que $\lim_{n \rightarrow \infty} p(n) = 0$, los llamados **grafos aleatorios sparse**.

A diferencia de los grafos aleatorios $G(n, p)$, que mostramos en el capítulo anterior que se comportan de modo muy similar para todo $p \in (0, 1)$, los grafos aleatorios sparse $G(n, p(n))$ tienen un comportamiento extremadamente rico, que depende mucho de la función $p(n)$. Dos ejemplos de ello son los siguientes teoremas clásicos de Erdős y Renyi:

Teorema 3.1. (Threshold de conexión para grafos aleatorios sparse)[8]

- Si $p(n) = (1 + O(1)) \cdot \frac{\ln(n)}{n}$, la probabilidad de que $G(n, p(n))$ sea conexo tiende a 1 cuando n tiende infinito.
- Si $p(n) = (1 - O(1)) \cdot \frac{\ln(n)}{n}$, la probabilidad de que $G(n, p(n))$ sea conexo tiende a 0 cuando n tiende infinito.

Teorema 3.2. (Aparición del componente gigante en grafos aleatorios sparse)[9]

Sea $c > 0$ y $p(n) = \frac{c}{n}$. Luego:

- Si $c < 1$, entonces cuando n tiende a infinito casi seguro todo componente conexo $G(p(n), n)$ es de tamaño $O(\ln(n))$.
- Si $c > 1$ entonces cuando n tiende a infinito casi seguro $G(p(n), n)$ tiene un componente conexo de tamaño $O(n)$.

Naturalmente, eso lleva a la pregunta de si es posible generalizar la ley 0 – 1 de Fagin para grafos aleatorios de la forma $G(n, p(n))$. La respuesta es afirmativa y fue hallada por Shelah y Compton, que investigaron las propiedades de grafos aleatorios de la forma $p(n) = c \cdot n^{-\alpha}$ para $\alpha \in (0, 1)$.

Teorema 3.3. (Ley 0-1 de Shelah-Spencer) [35] Sea $c > 0$, $\alpha \in (0, 1)$ irracional. Luego, para toda oración ϕ en la lógica de primer orden de grafos, vale que $\lim_{n \rightarrow \infty} \Pr(\mathbf{G}(n, c \cdot n^{-\alpha}) \models \phi) \in \{0, 1\}$ y mas aun, dicho límite no depende de c .

La demostración del teorema anterior, que Spencer simplificaría en [37], es similar a la de la ley 0-1 de Fagin, y consiste en ver que cuando n tiende a infinito, si $G_1, G_2 \sim G(n, c \cdot n^{-\alpha})$ casi seguro que el Duplicator tiene una estrategia ganadora en el juego de Ehrenfeucht-Fraïsse de k turnos entre G_1 y G_2 . A diferencia de la demostración de la ley de Fagin, sin embargo, la demostración es larga y requiere una atención minuciosa a los detalles. Como vimos anteriormente, los grafos $G(n, p)$ cumplen asintóticamente los axiomas de extensión, que garantizan siempre al Duplicator una respuesta adecuada, pero los grafos G tienen una estructura mucho mas complicada y la estrategia del Duplicator tendrá que adelantarse a todas las jugadas posibles del Spoiler.

Teorema 3.4. (Converso de la ley 0-1 de Shelah-Spencer)[35] Sea $c > 0$, $\alpha \in (0, 1)$ racional. Luego, existe una oración ϕ en la lógica de primer orden de grafos tal que el límite $\lim_{n \rightarrow \infty} \Pr(\mathbf{G}(n, c \cdot n^{-\alpha}) \models \phi)$ no existe.

Sin embargo, hay un caso muy importante que no estaba cubierto por en análisis de Shelah y Spencer, que es el de $\alpha = 1$, es decir, de grafos aleatorios de la forma $G(n, \frac{c}{n})$, los llamados **grafos aleatorios muy sparse**. Era bien sabido que ese tipo de grafos no cumplía una ley 0-1 para la lógica de primer orden, pero Lynch mostró que a diferencia de los grafos con α racional al menos cumplen una ley de convergencia:

Teorema 3.5. (Lynch)[29] Sea $c > 0$ y sea ϕ una oración en la lógica de primer orden de grafos. Luego, el límite $\lim_{n \rightarrow \infty} \Pr(\mathbf{G}(n, \frac{c}{n}) \models \phi)$ existe y mas aun, se puede escribir a partir del conjunto $\{1, c\}$ usando las operaciones $\{+, -, \cdot, e^x\}$

La demostración de Lynch se divide en dos partes, una combinatoria y una lógica. La combinatoria es para demostrar las propiedades locales del grafo $G(n, \frac{c}{n})$, que Lynch logra usando versiones generalizadas del principio de exclusión-inclusión, y en la parte lógica demuestra que esas propiedades locales caracterizan el grafo hasta equivalencia, que logra usando los juegos de Ehrenfeucht-Fraïsse. Nosotros seguiremos una estrategia muy similar pero con herramientas distintas, lo que dará lugar a una demostración mas corta. Para caracterizar las propiedades locales de $G(n, \frac{c}{n})$ usaremos el llamado método de los momentos y para demostrar la parte lógica usaremos como herramienta principal el teorema de Gaifman, demostrado en el capítulo 1.

3.2. Preliminares probabilísticos

3.2.1. El método de los momentos, ó, Cómo aprendí a dejar de preocuparme y a amar la esperanza

Un problema frecuente en probabilidad es que calcular exactamente la distribución de una variable aleatoria puede ser extremadamente difícil, cuando no imposible. Por otro lado, calcular la esperanza y, mas en general, los momentos de una variable aleatoria suele ser una tarea mas fácil, debido a las propiedades de linealidad. Por esa razón, son de gran utilidad los resultados que nos permiten extraer cierta información sobre la distribución a partir de los momentos. Las desigualdades de Markov y Chebyshev, que acotan las colas de una

distribución a partir de su primer y segundo momento respectivamente, son los ejemplos mas celebres de resultado del estilo. En esta sección estudiaremos cuando conocer todos los momentos de una variable nos permite determinarla inequívocamente.

Definición 3.6. Sea X una variable aleatoria con medida de probabilidad μ_X . Su función característica $\phi_X : \mathbb{R} \rightarrow \mathbb{C}$ esta definida como:

$$\phi_X(t) = \mathbb{E}(e^{itX}) = \int e^{itx} d\mu_X(x)$$

Proposición 3.7. Sea X una variable aleatoria con los primeros k momentos finitos. Luego ϕ_X es k veces diferenciable y $(\phi_X)^{(k)}(0) = i^k \cdot \mathbb{E}(X^k)$.

Teorema 3.8. [3] Si X e Y son dos variables aleatorias tal que $\phi_X = \phi_Y$, se sigue que X e Y tienen la misma distribución.

Definición 3.9. Sea X una variable aleatoria con todos los momentos finitos. Se dice que X esta **determinada por sus momentos** si para toda variable Y tal que para todo k vale que $\mathbb{E}(X^k) = \mathbb{E}(Y^k)$ se sigue X e Y tienen la misma distribución.

No toda variable aleatoria esta determinada por sus momentos. El contraejemplo mas conocido es el siguiente, descubierto por Stieljes en 1894.

Proposición 3.10. Sea $X \sim \mathcal{N}(0, 1)$ y sea $Y = e^X$, es decir, una variable con la distribución log-normal. Luego Y **no** esta determinada por sus momentos.

Demostración. La densidad de probabilidad de Y esta dada por $p_Y(x) = \frac{1}{x \cdot \sqrt{2\pi}} \cdot e^{-\frac{\ln(x)^2}{2}}$. Veamos que para todo $k \geq 0$ vale que

$$\int_{-\infty}^{\infty} x^k \cdot p_Y(x) \cdot \sin(2\pi \cdot \ln(x)) dx = 0$$

De lo que se sigue que la función de densidad dada por

$$p_Z(x) = p_Y(x) \cdot (\sin(2\pi \cdot \ln(x)) + 1)$$

define una variable aleatoria Z con los mismos momentos que Y , aunque claramente tienen distinta distribución. Empecemos haciendo el cambio de variables $u = \ln(x)$:

$$\int_{-\infty}^{\infty} x^k \cdot \frac{e^{-\frac{\ln(x)^2}{2}}}{x \cdot \sqrt{2\pi}} \cdot \sin(2\pi \cdot \ln(x)) dx = \frac{1}{\sqrt{2\pi}} \cdot \int_0^{\infty} e^{ku} \cdot e^{-\frac{u^2}{2}} \cdot \sin(2\pi \cdot u) du$$

Ahora haciendo el cambio de variable $v = u - k$ y recordando que $\sin(x - 2k\pi) = \sin(x)$ nos queda:

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(k-u)^2}{2}} \cdot \sin(2\pi \cdot u) du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{v^2}{2}} \cdot \sin(2\pi \cdot v) dv$$

Y como la ultima función es impar es claro que la integral debe ser 0. \square

Sin embargo, la mayoría de las distribuciones con las que trabajamos normalmente si están determinadas por sus momentos y se han descubiertos varios criterios para determinar cuando una variable aleatoria esta determinada por sus momentos. El mas conocido de ellos es el llamado "criterio de Carleson". Nosotros demostraremos una versión mas débil, que denominaremos el criterio de Riesz. Expresa la idea de que toda distribución cuyos momentos no crezcan "demasiado rápido" esta determinada por sus momentos.

Teorema 3.11. (Criterio de Riesz)[3] Sea X una variable aleatoria positiva con medida de probabilidad μ . Supongamos que existe una constante $C > 0$ tal que para todo $k \in \mathbb{N}$ vale que:

$$\mathbb{E}(X^k) < C^k \cdot k!$$

Luego, se sigue que X esta determinada por sus momentos.

Demostración. Denotemos al k -esimo momento como α_k . Si se cumple la cota $\alpha_k < C^k \cdot k!$ sabemos que se sigue que la serie $f(x) = \sum \frac{\alpha_k}{k!} \cdot x^k$ tiene radio de convergencia mayor a $\frac{1}{C}$. Vamos a usar eso para demostrar que, dada una variable aleatoria Y con los mismos momentos que X se cumple que $\phi_X = \phi_Y$, de lo que se sigue por el teorema 3.8 que X e Y tienen la misma distribución.

Empecemos calculando el residuo de Taylor de ϕ_X :

$$\begin{aligned} & \left| \phi_X(t+h) - \sum_{k=0}^n \frac{\phi^k(t)}{k!} \cdot h^k \right| = \left| \int e^{i(t+h)x} d\mu - \left(\sum_{k=0}^n \int (ihx)^k \cdot e^{itx} d\mu \right) \right| \\ & = \left| \int (e^{i(t+h)x} - \sum_{k=0}^n (ihx)^k \cdot e^{itx}) d\mu \right| \leq \int |(e^{i(t+h)x} - \sum_{k=0}^n (ihx)^k \cdot e^{itx})| d\mu \\ & = \int |e^{itx}| \cdot |e^{ihx} - \sum_{k=0}^n (ihx)^k| d\mu = \int |e^{ihx} - \sum_{k=0}^n (ihx)^k| d\mu \end{aligned}$$

Pero notemos que el termino siendo integrando no es mas que el residuo de Taylor de orden n de la función e^{ihx} centrado en $x = 0$, que por la fórmula del residuo sabemos que es igual a $\frac{(ih\zeta)^{n+1}}{(n+1)!} \cdot e^{ih\zeta}$ para algún $\zeta \in (0, x)$, luego su modulo es menor a $\frac{(hx)^{n+1}}{(n+1)!}$. Concluimos entonces que:

$$\left| \phi_X(t+h) - \sum_{k=0}^n \frac{\phi^k(t)}{k!} \cdot h^k \right| \leq \int \frac{(hx)^{k+1}}{(k+1)!} = \left(\int x^{n+1} d\mu \right) \cdot \frac{h^{n+1}}{(n+1)!}$$

Y ese ultimo termino no es otra cosa que $\alpha_{n+1} \cdot \frac{h^{n+1}}{(n+1)!}$. Como esa serie tiene radio de convergencia $r = \frac{1}{C} > 0$, podemos concluir que ϕ_X es analítica y tiene radio de convergencia al menos r en todo punto. Pero lo único que usamos en particular de X es que $\mathbb{E}(X) = \alpha_k$ y eso también vale para Y , luego se sigue que Y también es analítica y el radio de convergencia de todo punto es mayor a r . Ahora sabemos que para todo n vale que $\phi_X^{(n)}(0) = \phi_Y^{(n)}(0) = \alpha_n$. Luego ϕ_X y ϕ_Y tienen la misma serie de Taylor en $x = 0$ y por tanto son iguales en $(-r, r)$. Pero también tienen la misma serie de Taylor en $x = r - \epsilon$, luego son iguales en $(-r, 2r - \epsilon)$. Aplicando el mismo razonamiento se sigue que ϕ_X y ϕ_Y son iguales en todo punto. \square

Corolario 3.12. *Toda variable aleatoria acotada esta determinada por sus momentos.*

Demostración. Sea X variable aleatoria tal que casi seguro $|X| < M \in \mathbb{R}$. Se sigue que $\alpha_k = |\int x^k d\mu| < \int |x^k| d\mu < M^k$. \square

Teorema 3.13. (Fréchet-Shohat)[3] *Sea X una variable aleatoria determinada por sus momentos y X_n una sucesión de variables aleatorias tal que para todo k vale que $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^k) = \mathbb{E}(X^k)$. Se sigue que $X_n \rightarrow_{\mathcal{D}} X$.*

Demostración. La demostración clásica usa el teorema de selección de Helly. Daremos una demostración de espíritu similar pero con herramientas de análisis funcional.

Sea $E = C_0(\mathbb{R})$ el espacio Banach de funciones continuas que se anulan en el infinito, con norma infinito. Notemos que E es separable. Podemos ver a cualquier variable aleatoria Y como elementos del dual E^* , definiendo el funcional $Y^* \in E^*$ dado por $Y^*(f) = \mathbb{E}(f(Y))$. Es claro que es lineal y ademas pertenece a la bola unitaria de E^* , ya que $\mathbb{E}(f(Y)) = \int f(x) d\mu_y(x) \leq \|f\|_{\infty}$. Ademas, es fácil ver que $Y_n \rightarrow Y$ converge en distribución si y solo si $Y_n^* \rightarrow Y^*$ converge en E^* con la topología w^* .

Supongamos que el teorema no se cumple. Luego, debe existir una sucesión de variables aleatorias X_n tal que $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^a) = \mathbb{E}(X^a)$ pero ninguna subsucesión de X_n converge a X .

Por el teorema de Banach-Aloaglu secuencial, sabemos que la bola unitaria en E^* es compacta respecto a la topología w^* . Vimos anteriormente que los funcionales X_n^* están en la bola unitaria de E^* , luego debe existir una subsucesión $X_{n_k}^*$ que converge en la topología w^* , de lo que se sigue que X_{n_k} converge en distribución. Supongamos que converge a Y . Se sigue que $\mathbb{E}(Y^a) = \lim_{k \rightarrow \infty} \mathbb{E}(X_{n_k}^a) = \lim_{n \rightarrow \infty} \mathbb{E}(X_n^a) = \mathbb{E}(X^a)$ para todo a y como X esta determinado por sus momentos se sigue que $X = Y$. Pero eso contradice la hipótesis de que ninguna subsucesión de X_n converge a X . \square

Podemos hacer el mismo desarrollo para vectores aleatorios:

Definición 3.14. Sea $X = (X_1, \dots, X_n)$ un vector aleatorio. Su función característica $\phi_X : \mathbb{R}^n \rightarrow \mathbb{C}$ esta definida como $\phi_X(t_1, \dots, t_n) = \mathbb{E}(e^{t_1 X_1 + \dots + t_n X_n})$

Teorema 3.15. [3] *Sean X e Y dos vectores aleatorios tal que $\phi_X = \phi_Y$. Luego X e Y tienen la misma distribución.*

Proposición 3.16. *Sean $X = (X_1, \dots, X_n)$ y $Y = (Y_1, \dots, Y_n)$ dos vectores aleatorios tal que para todo $c_1, \dots, c_n \in \mathbb{R}$ la variable aleatoria $c_1 X_1 + \dots + c_n X_n$ tiene la misma distribución que $c_1 Y_1 + \dots + c_n Y_n$. Luego se sigue que X e Y tienen la misma distribución.*

Demostración. Como $c_1 X_1 + \dots + c_n X_n$ tiene la misma distribución que $c_1 Y_1 + \dots + c_n Y_n$ se sigue que $\mathbb{E}(e^{c_1 X_1 + \dots + c_n X_n}) = \mathbb{E}(e^{c_1 Y_1 + \dots + c_n Y_n})$, es decir, $\phi_X(c_1, \dots, c_n) = \phi_Y(c_1, \dots, c_n)$. Como sus funciones características son iguales, se sigue que X e Y tienen la misma distribución. \square

Definición 3.17. Decimos que un vector aleatorio $X = (X_1, \dots, X_n)$ está **determinado por sus momentos** si para todo vector aleatorio $Y = (Y_1, \dots, Y_n)$ si $\mathbb{E}(X_1^{k_1} \cdot \dots \cdot X_n^{k_n}) = \mathbb{E}(Y_1^{k_1} \cdot \dots \cdot Y_n^{k_n})$ para todo conjunto $k_1, \dots, k_n \in \mathbb{N}$ se sigue que X e Y tienen la misma distribución.

Podemos probar ahora un análogo del Teorema 3.11 y del teorema de Fréchet-Shohat para vectores aleatorios:

Teorema 3.18. (Criterio de Riesz vectorial)[3] Sea $X = (X_1, \dots, X_n)$ un vector aleatorio. Supongamos que existe un $C > 0$ tal que para todo k_1, \dots, k_n se cumple que

$$\mathbb{E}(X_1^{k_1} \cdot (\dots) \cdot X_n^{k_n}) < C^{k_1 + \dots + k_n} \cdot (k_1 + \dots + k_n)!$$

Luego se sigue que X está determinado por sus momentos.

Demostración. Denotemos a $\mathbb{E}(X_1^{k_1} \cdot (\dots) \cdot X_n^{k_n})$ con α_{k_1, \dots, k_n} . Vamos a ver que $c_1 X_1 + \dots + c_n X_n$ tiene la misma distribución que $c_1 Y_1 + \dots + c_n Y_n$ para todo $c_1, \dots, c_n \in \mathbb{R}$, de lo que se sigue por la proposición 3.16 que X e Y tienen la misma distribución.

Empecemos viendo que $c_1 X_1 + \dots + c_n X_n$ está determinada por sus momentos, calculando sus momentos y usando el criterio de Riesz. Sin pérdida de generalidad podemos asumir que $c_1 = \max(c_1, \dots, c_n)$.

$$\mathbb{E}((c_1 X_1 + \dots + c_n X_n)^k) = \sum_{s_1 + \dots + s_n = k} \binom{k}{s_1, \dots, s_n} \cdot \alpha_{s_1, \dots, s_n} \cdot c_1^{s_1} \cdot (\dots) \cdot c_n^{s_n}$$

Sabemos, por hipótesis, que podemos acotar α_{s_1, \dots, s_n} por $C^{s_1 + \dots + s_n} \cdot (s_1 + \dots + s_n)! = C^k \cdot k!$ y como asumimos que c_1 era el mayor de los c_i tenemos también que $c_1^{s_1} \cdot (\dots) \cdot c_n^{s_n} < c_1^k$. Luego nos queda:

$$\mathbb{E}((c_1 X_1 + \dots + c_n X_n)^k) < \sum_{s_1 + \dots + s_n = k} \binom{k}{s_1, \dots, s_n} \cdot C^k \cdot c_1^k \cdot k! = \left(\sum_{s_1 + \dots + s_n = k} \binom{k}{s_1, \dots, s_n} \right) \cdot C^k \cdot c_1^k \cdot k!$$

Como la suma de todos los multinomiales $\binom{k}{s_1, \dots, s_n}$ es n^k tenemos:

$$\mathbb{E}((c_1 X_1 + \dots + c_n X_n)^k) < n^k \cdot C^k \cdot c_1^k \cdot k! = (n \cdot C \cdot c_1)^k \cdot k!$$

Como $c_1 X_1 + \dots + c_n X_n$ está determinado por sus momentos y $c_1 Y_1 + \dots + c_n Y_n$ tiene los mismos momentos que $c_1 X_1 + \dots + c_n X_n$ se sigue que ambos tienen la misma distribución. Como eso vale para todo c_1, \dots, c_n se sigue que X e Y tienen la misma distribución. \square

Para el caso en el que las coordenadas son independientes, tenemos una condición más simple:

Teorema 3.19. [3] Sea $X = (X_1, \dots, X_n)$ un vector aleatorio con coordenadas independientes. Supongamos que existe un $C > 0$ tal que para todo k y todo $i \in \{1, \dots, n\}$ vale

$$\mathbb{E}(X_i^k) < C^k \cdot k!$$

Luego se sigue que X está determinado por sus momentos.

Demostración. Como las coordenadas de X son independientes tengo $\mathbb{E}(X_1^{k_1} \cdot \dots \cdot X_n^{k_n}) = \mathbb{E}(X_1^{k_1}) \cdot \dots \cdot \mathbb{E}(X_n^{k_n}) < C^{k_1} \cdot k_1! \cdot \dots \cdot C^{k_n} \cdot k_n! < C^{k_1 + \dots + k_n} \cdot (k_1 + \dots + k_n)!$. Luego por el teorema anterior se sigue que X esta determinado por sus momentos. \square

Teorema 3.20. (Frechet-Shohat vectorial)[3] Sea $X = (X_1, \dots, X_n)$ un vector aleatorio determinado por sus momentos y $Y_i = (Y_{(i,1)}, \dots, Y_{(i,n)})$ una serie de vectores aleatorios tal que para todo $\alpha \in \mathbb{N}^n$ el α -esimo momento de Y_i converge al α -esimo momento de X . Se sigue que Y converge a X en distribución.

Demostración. Análogo a la demostración que dimos de Frechet-Shohat: Sea $E = C_0(\mathbb{R}^n)$ el espacio Banach de funciones continuas que se anulan en el infinito, con norma infinito. Notemos que E es separable. Podemos ver a cualquier variable aleatoria Y como elementos del dual E^* , definiendo $Y^*(f) = \mathbb{E}(f(Y))$. Esos funcionales pertenecen a la bola unitaria de E^* y dada sucesión de variables aleatorias Y_n , Y_n converge en distribución a Y si y solo si $Y_n^* \rightarrow Y^*$ con la topología w^* .

Supongamos que el teorema no se cumple. Luego, debe existir una sucesión de vectores aleatorios X_n tal que $\lim_{n \rightarrow \infty} \mathbb{E}(X_n^a) = \mathbb{E}(X^a)$ pero **ninguna** subsucesión de X_n converge a X . Por el teorema de Banach-Aloaglu secuencial, sabemos que la bola unitaria en E^* es compacta respecto a la topología w^* . Vimos anteriormente que los funcionales X_m^* están en la bola unitaria de E^* , luego debe existir una subsucesión $X_{n_k}^*$ que converge en la topología w^* , de lo que se sigue que X_{n_k} converge en distribución. Supongamos que converge a Y . Se sigue que $\mathbb{E}(Y^a) = \lim_{k \rightarrow \infty} \mathbb{E}(X_{n_k}^a) = \lim_{n \rightarrow \infty} \mathbb{E}(X_n^a) = \mathbb{E}(X^a)$ para todo a y como X esta determinado por sus momentos se sigue que $X = Y$. Pero eso contradice la hipótesis de que ninguna subsucesión de X_n converge a X \square

3.2.2. Cuatro versiones de la ley de convergencia Poisson

En esta sección usaremos los resultados anteriores para dar para dar una serie de criterios de creciente nivel de generalidad de cuando una sucesión variables aleatorias converge en distribución a una variable Poisson, que usaremos mas adelante. Este tipo de método es estándar en probabilidad y en particular, en el estudio de grafos aleatorios[22][4], aunque no así pero la formulación específica de los dos últimos dos resultados (que denominamos ley de convergencia Poisson III y IV), que elegimos para simplificar demostraciones posteriores. Empezaremos demostrando resultados sobre las distribuciones Poisson, que nos llevaran a ver que están determinadas por sus momentos.

Proposición 3.21. Sea X una variable aleatoria que sigue una distribución Poisson de parámetro λ . Luego:

$$\mathbb{E}(X \cdot (X - 1) \cdot \dots \cdot (X - k + 1)) = \lambda^k$$

Demostración. Tenemos:

$$\begin{aligned} \mathbb{E}(X \cdot \dots \cdot (X - k + 1)) &= \sum_{a=k}^{\infty} a \cdot \dots \cdot (a - k + 1) \cdot Pr(X = a) \\ &= \sum_{a=k}^{\infty} a \cdot \dots \cdot (a - k + 1) \cdot e^{-\lambda} \cdot \frac{\lambda^a}{a!} = \sum_{a=k}^{\infty} e^{-\lambda} \cdot \frac{\lambda^a}{(a - k)!} = \sum_{a=k}^{\infty} e^{-\lambda} \cdot \lambda^k \cdot \frac{\lambda^{a-k}}{(a - k)!} \end{aligned}$$

Reordenando y haciendo el cambio de variable $u = a - k$ nos queda:

$$= \lambda^k \cdot e^{-\lambda} \cdot \sum_{u=0}^{\infty} \frac{\lambda^u}{u!} = \lambda^k \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda^k$$

□

Para continuar debemos introducir un concepto de combinatoria, los números de Stirling del segundo tipo.

Definición 3.22. Los números de Stirling del segundo tipo $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ son la cantidad de formas de partir un conjunto numerado de n elementos en k conjuntos no numerados. Dicho de otra forma, cuentan la cantidad de funciones sobreyectivas $f : \{1, \dots, n\} \rightarrow \{1, \dots, k\}$ bajo la relación de equivalencia dada por reordenamientos de k .

Observación 3.23. Se sigue que $k! \cdot \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ cuenta la cantidad de funciones sobreyectivas $f : \{1, \dots, n\} \rightarrow \{1, \dots, k\}$.

Proposición 3.24. Los números de Stirling del segundo tipo cumplen la siguiente recurrencia:

$$\left\{ \begin{smallmatrix} n+1 \\ k+1 \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} + (k+1) \cdot \left\{ \begin{smallmatrix} n \\ k+1 \end{smallmatrix} \right\}$$

Demostración. Por doble conteo. Por definición, $\left\{ \begin{smallmatrix} n+1 \\ k+1 \end{smallmatrix} \right\}$ cuenta la cantidad de forma de partir $\{1, \dots, n+1\}$ en $k+1$ conjuntos no numerados. Hay dos posibilidades, o bien el 1 queda en un conjunto aislado, de lo que se sigue que hay $\left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\}$ de terminar de partir $\{2, \dots, n+1\}$ en k conjuntos no numerados, o bien el 1 no queda en un conjunto aislado, luego hay $\left\{ \begin{smallmatrix} n \\ k+1 \end{smallmatrix} \right\}$ formas de partir $\{2, \dots, n+1\}$ en $k+1$ conjuntos no numerados, multiplicado por $k+1$ formas de asignar el 1. Luego, nos queda $\left\{ \begin{smallmatrix} n+1 \\ k+1 \end{smallmatrix} \right\} = \left\{ \begin{smallmatrix} n \\ k \end{smallmatrix} \right\} + (k+1) \cdot \left\{ \begin{smallmatrix} n \\ k+1 \end{smallmatrix} \right\}$ □

Proposición 3.25. Sean A_1, \dots, A_n eventos aleatorios y sea $X = \sum 1_{A_i}$. Luego:

$$\mathbb{E}(X^k) = \sum_{a=1}^k \left\{ \begin{smallmatrix} n \\ a \end{smallmatrix} \right\} \cdot \left[\sum_{i_1, \dots, i_a \text{ distintos}} Pr(A_{i_1} \cap \dots \cap A_{i_a}) \right]$$

Donde la suma interna tiene en cuenta el orden de los términos i_s .

Demostración. Como para todo i vale que $1_{A_i}^2 = 1_{A_i}$, la expresión $(\sum 1_{A_i})^k$ expandida consistirá en suma de términos de la forma $Pr(A_{i_1} \cap \dots \cap A_{i_a})$ con $a \leq k$ y i_1, \dots, i_a distintos. Si no tenemos en cuenta el orden, cada término aparecerá una vez por cada función sobreyectiva $f : \{1, \dots, n\} \rightarrow \{1, \dots, a\}$ y sabemos que hay $a! \cdot \left\{ \begin{smallmatrix} n \\ a \end{smallmatrix} \right\}$ funciones de esa naturaleza. Si tenemos en cuenta el orden, como cada a -tupla se puede ordenar de $a!$ formas distintas, cada término aparecerá acompañado solamente por $\left\{ \begin{smallmatrix} n \\ a \end{smallmatrix} \right\}$. □

Definición 3.26. El **factorial descendiente** de x , denotado $(x)_n$, es el polinomio $x \cdot (x-1) \cdot \dots \cdot (x-n+1)$

Proposición 3.27. Para todo $n \in \mathbb{N}$, vale la siguiente igualdad de polinomios:

$$x^n = \sum_{k=0}^n \binom{n}{k} \cdot (x)_k$$

Demostración. Por inducción en n . Es obvio que vale para $n = 0$, veamos el caso inductivo.

$$\begin{aligned} x^{n+1} &= x^n \cdot x = \left(\sum_{k=0}^n \binom{n}{k} \right) \cdot (x)_k \cdot x = \left(\sum_{k=0}^n \binom{n}{k} \right) \cdot (x)_k \cdot x \\ &= \left(\sum_{k=0}^n \binom{n}{k} \right) \cdot (x)_k \cdot (x - k + k) = \sum_{k=0}^n \binom{n}{k} \cdot (x)_{k+1} + k \cdot \binom{n}{k} \cdot (x)_k \end{aligned}$$

Reordenando y usando la identidad 3.24 llegamos a:

$$= \sum_{s=0}^{n+1} \left(\binom{n-1}{s-1} + s \cdot \binom{n-1}{s} \right) \cdot (x)_s = \sum_{s=0}^{n+1} \binom{n}{s} \cdot (x)_s$$

□

Proposición 3.28. (Momentos de variables Poisson) Sea X una variable aleatoria con distribución Poisson de parámetro λ . Luego:

$$\mathbb{E}(X^k) = \sum_{a=0}^k \binom{k}{a} \cdot \lambda^a$$

Demostración. Por la proposición 3.27 sabemos que:

$$\mathbb{E}(X^k) = \mathbb{E}\left(\sum_{a=0}^k \binom{k}{a} \cdot (X)_a\right) = \sum_{a=0}^k \binom{k}{a} \cdot \mathbb{E}((X)_a)$$

Por la proposición 3.21 sabemos que $\mathbb{E}(X \cdot \dots \cdot (X - k + 1)) = \lambda^k$, es decir, que $\mathbb{E}((X)_a) = \lambda^a$, luego la expresión anterior es igual a $\sum_{a=0}^k \binom{k}{a} \cdot \lambda^a$. □

Proposición 3.29. Sea X una variable con distribución Poisson con parámetro λ . Luego X esta determinada por sus momentos.

Demostración. Veamos que para todo k vale la desigualdad $\mathbb{E}(X^k) < k! \cdot (e \cdot \lambda)^k$, de lo que se sigue por el criterio de Riesz que X esta determinada por sus momentos.

$$\mathbb{E}(X^k) = \sum \binom{k}{a} \cdot \lambda^a \leq \sum \binom{k}{a} \cdot \lambda^a = \left(\sum \binom{k}{a} \right) \cdot \lambda^k$$

Ahora, $\sum \binom{k}{a}$ es la cantidad de formas de partir un conjunto de k elementos en conjuntos no numerados, por lo que se puede acotar por la manera de partir k en conjuntos numerados, que es k^k . Usando esa cota y la celebre aproximación de Stirling nos queda:

$$\mathbb{E}(X^k) < k^k \cdot \lambda^k < k! \cdot e^k \cdot \lambda^k = k! \cdot (e \cdot \lambda)^k$$

□

Mas en general:

Proposición 3.30. Sean X_1, \dots, X_n variables aleatorias independientes tal que X_i tiene una distribución Poisson de parámetro λ_i . Luego el vector aleatorio $X = (X_1, \dots, X_n)$ esta determinado por sus momentos.

Demostración. Consecuencia directa de la proposición anterior y el teorema 3.19. \square

Ahora podemos empezar a enunciar las leyes de convergencia Poisson. Empecemos con la versión mas clásica, conocida también como **la ley de los eventos raros**.

Teorema 3.31. (Ley de convergencia Poisson I) Sea $A_{n,i}$ una serie de eventos independientes entre si tal que para todo i , la probabilidad de $A_{n,i}$ es $\frac{\lambda}{n}$. Luego las variables aleatorias $X_n = \sum_{i=1}^n 1_{A_{n,i}}$ convergen en distribución a una variable Poisson de parámetro λ .

Demostración. Veamos que para todo k , el k -esimo momento de X_n converge al k -esimo momento de $Poi(\lambda)$. Como las variables Poisson están determinadas por sus momentos, por el teorema de Frechet-Shohat tenemos que $X_n \rightarrow_{\mathcal{D}} Poi(\lambda)$. Supongamos que X_n es suma de variables Bernoulli A_1^n, \dots, A_n^n . Por la proposición 3.25 tenemos, para todo $k < n$:

$$\mathbb{E}(X_n^k) = \left(\sum_{a=0}^k \binom{k}{a} \right) \cdot \left(\sum_{i_1, \dots, i_a \text{ distintos}} Pr(A_{i_1} \cap \dots \cap A_{i_a}) \right)$$

Como los eventos A_i son independientes con probabilidad $\frac{\lambda}{n}$, la probabilidad de $A_{i_1} \cap \dots \cap A_{i_a}$ es $(\frac{\lambda}{n})^a$. Por otro lado, hay $n \cdot (n-1) \cdot \dots \cdot (n-a+1) = (n)_a$ formas de elegir i_1, \dots, i_a distintos y ordenados.

$$\mathbb{E}(X_n^k) = \sum_{a=0}^k \binom{k}{a} \cdot (n)_a \cdot \left(\frac{\lambda}{n}\right)^a \rightarrow \sum_{a=0}^k \binom{k}{a} \cdot \lambda^a = \mathbb{E}(Poi(\lambda)^k)$$

\square

La primer generalización que consideramos es el caso en lo que los eventos $A_{n,i}$ no son variables independientes de parámetro $\frac{\lambda}{n}$ pero se comportan asintoticamente como si lo fueran:

Teorema 3.32. (Ley de convergencia Poisson II) Sean $A_{n,i}$ una serie de eventos con $i \leq n$ que cumplen que para todo n y todo $i_1, \dots, i_k \leq n$ la probabilidad de $A_{i_1} \cap \dots \cap A_{i_k}$ es la misma para toda elección de i_1, \dots, i_k y si denotamos a dicha probabilidad como $p_{n,k}$ vale que:

$$\lim_{n \rightarrow \infty} p_{n,k} \cdot n^k = \lambda^k.$$

Luego, se sigue que las variables aleatorias $X_n = \sum_{i=1}^n 1_{A_{n,i}}$ convergen en distribución a una variable Poisson de parámetro λ

Demostración. Veamos que para todo k , el k -esimo momento de X_n converge al k -esimo momento de $Poi(\lambda)$. Como las variables Poisson están determinadas por sus momentos, por el teorema de Frechet-Shohat tenemos que

$X_n \rightarrow_{\mathcal{D}} Poi(\lambda)$. Supongamos que X_n es suma de variables Bernoulli A_1^n, \dots, A_n^n . Por la proposición 3.25 tenemos, para todo $k < n$:

$$\mathbb{E}(X_n^k) = \left(\sum_{a=0}^k \binom{k}{a} \right) \cdot \left(\sum_{i_1, \dots, i_a \text{ distintos}} Pr(A_{n,i_1} \cap (\dots) \cap A_{n,i_a}) \right)$$

Por hipótesis sabemos que la probabilidad de $A_{i_1} \cap (\dots) \cap A_{i_k}$ es $p_{n,k}$

$$\mathbb{E}(X_n^k) = \sum_{a=0}^k \binom{k}{a} \cdot \binom{n}{a} \cdot p_{n,a} = \sum_{a=0}^k \binom{k}{a} \cdot \frac{\binom{n}{a}}{n^a} \cdot n^a \cdot p_{n,a}$$

Ahora, sabemos que $\lim_{n \rightarrow \infty} \binom{n}{a}/n^a = 1$ y sabemos por hipótesis se cumple que $\lim_{n \rightarrow \infty} p_{n,a} \cdot n^a = \lambda^a$. Luego nos queda:

$$\mathbb{E}(X_n^k) \rightarrow \sum_{a=0}^k \binom{k}{a} \cdot \lambda^a$$

□

La siguiente generalización resultara un tanto ad-hoc, pero resultara de crucial importancia para la siguientes sección. En lugar de considerar eventos de la forma $A_{i,n}$ con $i < n$ consideraremos eventos de la forma $A_{n,S}$ donde S es cualquier subconjunto de $\{1, \dots, n\}$ de tamaño dado, que, al igual que en la ley de Poisson II, se comportaran asintoticamente como variables independientes.

Teorema 3.33. (Ley de convergencia Poisson III) Supongamos que, para todo n y para cada subconjunto S de $\{1, \dots, n\}$ de tamaño s tenemos asociado un evento A_S , que cumplen las siguientes propiedades:

- Para todo n y todo a , la probabilidad de $A_{S_1} \cap (\dots) \cap A_{S_a}$ es la misma para toda colección de subconjuntos disjuntos $S_1, \dots, S_a \subset \{1, \dots, n\}$ y mas aun, si denotamos a dicha probabilidad $p_{n,a}$ se cumple que:

$$\lim_{n \rightarrow \infty} (p_{n,a} \cdot n^{sa}) = \lambda^a$$

- La probabilidad de que ocurran dos eventos A_S y A_T tal que $S \neq T$ y $S \cap T \neq \emptyset$ tiende a 0 cuando n tiende a infinito.

Luego las variables $X_n = \sum_S 1_{A_S}$ convergen en distribución a una variable Poisson con parámetro $\frac{\lambda}{s!}$.

Demostración. Vamos a ver que para todo k , el k -esimo momento de X_n converge al k -esimo momento de una Poisson de parámetro $c = \frac{\lambda}{s!}$, y como las variables Poisson están determinadas por sus momentos, se sigue por el teorema de Frechet-Shohat que $X_n \rightarrow_{\mathcal{D}} Poi(c)$.

$$\mathbb{E}(X_n^k) = \mathbb{E}\left(\left(\sum_S 1_{A_S}\right)^k\right) = \sum_{a=0}^k \binom{k}{a} \cdot \sum_{S_1, \dots, S_a \text{ distintos}} Pr(A_{S_1} \cap (\dots) \cap A_{S_a})$$

Ahora, por la segunda condición sabemos que la probabilidad de que ocurran dos eventos distintos y no disjuntos tiende a 0, luego asintoticamente esto es igual a:

$$= \sum_{a=0}^k \binom{k}{a} \cdot \sum_{S_1, \dots, S_a \text{ disjuntos}} Pr(A_{S_1} \cap (\dots) \cap A_{S_a})$$

Por $Pr(A_{S_1} \cap (\dots) \cap A_{S_a}) = p_{n,a}$ para toda elección de a subconjuntos de $\{1, \dots, n\}$ disjuntos. Además, la cantidad de tuplas (S_1, \dots, S_a) donde los S_i son subconjuntos disjuntos de $\{1, \dots, n\}$ de tamaño s es $\binom{n}{s} \cdot \binom{n-s}{s} \cdot (\dots) \cdot \binom{n-(a-1)s}{s} = \frac{(n)_{sa}}{s!^a}$

$$= \sum_{a=0}^k \binom{k}{a} \cdot (n)_{as} \cdot p_{n,a} = \sum_{a=0}^k \binom{k}{a} \cdot \frac{(n)_{as}}{n^{as}} \cdot n^{as} \cdot p_{n,a} \rightarrow \sum_{a=0}^k \binom{k}{a} \cdot c^a$$

□

Para nuestra última generalización, consideramos eventos de la forma $A_{n,S}$ onde $S \subset \{1, \dots, n\}$ y consideraremos múltiples tamaños posibles de S simultáneamente.

Teorema 3.34. (Ley de convergencia Poisson IV) Sean s_1, \dots, s_d una serie de naturales (no necesariamente distintos). Supongamos que para todo n y para todo subconjunto $S \subset \{1, \dots, n\}$ de tamaño s_i tenemos asociado un evento A_S^i , que cumplen las siguientes propiedades:

- Sea $\bar{k} = (k_1, \dots, k_d) \in \mathbb{N}^d$. Para todo n , dados k_1 subconjuntos de $\{1, \dots, n\}$ de tamaño s_1 , k_2 subconjuntos de $\{1, \dots, n\}$ de tamaño s_2 , ..., k_d subconjuntos de $\{1, \dots, n\}$ de tamaño s_d , la probabilidad de que sus respectivos eventos ocurran simultáneamente es la misma para cualquier elección de subconjuntos disjuntos. Mas aun, si denotamos a dicha probabilidad $p_{n, \bar{k}}$, vale que:

$$\lim_{n \rightarrow \infty} [p_{n,d} \cdot n^{\sum s_i \cdot k_i}] = \lambda_1^{k_1} \cdot (\dots) \cdot \lambda_d^{k_d}$$

- La probabilidad de que ocurran eventos A_S^i y A_T^j con S y T subconjuntos distintos pero no disjuntos tiende a 0 cuando n tiende a infinito.

Sea $X_{(n,i)} = \sum_S 1_{A_S^i}$. Luego las variables $X_{(n,i)}$ converge a una distribución Poisson de parámetro $\lambda_i/s_i!$ **independientes entre si**.

Demostración. La demostración es muy similar a las anteriores, pero se vuelve bastante mas difícil de seguir por la gran cantidad de subíndices involucrados. Para simplificar un poco la notación, vamos a denotar al evento A_S^i simplemente como S^i .

Sea X_n el vector aleatorio dado por $(X_{n,1}, (\dots), X_{n,d})$ y sea Y el vector aleatorio dado por d variables Poisson independientes de parámetros $\lambda_1, \dots, \lambda_d$. Vamos a ver que para todo $\bar{k} \in \mathbb{N}^d$ el \bar{k} -esimo momento de X_n converge al \bar{k} -esimo momento de Y , es decir, que para todo k_1, \dots, k_d vale que:

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_{(n,1)}^{k_1} \cdot (\dots) \cdot X_{(n,d)}^{k_d}) = \mathbb{E}(Poi(\lambda_1)^{k_1}) \cdot (\dots) \cdot \mathbb{E}(Poi(\lambda_d)^{k_d})$$

Como por el teorema 3.30 el vector aleatorio Y esta determinado por sus momentos, se sigue que X_n converge en distribución Y . Empecemos expandiendo cada termino X_n en suma de indicatoras:

$$\mathbb{E}(X_{(n,1)}^{k_1} \cdot (\dots) \cdot X_{(n,d)}^{k_d}) = \mathbb{E}\left(\prod_{i=1}^d \left(\sum_S 1_{S^i}\right)^{k_i}\right)$$

Expandiendo cada termino de la forma $(\sum 1_S)^{k_i}$ de acuerdo con la fórmula dada en la proposición 3.25 nos queda:

$$= \mathbb{E}\left(\prod_{i=1}^d \left(\sum_{a_i=0}^{k_i} \binom{k_i}{a_i}\right) \cdot \sum_{S_1, \dots, S_{a_i} \text{ distintos}} 1_S \cdot (\dots) 1_{S_i}\right) =$$

Expandiendo el producto usando la propiedad distributiva llegamos a la siguiente expresión, donde la suma interior es sobre todos las posibles elecciones de a_1 subconjuntos distintos de tamaño s_1 , a_2 subconjuntos distintos de tamaño s_2 , ..., a_d subconjuntos de tamaño s_d .

$$\sum_{a_1, \dots, a_d} \binom{k_1}{a_1} \cdot (\dots) \cdot \binom{k_d}{a_d} \sum Pr(\bigcap S^i)$$

Ahora, por la segunda condición, la probabilidad de que ocurran simultáneamente eventos S^i y T^j con S y T distintos pero no disjuntos es asintoticamente 0, luego es equivalente considerar la suma sobre subconjuntos **disjuntos**. Por la primer condición, la probabilidad es uniforme para todo elección de subconjuntos, dada por $p_{n, \bar{a}}$.

$$\begin{aligned} &= \sum_{a_1, \dots, a_d} \binom{k_1}{a_1} \cdot (\dots) \cdot \binom{k_d}{a_d} \cdot \sum p_{n, \bar{a}} = \sum_{a_1, \dots, a_d} \binom{k_1}{a_1} \cdot (\dots) \cdot \binom{k_d}{a_d} \cdot \lambda_1^{k_1} \cdot (\dots) \cdot \lambda_d^{k_d} \\ &= \prod_{i=1}^d \left(\sum_{a_i=0}^{k_i} \binom{k_i}{a_i} \lambda_i^{k_i}\right) = \prod_{i=1}^d \mathbb{E}(Poi(\lambda_i)^{k_i}) \end{aligned}$$

□

Por supuesto, el teorema anterior no es el mas general posible: Siempre se puede descender mas y mas hacia las profundidades de la abstracción. Resultados de carácter muy general se encuentran en [2]. Dichos resultados no se demuestran usando el método de los momentos, que como muestra la última demostración se vuelve extremadamente engorroso de usar en casos generales, sino usando el método de Chen-Stein, que ademas de ser mas elegante permite hallar cotas explicitas en la velocidad de convergencia.

3.3. Grafos muy Sparse

3.3.1. Estructura local

Usando el método de los momentos, podemos empezar a describir la naturaleza asintótica de los grafos aleatorios muy sparse.

Proposición 3.35. Sea $G \sim \mathbf{G}(n, \frac{c}{n})$ y x un vértice de G . Luego, la cantidad de vecinos de x converge a una distribución Poisson de parámetro c cuando $n \rightarrow \infty$.

Demostración. Cada $z \in G - x$ esta conectado a x con probabilidad $\frac{c}{n}$ y cada una de esas variables es independiente entre si. Luego, el numero de vecinos de x es la suma de $n - 1$ variables Bernoulli independientes con probabilidad $\frac{c}{n}$ y por el teorema de convergencia Poisson I se sigue que converge a $Poi(c)$ cuando $n \rightarrow \infty$. \square

Proposición 3.36. *Sea H un grafo con a aristas y v vértices tal que $a > v$. Luego la probabilidad de que H sea un subgrafo de $G(\frac{c}{n}, n)$ tiende a 0 cuando $n \rightarrow \infty$*

Demostración. Sea X_H la cantidad de conjuntos de v elementos de G que son isomorfos a H . $X_H = \frac{1}{|Aut(H)|} \sum S_A$ donde la suma es sobre todos los subconjuntos A de v vértices de $\{1, \dots, n\}$ y S_A es el evento que dicho subgrafo sea isomorfo a H . Fijado un orden de los vértices de A , la probabilidad de que el isomorfismo se realice en ese orden en particular es $\frac{v!}{|Aut(H)|}$ y hay $\frac{v!}{|Aut(H)|}$ ordenes distintos a considerar. Nos queda:

$$\begin{aligned} \mathbb{E}(X_H) &= \mathbb{E}\left(\sum_S 1_{S_A}\right) = \sum_S Pr(A \cong H) = \frac{v!}{|Aut(H)|} \sum_S \left(\frac{c}{n}\right)^a \cdot \left(1 - \frac{c}{n}\right)^{\binom{v}{2}-a} \\ &= \binom{n}{v} \cdot \frac{v!}{|Aut(H)|} \cdot \left(\frac{c}{n}\right)^a \cdot \left(1 - \frac{c}{n}\right)^{\binom{v}{2}-a} \rightarrow \frac{1}{|Aut(H)|} \cdot c^a \cdot n^{v-a} \end{aligned}$$

Como $v > a$, es claro que ese termino tiende a 0 cuando n tiende a infinito. Luego, como por la desigualdad de Markov tenemos que $Pr(X_H \geq 1) \geq \mathbb{E}(X_H)$ se sigue que $\lim_{n \rightarrow \infty} Pr(X_H \geq 0) = 0$. \square

La proposición anterior limita muy fuertemente que tipo de grafos pueden aparecer asintoticamente como subgrafos de un grafo aleatorio muy sparse, como muestra la siguiente proposición. En particular, asintoticamente no puede aparecer ningún grafo completo K_l con $l \geq 4$.

Proposición 3.37. *Sea G un grafo conexo con a aristas y v vértices. Luego, $a \geq v - 1$. Mas aun, si $a = v - 1$ el grafo G es un árbol y si $a = v$ el grafo G es unicyclico, es decir, tiene un único ciclo.*

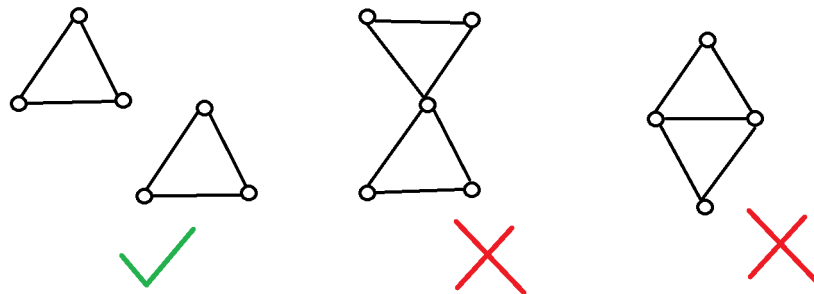
Demostración. Por inducción en la cantidad de vértices. Sea G un grafo con v vértices y $v - 1$ aristas. Como cada arista cubre solo dos vértices, debe haber al menos un vértice x que tiene un solo vecino. El grafo $G - \{x\}$ tiene $v - 1$ vértices y $v - 2$ aristas, luego por hipótesis inductiva es un árbol. Como x tiene un solo vecino, no puede pertenecer a ningún ciclo y por tanto G también es un árbol. Si le agregamos una arista ($a = v$) el grafo se volverá unicyclico y si le sacamos una arista ($a < v - 1$) el grafo se vuelve desconexo. \square

Luego, localmente los grafos aleatorios muy sparse se componen unicamente de arboles y de ciclos. Vamos a clasificar la distribución de ambos, empezando por los ciclos:

Teorema 3.38. *Dado $k \geq 3$, sea $X_{k,n}$ la variable aleatoria que cuenta la cantidad de k -ciclos en una grafo aleatorio sparse $G(n, \frac{c}{n})$. Luego $X_{k,n}$ converge en distribución a una Poisson de parámetro $\lambda = \frac{c^k}{2k}$.*

Demostración. Para todo subconjunto $S \subset \{1, \dots, n\}$ de tamaño k sea A_S el evento " S forma un k -ciclo en $G(n, \frac{c}{n})$. Por definición $X_{k,n} = \sum 1_{A_S}$. Veamos que se cumplen las hipótesis de la ley de convergencia Poisson III, lo que nos garantiza que $X_{k,n}$ converge a una distribución Poisson.

Empecemos con la segunda hipótesis. Dados dos subconjuntos $S, T \subset \{1, \dots, n\}$ de tamaño k distintos pero no disjuntos, si tanto S como T forman un k -ciclo es fácil ver que $S \cup T$ forma un subgrafo de más aristas que vértices, como ilustra la imagen para el caso $k = 3$. Pero vimos en 3.36 que la probabilidad de que $G(n, \frac{c}{n})$ contenga un subgrafo de esas características tiende a 0 cuando n tiende a infinito. Luego, la probabilidad de que ocurran simultáneamente A_S y A_T con S, T distintos pero no disjuntos tiende a 0 cuando n tiende a infinito.



Veamos ahora que se cumple la primera hipótesis. Dados elementos x_1, \dots, x_k la probabilidad de que formen un ciclo en ese orden, es decir, de que $x_1 \sim x_2, x_2 \sim x_3, \dots, x_k \sim x_1$ y $x_i \not\sim x_j$ para otros valores de i, j es:

$$\left(\frac{c}{n}\right)^k \cdot \left(1 - \frac{c}{n}\right)^{\binom{k}{2} - k}$$

Dado un subconjunto $S \subset \{1, \dots, n\}$ de tamaño k hay $(k-1)!/2$ ordenamientos posibles hasta traslaciones y reflexiones. Luego, la probabilidad de que un subconjunto S de tamaño k forme un ciclo es:

$$\left(\frac{c}{n}\right)^k \cdot \left(1 - \frac{c}{n}\right)^{\binom{k}{2} - k} \cdot \frac{(k-1)!}{2}$$

Además, dados subconjuntos disjuntos $S_1, \dots, S_a \subset \{1, \dots, n\}$ los eventos A_{S_i} son independientes entre sí. Luego, se sigue que la probabilidad de $A_{S_1} \cap \dots \cap A_{S_a}$ es la misma para toda elección de subconjuntos disjuntos y si denotamos a dicha probabilidad $p_{n,a}$ vale que:

$$p_{n,a} \cdot n^{ka} = \left[\left(\frac{c}{n}\right)^k \cdot \left(1 - \frac{c}{n}\right)^{\binom{k}{2} - k} \cdot \frac{(k-1)!}{2}\right]^a \cdot n^{ka} = \left(c^k \cdot \frac{(k-1)!}{2}\right)^a \cdot \left(1 - \frac{c}{n}\right)^{a \cdot (\binom{k}{2} - k)}$$

Como $\lim_{n \rightarrow \infty} \left(1 - \frac{c}{n}\right)^{a \cdot (\binom{k}{2} - k)} = 1$ se sigue que $\lim_{n \rightarrow \infty} [p_{n,a} \cdot n^{ka}] = \left(c^k \cdot \frac{(k-1)!}{2}\right)^a$ se sigue por la ley de convergencia Poisson III que $X_{k,n}$ converge en distribución a una variable Poisson de parámetro $\lambda = \left(c^k \cdot \frac{(k-1)!}{2}\right)/k! = c^k/2k$ \square

Observación 3.39. Del teorema anterior se sigue que los grafos aleatorios sparse **no** siguen una ley 0-1 respecto a la lógica de primer orden, ya que la oración $\phi = \exists x \exists y \exists z : (x \sim y) \wedge (x \sim z) \wedge (y \sim z)$, que establece la existencia de un 3-ciclo, tiene probabilidad asintótica $1 - e^{-\frac{c^3}{6}}$

Observación 3.40. Como mencionamos anteriormente, la gran desventaja del método de los momentos es que no nos da cotas explícitas en cuanto rápido una serie de variables converge. Usando el método de Chen-Stein, se puede ver que si $X_{k,n}$ es la cantidad de k -ciclos en $G(n, \frac{c}{n})$, entonces $d_{VA}(X_{k,n}, Poi(\frac{c^k}{2^k})) = O(\frac{1}{n})$ [4]

Mas en general:

Teorema 3.41. (Distribución de ciclos en grafos sparse) Dado $k \geq 3$, sea $X_{k,n}$ la variable aleatoria que cuenta la cantidad de k -ciclos en un grafo aleatorio sparse $G(n, \frac{c}{n})$. Luego $X_{k,n}$ converge en distribución a una Poisson de parámetro $\lambda = \frac{c^k}{2^k}$ y mas aun, para distintos valores de k las $X_{k,n}$ convergen a variables independientes entre si.

Demostración. Muy similar a la demostración anterior. Para todo subconjunto $S \subset \{1, \dots, n\}$ de tamaño k_i sea A_S el evento " S forma un k_i -ciclo en $G(n, \frac{c}{n})$ ". Por definición $X_{k_i,n} = \sum 1_{A_S}$. Veamos que se cumplen las hipótesis de la ley de convergencia Poisson IV, lo que nos garantiza que $X_{k_i,n}$ convergen a distribuciones Poisson independientes entre si.

Empecemos con la segunda hipótesis. Dados dos subconjuntos $S, T \subset \{1, \dots, n\}$ de tamaño k_i, k_j distintos pero no disjuntos, si S forma un k_i -ciclo y T forma un k_j -ciclo es fácil ver que $S \cup T$ forma un subgrafo con mas aristas que vértices. Pero vimos en 3.36 que la probabilidad de que $G(n, \frac{c}{n})$ contenga un subgrafo de esas características tiende a 0 cuando n tiende a infinito. Luego, la probabilidad de que ocurran simultaneamente A_S y A_T con S, T distintos pero no disjuntos tiende a 0 cuando n tiende a infinito.

Veamos ahora la segunda hipótesis. Vimos en la demostración anterior que la probabilidad de que un subconjunto de tamaño k_i formara un ciclo es $(\frac{c}{n})^{k_i} \cdot (1 - \frac{c}{n})^{\binom{k_i}{2} - k} \cdot \frac{(k_i - 1)!}{2}$ y que dichas probabilidades son independientes para subconjuntos disjuntos. Luego, dados a_1 subconjuntos de tamaño k_1 , a_2 subconjuntos de tamaño k_2 , (...), a_d subconjuntos de tamaño k_d , la probabilidad de que todos formen un ciclo es:

$$p_{n,\bar{a}} = \prod_{i=1}^d [(\frac{c}{n})^{k_i} \cdot (1 - \frac{c}{n})^{\binom{k_i}{2} - k} \cdot \frac{(k_i - 1)!}{2}]^{a_i}$$

Como los términos de la forma $(1 - \frac{c}{n})^{\binom{k_i}{2} - k}$ convergen uniformemente a 1 cuando n tiende a infinito nos queda:

$$p_{n,\bar{a}} \cdot n^{\sum a_i \cdot k_i} \rightarrow \prod [c_i^{k_i} \cdot \frac{(k_i - 1)!}{2}]^{a_i}$$

Luego por la ley de Convergencia Poisson IV se sigue que las $X_{k_i,n}$ convergen en distribución a variables Poisson independientes, de parámetro:

$$\lambda_{k_i} = \frac{c_i^{k_i} \cdot \frac{(k_i - 1)!}{2}}{k_i!} = \frac{c_{k_i}}{2 \cdot k_i}$$

□

Proposición 3.42. Sea $G(n, \frac{c}{n})$ un grafo aleatorio sparse y $x \in G$ un vértice dado. La probabilidad de que x este en un ciclo de tamaño k tiende a 0 cuando $n \rightarrow \infty$ para todo k .

Demostración. Si la probabilidad asintótica tendiera a $p > 0$, el numero esperado de ciclos de tamaño menor a k en $G(n, \frac{c}{n})$ tendería a $p \cdot n$. Pero por la proposición 3.41 sabemos que esa esperanza tiende a $\frac{c^k}{2 \cdot k}$. \square

Luego, como heurística, tenemos que un grafo aleatorio sparse $G(n, \frac{c}{n})$ consiste en (muchos) arboles y (algunos) ciclos.

Proposición 3.43. Sean k_1, k_2 y d números naturales. Luego, existe una constante K tal que la probabilidad de que en $G(n, \frac{c}{n})$ existan un k_1 -ciclo y un k_2 -ciclo a distancia menor a d es menor a $\frac{K}{n}$. En particular, la probabilidad de en $G(n, \frac{c}{n})$ existan un k_1 -ciclo y un k_2 -ciclo a distancia menor a d tiende a 0 cuando n tiende a infinito.

Demostración. La existencia de un k_1 -ciclo y un k_2 -ciclo a distancia r se corresponde a la aparición de un subgrafo de $k_1 + k_2 + r - 1$ vértices y $k_1 + k_2 + r$ aristas. Vimos en la demostración de 3.36 que la probabilidad de que eso pase esta acotado por $\frac{c^{k_1+k_2+r}}{n}$. Luego, la probabilidad de que existan un k_1 -ciclo y un k_2 -ciclo a distancia menor a d es menor a:

$$\sum_{r=1}^d \frac{c^{k_1+k_2+r}}{n} = \frac{1}{n} \cdot \left(\sum_{r=1}^d c^{k_1+k_2+r} \right) = \frac{1}{n} \cdot \frac{c^{k_1+k_2+d} - c^{k_1+k_2}}{c - 1}$$

\square

Proposición 3.44. Sea $c < 1$. Cuando n tiende a infinito, la probabilidad de que existan dos ciclos conectados en $G(n, \frac{c}{n})$ tiende a 0.

Demostración. Dados k_1, k_2 , la demostración muestra que la probabilidad de que exista un k_1 -ciclo y un k_2 -ciclo conectados en $G(n, \frac{c}{n})$ es menor a:

$$\frac{1}{n} \cdot \left(\sum_{r=1}^{\infty} c^{k_1+k_2+r} \right) = \frac{1}{n} \cdot \frac{c^{k_1+k_2}}{1 - c}$$

Luego, la probabilidad de que existan dos ciclos conectados es menor a :

$$\frac{1}{n} \left(\sum_{k_1, k_2} \cdot \frac{c^{k_1+k_2}}{1 - c} \right) < \frac{1}{n} \sum_s \left(\frac{c^s}{1 - c} \cdot s \right)$$

Como la suma converge, se sigue que la probabilidad tiende a 0 \square

El resultado anterior no es cierto para $c > 1$, ya que como establece el teorema 3.2 dichos grafos tienen asintoticamente un componente conexo gigante de tamaño $O(n)$, que contiene muchos ciclos. Sin embargo, podemos demostrar la siguiente proposición, que muestra que incluso si hay ciclos conexos no pueden estar demasiado cerca entre si:

Proposición 3.45. Sea $c > 1$ y $k_1, k_2 \in \mathbb{N}$. Luego, la probabilidad de que $G(n, \frac{c}{n})$ tenga un k_1 -ciclo y un k_2 -ciclo a distancia menor a $d = \frac{\ln(n)}{2 \cdot \ln(c)}$ tiende a 0 cuando n tiende a infinito.

Demostración. Por la demostración de 3.43, sabemos que la probabilidad de que exista un k_1 -ciclo y un k_2 -ciclo a distancia menor a d es menor a:

$$\frac{1}{n} \cdot \frac{c^{k_1+k_2+d} - c^{k_1+k_2}}{c-1} < \frac{1}{n} \cdot c^{k_1+k_2+d} = \frac{1}{n} \cdot c^{k_1+k_2} \cdot \sqrt{n} \rightarrow 0$$

□

Ahora clasificaremos la aparición de arboles aislados:

Proposición 3.46. Sea $c > 0$ y $G \sim \mathbf{G}(n, \frac{c}{n})$ un grafo aleatorio sparse. Sea T un árbol de v vértices y sea X la variable aleatoria que cuenta la cantidad de copias aisladas de T en G . Luego, para todo ε cuando n tiende a infinito casi seguro se cumple:

$$(1 - \varepsilon) \cdot e^{-c \cdot v} \cdot \frac{c^{v-1}}{|Aut(T)|} \cdot n < X < (1 + \varepsilon) \cdot e^{-c \cdot v} \cdot \frac{c^{v-1}}{|Aut(T)|} \cdot n$$

Demostración. Dado un subconjunto $S \subset G$, sea A_S el evento “<El subconjunto S forma una copia aislada de T >”. Se sigue que $X = \sum_S 1_{A_S}$. Luego:

$$\mathbb{E}(X) = \sum Pr(A_S) = \binom{n}{v} \cdot \left(\frac{c}{n}\right)^{v-1} \cdot \left(1 - \frac{c}{n}\right)^{\binom{v}{2}-v+1} \cdot \frac{1}{|Aut(T)|}$$

Como el término tiende a 1 cuando n tiende a infinito nos queda:

$$\mathbb{E}(X) \rightarrow e^{-c \cdot v} \cdot \frac{c^{v-1}}{|Aut(T)|} \cdot n$$

Por otro lado:

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum_S 1_{A_S}\right)^2 = \sum_S Pr(A_S) + \sum_{S,U \text{ distintos}} Pr(A_S \cap A_U)$$

Ahora, como el evento A_S significa que S es una copia aislada de T , es imposible que ocurran A_S y A_U donde S y U son conjuntos distintos pero no aislados. Luego nos queda:

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum_S 1_{A_S}\right)^2 = \sum_S Pr(A_S) + \sum_{S,U \text{ disjuntos}} Pr(A_S \cap A_U)$$

Notemos que el término no es otra cosa que $\mathbb{E}(X)$. Por otro lado, si U y A son disjuntos los eventos A_S y A_U son independientes. Luego nos queda:

$$\mathbb{E}(X^2) = \mathbb{E}(X) + \binom{n}{v,v} \cdot \left(\left(\frac{c}{n}\right)^{v-1} \cdot \left(1 - \frac{c}{n}\right)^{\binom{v}{2}-v+1} \cdot \frac{1}{|Aut(T)|}\right)^2$$

$$\mathbb{E}(X^2) \rightarrow \mathbb{E}(X) + \mathbb{E}(X)^2$$

Luego:

$$Var(X) \rightarrow \mathbb{E}(X)$$

$$\sigma(X) \rightarrow \sqrt{\mathbb{E}(X)}$$

Concluimos mediante la desigualdad de Chebyshev que para todo $\varepsilon > 0$ vale que:

$$Pr(|X - \mu| > \varepsilon \cdot \mu) = Pr(|X - \mu| > \varepsilon \cdot \sigma^2) = \frac{1}{\varepsilon \cdot \sigma} \rightarrow 0$$

De lo que se sigue el enunciado. \square

Corolario 3.47. Para todo árbol T , casi seguro $G(n, \frac{c}{n})$ contiene una cantidad arbitraria de copias de T cuando n tiende a infinito.

Para clasificar mejor el comportamiento local de $G(n, \frac{c}{n})$, introduciremos el conjunto de tipo de un ciclo, que clasifica que estructura aparece alrededor de un ciclo dado:

Definición 3.48. Sea $k > 0, T_1, \dots, T_k$ arboles con raíz de altura r . Decimos que elementos $S = (s_1, \dots, s_k)$ forman un k -ciclo de r -tipo (T_1, \dots, T_k) si forman un k -ciclo y para todo i se cumple que $B_r(s_i) - S \cong T_i$

Teorema 3.49. (Distribución asintótica de ciclos de r -tipo dado) Sea $k \geq 3, r \geq 0, (T_1, \dots, T_k)$ arboles con raíz de altura menor o igual a r tal que T_i tiene a_i vértices y $b_i \leq a_i$ vértices a distancia de la raíz menor a r . Sea $X_{n,\tau}$ la variable aleatoria que cuenta la cantidad de k -ciclos de r -tipo $\tau = (T_1, \dots, T_r)$ en el grafo aleatorio sparse $G(n, \frac{c}{n})$. Luego $X_{n,\tau}$ converge a una variable Poisson de parámetro

$$\lambda = \frac{c^{\sum a_i} \cdot e^{-c \cdot \sum b_i}}{|Aut(\tau)|}$$

Y mas aun, las variables convergen a variables independientes para distintos τ .

Demostración. Muy similar a la demostración del teorema 3.41. Para todo subconjunto $U \subset \{1, \dots, n\}$ de tamaño $\sum a_i$ sea A_U el evento "Existen z_1, \dots, z_k en U tal que forman un k -ciclo de r -tipo τ y $B_r(z_i) = U$ ". Por definición $X_\tau = \sum 1_{A_U}$. Veamos que se cumplen las hipótesis de la ley de convergencia Poisson IV, lo que nos garantiza que $X_{\tau,n}$ convergen a distribuciones Poisson independientes entre sí.

Empecemos viendo que se cumple la segunda hipótesis, es decir, que la probabilidad de que ocurran A_U y A_V donde U y V son distintos pero no disjuntos tiende a 0. Si los ciclos son distintos se sigue por el mismo argumento que en la demostración 3.41. Si los ciclos son iguales, se sigue porque los tipos de arboles requeridos serian distintos, luego es imposible que se cumplan simultáneamente.

Ahora veamos que se cumple la primer hipótesis. Si U_1, \dots, U_n son disjuntos, se sigue que los eventos A_{U_1}, \dots, A_{U_s} son independientes. Luego basta ver que la probabilidad de que dado un $U \subset \{1, \dots, n\}$ de tamaño $\sum a_i$ se cumpla A_U tiende a $\lambda \cdot n^{-\sum a_i}$.

Hay .Como el grafo inducido que debe realizar U es unicyclico sabemos que hay $\sum a_i$ aristas. Además, todos los elementos a distancia menor a r , que son $\sum b_i$ no deben tener ningún vecino salvo los de U , y la probabilidad asintótica de que pase eso es igual a la probabilidad de que sea un vértice aislado, que sabemos que es e^{-c} . Juntando todo, nos queda:

$$p = \frac{\sum a_i!}{|Aut(\tau)|} \cdot \left(\frac{c}{n}\right)^{\sum a_i} \cdot \left(1 - \frac{c}{n}\right)^{\binom{\sum a_i}{2} - \sum a_i} \cdot (e^{-c})^{\sum b_i}$$

El termino $(1 - \frac{c}{n})^{\binom{\sum a_i}{2} - \sum a_i}$ converge a 1 cuando n tiende a infinito .Luego por la ley de Convergencia Poisson IV se sigue que converge a una variable Poisson de parámetro

$$\lambda = \frac{\sum a_i!}{|Aut(\tau)|} \cdot c^{\sum a_i} \cdot (e^{-c})^{\sum b_i} \cdot \frac{1}{(\sum a_i)!} = \frac{1}{|Aut(\tau)|} \cdot c^{\sum a_i} \cdot (e^{-c})^{\sum b_i}$$

□

Ejemplo 3.50. Un k -ciclo aislado es un k -ciclo con 1-tipo $(\emptyset, \emptyset, \dots, \emptyset)$, luego por el teorema anterior la cantidad de k -ciclos aislados en $G(n, \frac{c}{n})$ converge a una variable Poisson con parámetro $\lambda = e^{-ck} \cdot \frac{c^k}{2^k}$. En particular, se cumple que:

$$\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \text{ tiene un } k\text{-ciclo aislado}) = 1 - e^{-e^{-ck} \cdot \frac{c^k}{2^k}}$$

Definición 3.51. Sea $k > 0$, T_1, \dots, T_k clases de m -equivalencia de arboles con raíz de altura r . Decimos que elementos $S = (s_1, \dots, s_k)$ forman un k ciclo de (r, m) -**pseudotipo** (T_1, \dots, T_k) si forman un k -ciclo y para todo i se cumple que $B_r(s_i) - S \equiv_m T_i$

Observación 3.52. Para todo $k \geq 3$, $r > 0, m > 0$, hay una cantidad finito de (m, r) -pseudotipos de k -ciclos.

El siguiente lema muestra que, si se cumple una cota uniforme en las colas, una sumatoria infinita de una sucesión de variables que convergen en distribución a variables Poisson independientes converge a su vez a una distribución Poisson.

Lema 3.53. Sean $X_{n,t}$ variables aleatorias discretas tal que:

- Para todo n , cuando t tiende a infinito $X_{n,t}$ converge en distribución a una variable Poisson de parámetro λ_n y mas aun, para distintos n convergen a variables independientes entre si.
- Para todo ε , existe un n_0 tal que la probabilidad de que para todo $n > n_0$ y todo t se cumpla que $X_{n,t} = 0$ es mayor a $1 - \varepsilon$.

Luego, la sumatoria $\sum_n \lambda_n < \infty$ y cuando t tiende a infinito $\sum_n X_{n,t}$ converge en distribución a una variable Poisson de parámetro $\sum_n \lambda_n$.

Demostración. Empecemos viendo que la sumatoria $\sum \lambda_n$ converge. Supongamos que no. Luego, para todo n_0 existe un n_1 tal que $\sum_{n=n_0}^{n_1} \lambda_n > 1$. Sabemos que para todo n en ese rango las variables $X_{n,t}$ convergen a variables Poisson independientes, se sigue que $\sum_{n=n_0}^{n_1} X_{n,t}$ converge a una variable Poisson con parámetro mayor a 1. Luego, la probabilidad de que algún $X_n > 0$ es mayor a $Pr(Poi(1) > 0) = 1 - \frac{1}{e}$, pero eso contradice la cota uniforme dada por la segunda hipótesis.

Veamos ahora que la suma de las variables converge en distribución. Como las variables son discretas, basta ver que para todo $m \in \mathbb{N}$ vale que:

$$\lim_{t \rightarrow \infty} |Pr(\sum_n X_{n,t} = m) - Pr(Poi(\sum_n \lambda_n) = m)| = 0$$

Para todo n_0 , podemos acotar dicha diferencia por la suma de los siguientes tres términos:

$$\begin{aligned} & |Pr(\sum_n X_{n,t} = m) - Pr(\sum_{n=1}^{n_0} X_{n,t} = m)| \\ & |Pr(\sum_{n=1}^{n_0} X_{n,t} = m) - Pr(Poi(\sum_{n=1}^{n_0} \lambda_n) = m)| \\ & |Pr(Poi(\sum_{n=1}^{n_0} \lambda_n) = m) - Pr(Poi(\sum_n \lambda_n))| \end{aligned}$$

Sea $\varepsilon > 0$. El primer término lo podemos acotar por la probabilidad de que $X_{n,t}$ sea distinto de 0 para algún $n \geq n_0$ y por la segunda hipótesis, existe un n_0 tal que dicha probabilidad es menor a $\frac{\varepsilon}{3}$ para todo t . Además, como la sumatoria $\sum \lambda_n$ converge y la distribución de las variables Poisson es continua respecto al parámetro λ , se sigue que para un n_0 suficientemente grande el tercer término es menor a $\frac{\varepsilon}{3}$. Por último, para un n_0 fijo podemos hallar un t_0 tal que para todo $t > t_0$ el segundo término es menor a $\frac{\varepsilon}{3}$. Se sigue que para t suficientemente grande podemos acotar la diferencia por ε y como este es arbitrario, que tiende a 0. \square

Teorema 3.54. (Distribución asintótica de ciclos de pseudotipo dado) Sean $k \geq 3$, $r > 0$ y $m > 0$. y sea τ un pseudotipo de k -ciclos. Sea X la variable aleatoria que cuenta la cantidad de ciclos en $G(n, \frac{c}{n})$ de pseudotipo τ . Luego, cuando n tiende a infinito, X converge a una distribución Poisson y más aun, para pseudotipos distintos las variables X convergen a distribuciones independientes.

Demostración. Cada (m, r) -pseudotipo de k -ciclos está realizado por un número, posiblemente infinito, de r -tipos de k -ciclos. Luego, si X es la variable que cuenta la cantidad de k -ciclos de un (m, r) -pseudotipo dado, se cumple que $X = \sum X_\tau$, donde la suma es sobre todos los r -tipos τ que realizan nuestro pseudotipo. Sabemos que para cada τ , X_τ converge en distribución a una Poisson y que para τ_1, τ_2 distintos convergen a variables independientes. Veamos que se cumple la segunda hipótesis del lema anterior, de lo que se sigue que X converge en distribución a una Poisson.

Queremos ver que, para todo $\varepsilon > 0$, existe un conjunto finito A de los tipos tal que la probabilidad de que para algún $\tau \notin A$ se cumpla $X_\tau > 0$ es menor a ε . Sabemos que la cantidad de vecinos de un vértice dado tiende a una distribución Poisson de parámetro c . Como $Pr(Poi(c) > d)$ tiende a 0 cuando d tiende a infinito, se sigue que para todo $\varepsilon > 0$ existe un d tal que para todo k -ciclo la probabilidad de que todo elemento del ciclo tenga menos de d vecinos, que a su vez tienen menos de d vecinos, y sucesivamente por r niveles es mayor $1 - \varepsilon$. En otras palabras, existe un d tal que la probabilidad de que se realiza un r -tipo τ de más de r^d elementos es menor a ε . Como solo hay finitos árboles de menos de r^d elementos, se sigue la proposición.

Por último, como pseudotipos distintos están realizados por tipos distintos, que sabemos que son independientes, se sigue que las para pseudotipos distintos las variables convergen a Poisson independientes. \square

3.3.2. Ley de convergencia

Ya tenemos todas las herramientas necesarias para demostrar el teorema principal de este capítulo, la ley de convergencia de Lynch que mencionamos en la introducción:

Teorema 3.55. (Ley de convergencia para grafos sparse) *Sea ϕ una oración en la lógica de primer orden de grafos. Luego el límite $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi)$ existe y mas aun, se puede escribir a partir del conjunto $\{1, c\}$ con las operaciones $\{+, -, \cdot, /, exp\}$.*

Demostración. Sea ϕ una oración dada y sea $k = qr(\phi)$. Por el teorema de Gaifman, dos grafos son equivalentes si cumplen las mismas oraciones locales de tipo $(2k^2, k + 1, 7^k)$. Pero sabemos por 3.36 que cuando n tiende a infinito la probabilidad de G contenga como subgrafo a un grafo H que no sea unicyclico o un árbol tiende a 0, y sabemos por 3.47 que la probabilidad de que tenga como arbitrarias de subgrafo a un árbol dado tiende a 1. Además, sabemos que cuando n tiende a infinito todo par de ciclos esta arbitrariamente lejos. Luego, se sigue que toda la variación de comportamiento local esta dada por la aparición de grafos unicyclicos, es decir, por los pseudotipos de ciclos.

Sabemos que la cantidad de $(2 \cdot k^2, 7^k)$ pseudotipos de ciclos de tamaño menor a 7^k es finita. Llamemos a esa cantidad N . Enumeremos todos esos pseudotipos y sea Z_i la variable aleatoria que cuenta la cantidad de ciclos de un pseudotipo dado en G y sea $S_i = \min(Z_i, k)$. Por la observación anterior, cuando n tiende a infinito, la probabilidad de que todas grafos sparse con los mismos valores de S_i para todo i sean m -equivalentes tiende a 1. Luego, para todo $\bar{a} \in \{0, 1, \dots, k - 1, k\}^N$ la probabilidad límite $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi | S_1 = a_1, \dots, S_N = a_N)$ es 0 o 1.

$$Pr(G(n, \frac{c}{n}) \models \phi) = \left(\sum_{\bar{a}} p_{\bar{a}} \cdot Pr(S_1 = a_1, \dots, S_N = a_N) \right) + \varepsilon(n)$$

Por ultimo, por el teorema 3.54 las Z_i convergen en distribución a variables Poisson independientes, con un parámetro que se puede escribir a partir del conjunto $\{1, c\}$ con las operaciones $\{+, -, \cdot, /, exp\}$. Se sigue que cada termino de forma $Pr(S_1 = a_1, \dots, S_N = a_N)$, que se pueden escribir como intersección de eventos de la forma $Z_i = a$ o $Z_i \geq l$, tiene una probabilidad límite que se se puede escribir a partir del conjunto $\{1, c\}$ con las operaciones $\{+, -, \cdot, /, exp\}$. Como la sumatoria es finita, se puede intercambiar con el límite y por tanto nos queda la siguiente igualdad, de la que se sigue el teorema:

$$\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = \sum_{\bar{a}} p_{\bar{a}} \cdot [\lim_{n \rightarrow \infty} Pr(S_1 = a_1, \dots, S_N = a_N)]$$

□

El teorema anterior nos permite demostrar varios resultados útiles sobre el comportamiento asintótico de oraciones sobre grafos sparse.

Corolario 3.56. *Sea ϕ una oración en la lógica de primer orden de grafos y sea $f_{\phi}(c) = \lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi)$. Luego la función f_{ϕ} es suave.*

El siguiente teorema es una suerte de análogo del resultado demostrado en el capítulo anterior de que el límite $\lim_{n \rightarrow \infty} Pr(G(n, p) \models \phi)$ no depende de p . Aunque el límite $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi)$ sí depende de c , lo que muestra el teorema es que no varía **cualitativamente** con c ,

Teorema 3.57. *Sea ϕ una oración en la lógica de primer orden de grafos. Luego:*

- *Si para algún $c > 0$ se cumple que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1$, se sigue que lo mismo pasa para todo $c > 0$.*
- *Si para algún $c > 0$ se cumple que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) \in (0, 1)$, se sigue que lo mismo pasa para todo $c > 0$*
- *Si para algún $c > 0$ se cumple que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 0$, se sigue que lo mismo pasa para todo $c > 0$*

Demostración. Para todo pseudotipo de ciclo τ , sabemos que la cantidad de k -ciclos de pseudotipo de τ en $G(n, \frac{c}{n})$ tiende a una Poisson cuyo parámetro depende de c , pero es fácil ver que si no es 0 en algún c no es 0 en ningún c . En la demostración de la ley de convergencia para grafos sparse, llegamos a que:

$$\lim_{n \rightarrow \infty} (G(n, \frac{c}{n}) \models \phi) = \sum_{\bar{a}} p_{\bar{a}} \cdot [\lim_{n \rightarrow \infty} Pr(S_1 = a_1, \dots, S_N = a_N)]$$

Donde S_i es el máximo entre m y la variable que cuenta la cantidad de ciclos de pseudotipo i . Por la observación anterior, los términos $Pr(S_1 = a_1, \dots, S_N = a_N)$ no son 0 ni 1 para ningún c . Luego se sigue que si para algún c vale que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1$ todos los términos $p_{\bar{a}}$ son iguales a 1, y luego para todo c vale que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1$.

Similarmente, si para algún c vale que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 0$ todos los términos $p_{\bar{a}}$ son iguales a 0, y luego para todo c vale que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 0$. \square

A diferencia de la ley 0-1 de Fagin para grafos, la ley de convergencia para grafos muy sparse **no** se puede extender a la lógica infinitaria. Lynch atribuye[29] dicho resultado a Ehrenfeucht, pero no pudimos encontrar dicha referencia, así que damos una demostración original.

Proposición 3.58. *Sea $c > 1$. Existe una oración ϕ en la lógica infinitaria $\mathcal{L}_{w, \infty}^4$ tal que $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi)$ **no** existe.*

Demostración. Dado $r > 0$, sea ψ_r la siguiente oración, que establece la existencia de dos 3-ciclos a distancia menor a r :

$$\psi_r = [\exists x \exists y : (\exists z \exists w : x \sim z \wedge x \sim w \sim z \sim w) \wedge (\exists z \exists w : y \sim z \wedge y \sim w \sim z \sim w) \wedge (d(x, y) < r)]$$

Como sabemos que $d(x, y) < r$ se puede expresar con una fórmula con variables libres x e y y dos variables cuantificadas, luego se sigue que la oración ψ_r pertenece a FO^4 . Vamos a considerar la siguiente oración $\phi \in \mathcal{L}_{w, \infty}^4$:

$$\phi = \bigvee_k (\neg \psi_{a_{2k}} \wedge \psi_{a_{2k+1}})$$

Donde a_k es la sucesión de naturales definida por $a_{k+1} = (2 \cdot c)^{a_k}$. El propo-sitivo de esa secuencia es que por el teorema 3.45, si $n = a_{k+1}$ la probabilidad

de que $G(n, \frac{c}{n})$ tenga dos 3-ciclos a distancia menor a a_k es 0. Luego, se sigue que si $n = a_{2k+1}$:

$$\lim_{k \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 0$$

Por otro lado, sabemos por el teorema 3.2 que asintoticamente el grafo $G(n, \frac{c}{n})$ tiene un componente conexo gigante de tamaño $d \cdot n$ para algún $d < 1$. Luego, la probabilidad de que existan dos 3-ciclos conectados se puede acotar por la probabilidad de que existan dos 3-ciclos dentro del componente conexo, evento que tiene probabilidad $1 - e^{-d \cdot c}$. Si $n = a_{2k}$, como casi seguro no hay 3-ciclos a distancia menor a a_{2k-1} , el grafo cumple ϕ si y solo si hay dos ciclos conectados, que por el argumento anterior sabemos que tiene probabilidad mayor a $1 - e^{-d \cdot c}$. Luego, si $n = a_{2k}$:

$$\lim_{k \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1 - e^{-d \cdot c}$$

Por tanto, la oración ϕ no tiene probabilidad limite. □

Ahora pasaremos a investigar los aspectos efectivos de la ley de convergencia para grafos sparse. Empezaremos con el de encontrar cotas para la velocidad de convergencia de dicha ley, tema que al igual de el de encontrar cotas para la ley 0-1, no recibió atención en la literatura. Podemos demostrar que la cota inferior es en este caso super-exponencial:

Proposición 3.59. (Cota inferior para la convergencia en grafos muy sparse)
Sea $c > 0$. Para todo k , existe una oración ϕ_k en la lógica de primer orden de grafos de longitud $O(k)$ tal que:

$$\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi_k) = 1$$

Pero si $n = 2^{2^k}$ se cumple que:

$$\lim_{k \rightarrow \infty} (Pr(G(n, \frac{c}{n}) \models \phi_k) = 0$$

Demostración. La oración ϕ_k sera de la forma $\exists x : \psi_k(x)$, donde $\psi_k(x)$ es la siguiente fórmula, que expresa que x tiene $\sqrt{k} + 1$ vecinos, los cuales tienen a su vez $\sqrt{k} + 1$ vecinos y asi sucesivamente por 2^k niveles.

$$\psi_k(x) = [(\forall y \ d(x, y) < 2^k \implies (\exists z_1 \dots \exists z_{\sqrt{k}+1} \bigwedge_i y \sim z_i \wedge \bigwedge_{i \neq j} z_i \neq z_j)]$$

Como vimos anteriormente que $d(x, y) < 2^k$ se puede expresar con una oración de longitud $O(k)$ y la segunda parte de la oración tiene longitud del orden de $(\sqrt{k} + 1)^2 = O(k)$, luego se sigue que ϕ_k tiene longitud $O(k)$. Por otro lado, es fácil ver que la oración es satisfecha si el grafo contiene como subgrafo un árbol de altura 2^k con \sqrt{k} vecinos en todos sus vértices salvo las hojas. Como por el corolario todo árbol aparece asintoticamente como subgrafo de $G(n, \frac{c}{n})$, se sigue que la oración tiene probabilidad límite 1.

Por otro lado, la probabilidad de que un vértice dado tenga mas de \sqrt{k} vecinos es, para $\sqrt{k} > 2c$:

$$Pr(Poi(c) > \sqrt{k}) = \sum_{m=\sqrt{k}}^{\infty} e^{-c} \cdot \frac{c^m}{m!} < \sum_{m=\sqrt{k}}^{\infty} e^{-c} \cdot c^{2c} \cdot \left(\frac{1}{2}\right)^{m-2c} = e^{-c} \cdot c^{2c} \cdot \sum_{s=\sqrt{k}-2c}^{\infty} \frac{1}{2^s}$$

$$Pr(Poi(c) > \sqrt{k}) < \frac{e^{-c} \cdot c^{2c}}{2^{\sqrt{k}-2c}} = \frac{C}{2^{\sqrt{k}}}$$

Ahora, la oración ϕ_k requiere no solo que x tenga al menos $\sqrt{k}+1$ vecinos, si no que esos vecinos también tenga al menos $\sqrt{k}+1$ vecinos, y así por 2^k niveles. Como cada nivel del árbol tiene al menos \sqrt{k} vértices mas que el anterior, la oración fuerza a mas de $(\sqrt{k})^{2^k}$ vértices a tener mas de \sqrt{k} vecinos. Luego, la probabilidad de que un vértice dado cumpla $\psi_k(x)$ es:

$$p < (Pr(Poi(c) > \sqrt{k}))^{\sqrt{k}^{2^k}} < C^{\sqrt{k}^{2^k}} \cdot (2^{-\sqrt{k}})^{\sqrt{k}^{2^k}}$$

Reescribiendo los términos, nos queda que:

$$p < 2^{-k^{(2^k)} + \log_2(c) \cdot (\sqrt{k})^{2^k}}$$

Finalmente, si $n = n = 2^{2^{2^k}}$ por la desigualdad de Markov tenemos:

$$Pr(G(n, \frac{c}{n}) \models \phi) < p \cdot 2^{2^{2^k}} = 2^{-k^{(2^k)} + \log_2(c) \cdot (\sqrt{k})^{2^k} + 2^{2^k}}$$

En el exponente hay tres términos, dos positivos y uno negativo, pero es evidente que el negativo crece mucho mas rápido que los otros dos. Luego, se sigue que la probabilidad de que $G(n, \frac{c}{n}) \models \phi$ tiende a 0 cuando n tiende a infinito. \square

Observación 3.60. Usando cotas superiores de la cola de la distribución Poisson, se puede ver que la oración ϕ_k definida en la proposición anterior se cumple asintóticamente para:

$$n = 2^{2^{2^k \cdot \ln(k)}}$$

Veamos ahora los aspectos computacionales de la ley de convergencia. Lynch observa en su artículo que el problema de determinar es decidible: Como vimos, lo que determina el valor asintótico de es la aparición de ciclos con ciertos arboles alrededor y es consecuencia de un teorema muy profundo de decidibilidad de teorías, el teorema del Arbol de Rabin [17], que la teoría de arboles finitos es decidible, luego existe un algoritmo que puede calcular explícitamente todos los términos de la fórmula dada en la demostración 3.55. Ese algoritmo, sin embargo, es de complejidad no elementaria ¿Se podrá encontrar uno mas eficiente?

No tenemos la respuesta, pero podemos al menos probar una cota inferior en la complejidad del problema:

Teorema 3.61. (Cota inferior para la complejidad de la ley de Lynch) El problema de determinar, dada oración ϕ en la lógica de primer orden de grafos, si $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1$ es NEXPTIME-hard.

Demostración. Vamos a reducir el problema de decidir la validez de una oración en un lenguaje monádico, que vimos en 1.40 que es NEXPTIME-completo, a la de decidir si una oración se cumple en casi todos los grafos muy sparse. Para $n \geq 1$, sea $\phi_n(x)$ la fórmula:

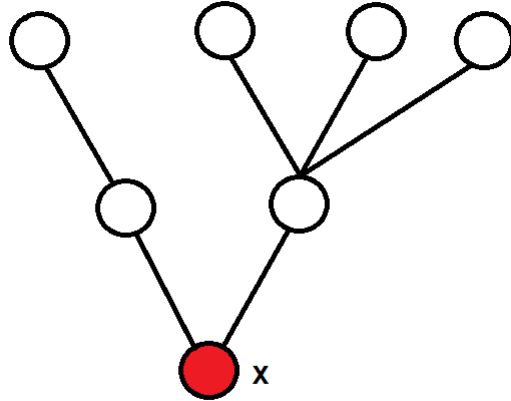
$$\phi_n(x) = [\exists y : x \sim y \wedge (\exists z_1 (\dots) \exists z_n : (\bigwedge_i y \sim z_i) \wedge (\bigwedge_{i,j} z_i \neq z_j) \wedge (\forall w (w \sim y) \rightarrow (w = x \vee \bigvee_i w = z_i)))]$$

Es decir, $\psi_n(x)$ es válida si y solo si x tiene un vecino y con exactamente otros n vecinos. Ahora, dada oración ϕ en un lenguaje que consta unicamente de predicados monádicos P_1, \dots, P_n podemos construir una oración ϕ^* en el lenguaje de grafos de cada ocurrencia de $P_i(x)$ con $\psi_i(x)$ y podemos construir la oración $\bar{\phi}$ como:

$$\bar{\phi} = \exists x : \phi^*_{\{y:y \sim x\}}$$

Es decir, $\bar{\phi}$ es cierta si existe algún x tal que sus vecinos cumplen ϕ^* . Notemos que si la longitud de ϕ es n usara como mucho n predicados y la longitud de ψ_n es $O(n^2)$, luego, la longitud de $\bar{\phi}$ sera $O(n^3)$, luego la reducción es polinómica.

Esta claro que si **algún** grafo G cumple $\bar{\phi}$ entonces ϕ es satisfacible, en particular, si casi todo grafo $G(n, \frac{c}{n})$ la cumple entonces es satisfacible. Veamos que vale la vuelta. Es fácil ver que para todo subconjunto de $\{1, \dots, n\}$ existe un árbol de altura 3 tal que su raíz cumple $\bar{\psi}_i$ si y solo i . Por ejemplo, la siguiente imagen muestra un x que cumple ψ_1 y ψ_3 pero no ψ_2



Luego, si ϕ es satisfacible se sigue que existe un árbol T de altura 4 tal que $T \models \bar{\phi}$. Como sabemos que los grafos $G(n, \frac{c}{n})$ contienen arbitrarias copias aisladas de cada árbol cuando n tiende a infinito, se sigue que si ϕ es satisfacible la probabilidad límite de $\bar{\phi}$ es 1. □

Observación 3.62. Como sabemos por el teorema de jerarquía temporal que $P \neq EXPTIME$, se sigue que el problema de determinar, dada oración ϕ en la lógica de primer orden de grafos, si $\lim_{n \rightarrow \infty} Pr(G(n, \frac{c}{n}) \models \phi) = 1$ no puede ser resuelto en tiempo polinómico.

3.4. Espacios con funciones

Uno de los problemas mas celebres y antiguos de la Teoría de Probabilidades, presente en todo curso introductorio de la materia, es el llamado "Problema de los sombreros", introducido por Montmort en su *Essay d'analyse sur les jeux de hazard*, uno de los primeros tratados matemáticos de probabilidad. Su formulación mas común es la siguiente:

Un grupo de n personas entra en un restaurante y dejan los sombreros en el guardarropa. El encargado del guardarropa es distraído y, al irse, redistribuye los sombreros a los comensales al azar. ¿Cuál es la probabilidad de que ninguna persona reciba su sombrero correcto?

Una formulación mas abstracta seria: Dada una permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ elegida de forma uniforme, ¿Cual es la probabilidad de que no tenga ningún punto fijo?.

Teorema 3.63. *Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación elegida de forma uniforme. Luego, la probabilidad de que σ no tenga ningún punto fijo tiende a $1 - \frac{1}{e}$ cuando n tiende a infinito.*

Demostración. Veamos que la probabilidad de que σ tenga al menos un punto fijo tiende a $\frac{1}{e}$. Sea A_i el evento $\sigma(i) = i$. Por el principio de inclusión-exclusión tenemos:

$$Pr(\bigcup A_i) = \sum Pr(A_i) - \sum Pr(A_i \cap A_j) + \sum Pr(A_i \cap A_j \cap A_k) - \dots$$

Ahora, dados i_1, \dots, i_k distintos hay $(n - k)!$ permutaciones $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ que dejan fijos a i_1, \dots, i_k , luego la probabilidad de $A_{i_1} \cap \dots \cap A_{i_k}$ es $\frac{n!}{(n-k)!}$. Nos queda entonces:

$$Pr(\bigcup A_i) = \binom{n}{1} \cdot \frac{(n-1)!}{n!} - \binom{n}{2} \cdot \frac{(n-2)!}{n!} + \binom{n}{3} \cdot \frac{(n-3)!}{n!} - \dots$$

$$Pr(\bigcup A_i) = \frac{1}{1!} - \frac{1}{2!} + \frac{1}{3!} - \dots = \sum_{a=0}^n (-1)^k \cdot \frac{1}{k!} \rightarrow \frac{1}{e}$$

□

Notemos que el problema anterior puede expresarse como una oración de primer orden sobre el lenguaje con una permutación $\mathcal{L} = \{\sigma\}$, de la forma $\phi = \forall x : \neg(\sigma(x) = x)$, lo que da lugar a la pregunta si es posible realizar un estudio general de los problemas asintóticos sobre dicho lenguaje. La respuesta es positiva y podemos demostrarla rápidamente con las herramientas que desarrollamos. La clave es el siguiente teorema, que generaliza ampliamente el resultado anterior:

Teorema 3.64. (Distribución de ciclos de una permutación aleatoria) *Sea $k \geq 1$ y sea $X_{k,n}$ la variable que cuenta la cantidad de k -ciclos de una permutación $\sigma \in S_n$ elegida de manera uniforme. Cuando n tiende a infinito, $X_{k,n}$ converge en distribución a una variable Poisson de parámetro $\lambda = \frac{1}{k}$ y mas aun, para distintos valores de k las variables convergen a variables independientes entre si.*

Demostración. Para todo $S \subset \{1, \dots, n\}$ denotemos con A_S el evento de que el conjunto S sea un ciclo de σ . Por definición, $X_{k,n} = \sum 1_{A_S}$, donde la sumatoria es sobre todos los S tal que $|S| = k$. Veamos que se cumplen las hipótesis de la ley de convergencia Poisson IV, de lo que se sigue el teorema. Para empezar, como los ciclos de σ son siempre disjuntos, es imposible que ocurra A_S y A_T si S y T son subconjuntos distintos pero no disjuntos, lo que corrobora la segunda hipótesis.

Para ver la primer hipótesis, consideremos s_1 subconjuntos de tamaño k_1, \dots, s_m subconjuntos de tamaño k_m y veamos que si denotamos con p la probabilidad de que todos formen un ciclo vale que $\lim_{n \rightarrow \infty} p \cdot n^{\sum k_i s_i} = \lambda_1^{k_1} \cdot (\dots) \cdot \lambda_m$. Dado un conjunto de tamaño r , hay $(r-1)!$ permutaciones en las que todo el conjunto forma un solo ciclo, luego tenemos:

$$p = \prod (k_i - 1)!^{s_i} \cdot \frac{(n - \sum k_i s_i)!}{n!}$$

$$p \cdot n^{\sum k_i s_i} = \prod (k_i - 1)!^{s_i} \cdot \frac{(n - \sum k_i s_i)!}{n!} \cdot n^{\sum k_i s_i} \rightarrow \prod (k_i - 1)!^{s_i}$$

Se sigue por la ley de convergencia Poisson IV que la variable $X_{k,n}$ converge en distribución a una Poisson de parámetro:

$$\lambda = \frac{(k-1)!}{k!} = \frac{1}{k}$$

□

Teorema 3.65. (Ley de convergencia para permutaciones aleatorias) Sea ϕ una oración sobre el lenguaje $\mathcal{L} = \{\sigma\}$ con una única función unaria. Luego la probabilidad de que una permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ elegida de forma uniforme cumpla ϕ converge cuando n tiende a infinito y mas aun, el límite se puede escribir a partir del conjunto $\{0, 1\}$ usando las operaciones $\{+, -, \cdot, /, e^x\}$

Demostración. Sea A un conjunto equipado con una permutación. Por el teorema de Hanf, la clase de k -equivalencia queda determinada por la aparición de 3^k -tipos. Pero en un conjunto con una permutación, los únicos tipos posibles son los ciclos, luego, la clase de k -equivalencia queda determinado por cuantos (hasta k) ciclos de tamaño r hay para cada $r \leq 3^k$. Luego, si denotamos con X_i la variable aleatoria que cuenta la cantidad de i -ciclos, se sigue que:

$$Pr((A, \sigma) \models \phi) = \sum_{\bar{a}} Pr(X_1 = a_1) \cdot (\dots) \cdot Pr(X_{3^r} = a_{3^r}) \cdot s_{\bar{a}}$$

Donde la sumatoria es sobre todas las tuplas $\bar{a} = a_1, \dots, a_{3^r}$ tomando valores en $\{0, 1, 2, \dots, k+\}$ y el termino $s_{\bar{a}} \in \{0, 1\}$ indica si esa clase de k -equivalencia cumple ϕ o no. Como sabemos por el teorema anterior que todos los términos X_i convergen a variables Poisson independientes se sigue que la probabilidad límite existe y como todas esas Poisson tienen parámetro racional, por la fórmula de densidad puntual de las variables Poisson se sigue que el límite se puede escribir a partir del conjunto $\{0, 1\}$ usando las operaciones $\{+, -, \cdot, /, exp\}$. □

Lynch demostró en [28] que se cumple un resultado mucho mas general, considerando espacios que cuentan con una cantidad arbitraria de permutaciones y funciones unarias:

Teorema 3.66. (Ley de convergencia para lenguajes con funciones)[28] Sea $\mathcal{L} = \{f_1, \dots, f_k, \sigma_1, \dots, \sigma_s, P_1, \dots, P_t\}$ un lenguaje consistente en un número de funciones unarias f , permutaciones σ y predicados unarios P . Luego si $A(n)$ representa una \mathcal{L} -estructura de n elementos elegida de forma uniforme, se cumple que el límite $\lim_{n \rightarrow \infty} Pr(A(n) \models \phi)$ existe y, mas aun, se puede escribir a partir del conjunto $\{0, 1\}$ usando las operaciones $\{+, \cdot, -, /, e^x\}$.

La idea de la demostración es muy similar en espíritu a la de la ley de convergencia para grafos muy sparse y la demostración previa para espacios con una única permutación: Consiste en usar los resultados de localidad de la lógica de primer orden para reducir el problema a clasificar el comportamiento local de la estructura y ver que el comportamiento local de la estructura esta determinado por la presencia de ciclos, que convergen en distribución a variables Poisson independientes. La aparición de funciones que no son permutaciones, sin embargo, nos obliga a considerar no solo los ciclos sino la estructura alrededor de los ciclos, que complica la demostración.

Resultados muy similares se demostraron también para el caso de grafos d -regulares uniformes. En general, parece que el mismo fenómeno ocurre en toda familia de estructuras aleatorias en la que dos objetos en una estructura de n elementos están relacionados con probabilidad $O(\frac{1}{n})$. Queda abierta la pregunta de si existe una formulación general que englobe todos esos casos. Lynch considera que la hay, pero es pesimista respecto a su valor: "We expect that there is a general theorem that implies all of these results as corollaries, however, it may be so awkward to state that its value would be questionable"[29].

Capítulo 4

No convergencia y no decidibilidad en grafos ordenados y otras estructuras

4.1. Grafos ordenados

4.1.1. Introducción

Este capítulo es una suerte de dual al resto de la tesis. En los capítulos anteriores analizamos la probabilidad asintótica de oraciones en distintas estructuras aleatorias y demostramos una serie de resultados positivos que clasifican su comportamiento, leyes de convergencia y la existencia de algoritmos para determinarla, y establecimos que esos resultados nos hablaban de las limitaciones de la lógica de primer orden en dichas estructuras.

En este capítulo demostraremos casi exclusivamente resultados **negativos** acerca de la probabilidad asintótica de oraciones en nuevos tipos de estructuras, mostrando que su probabilidad limite en general no existe, que el problema de determinar su limite no es decidible y que su comportamiento es en general inclasificable.

El primer tipo de estructura que analizaremos son los **grafos ordenados**. Denotaremos con $G_{<}(n)$ a un grafo aleatorio elegido de manera uniforme entre todos los grafos de $\{1, \dots, n\}$, equipado con el orden usual en dicho conjunto. Es fácil ver que la ley 0-1 de Fagin no se puede extender a este tipo de estructuras: Si ϕ es la oración dada por:

$$\phi = \forall x \forall y (\forall z x \leq z \wedge z \leq y) \implies (x \sim y)]$$

que afirma que el los vértices 1 y n están conectados, es fácil ver que la probabilidad asintótica de $G(n) \models \phi$ es $\frac{1}{2}$.

En [33], Shelah, Compton y Henson demostraron que no solo falla la ley 0-1, sino que tampoco hay una ley de convergencia y que el problema de determinar la probabilidad limite de una oración o si el limite existe es indecidible. El método que usaron para la demostración fue usar el lenguaje con orden para codificar la existencia de cuerpos finitos dentro del grafo y aprovechar la estructura de los cuerpos para construir una oración sin probabilidad limite. Dicho método, que los autores denominaron **aritmización**, es muy versátil y lo usaremos extensamente mas adelante para probar una serie de resultados

negativos. Sin embargo, las oraciones construidas de esta manera son largas y “artificiales”, lo que deja abierta la pregunta de si hay oraciones mas naturales que no tengan probabilidad limite.

La respuesta es afirmativa. En [37] Spencer muestra como un problema probabilístico, que el denomina “La maratón de danza”, se puede codificar con una oración muy simple en el lenguaje de grafos ordenados, que no tiene probabilidad asintótica.

4.1.2. La maratón de danza

Definición 4.1. La maratón de danza tiene n participantes y dura una cantidad entera de rondas, hasta que todos los participantes abandonan. En cada ronda, cada participante abandona el concurso con probabilidad $\frac{1}{2}$, de manera independiente al resto y a las rondas anteriores. Un participante se declara **ganador** si en cierta ronda es el único que se mantiene en el concurso. Notemos que es posible que no haya ningún ganador, por ejemplo, si los últimos dos participantes abandonan al mismo tiempo.

Observación 4.2. Una definición mas abstracta seria la siguiente: Sean, para $i \in \{1, \dots, n\}$, $t \geq 0$, $X_{i,t}$ variables Bernoulli independientes con $p = \frac{1}{2}$. Un $i \in \{1, \dots, n\}$ se declara ganador si existe un t_0 tal que para todo $t \leq t_0$ vale que $X_{i,t} = 1$ y para todo $j \neq i$ existe un $t < t_0$ tal que $X_{j,t} = 0$.

Proposición 4.3. Denotemos con $p_{MD}(n)$ la probabilidad de que haya un ganador en la maratón de danza de n participantes. Vale la siguiente fórmula:

$$p_{MD}(n) = n \cdot \sum_{s=1}^{\infty} 2^{-s-1} \cdot (1 - 2^{-s})^{n-1}$$

Demostración. Dados $i \in \{1, \dots, n\}$ y $t \in \mathbb{N}$ denotemos con $S_{i,t}$ el evento “El participante i es el ganador del concurso y abandona en la ronda t ”. Notemos que los eventos $S_{i,t}$ son disjuntos. Luego tenemos:

$$p_{MD}(n) = Pr\left(\bigcup_{t,i} S_{t,i}\right) = \sum_{t,i} Pr(S_{t,i})$$

Y por simetría, este ultimo término es igual a $n \cdot \sum_t Pr(S_{t,1})$. Por otro lado, la probabilidad de $S_{t,i}$ es igual a la probabilidad de que el primer participante se mantenga las primeras t rondas y abandona en la $t + 1$ -ésima, multiplicado por la probabilidad de que todos los otros participantes abandonen antes de la ronda t , es decir, tenemos:

$$Pr(S_{1,t}) = 2^{-t-1} \cdot (1 - 2^{-t})^{n-1}$$

Juntando todo:

$$p_{MD}(n) = n \cdot \sum Pr(S_{1,t}) = n \cdot \sum 2^{-t-1} \cdot (1 - 2^{-t})^{n-1}$$

□

Proposición 4.4. [37] El límite $\lim_{n \rightarrow \infty} p_{MD}(n)$ no existe.

Demostración. Mostraremos que $\lim_{k \rightarrow \infty} p_{MD}(\alpha \cdot 2^k)$ adopta distintos valores según el $\alpha \in \mathbb{R}^+$, de lo que se sigue que $\lim_{n \rightarrow \infty} p_{MD}(n)$ no existe. Sea $n = \alpha \cdot 2^k$. Tenemos:

$$\begin{aligned} p_{MD}(n) &= n \cdot \sum_{s=1}^{\infty} 2^{-s-1} \cdot (1-2^{-s})^{n-1} = \alpha \cdot 2^k \cdot \sum_{s=1}^{\infty} 2^{-s-1} \cdot (1-2^{-s})^{\alpha \cdot 2^k - 1} \\ &= \sum_{s=1}^{\infty} \alpha \cdot 2^{k-s-1} \cdot \frac{(1-2^{-s})^{\alpha \cdot 2^k}}{(1-2^{-s})} \end{aligned}$$

Cambiando la variable de sumatoria de s a $t = s - k$ nos queda:

$$p_{MD}(n) = \sum_{t=-k}^{\infty} \alpha \cdot 2^{-t-1} \cdot \frac{(1-2^{-k-t})^{\alpha \cdot 2^k}}{(1-2^{-k-t})}$$

Analizamos la sumatoria. Para todo $t > 0$ y todo k , el término se puede acotar por $\alpha \cdot 2^{-t}$. Por otro lado, usando la desigualdad $(1-x) < e^{-x}$, para $t < 0$ podemos acotar el término por $\alpha \cdot 2^{-t-1} \cdot e^{-\alpha 2^{-t}}$, que decrece superexponencialmente. Como ambas cotas son sumables, se sigue por el teorema de la convergencia dominada que podemos intercambiar el límite con la sumatoria, es decir:

$$\begin{aligned} \lim_{k \rightarrow \infty} p_{MD}(n) &= \lim_{k \rightarrow \infty} \sum_{t=-k}^{\infty} \alpha \cdot 2^{-t-1} \cdot \frac{(1-2^{-k-t})^{\alpha \cdot 2^k}}{(1-2^{-k-t})} \\ &= \sum_{t=-\infty}^{\infty} \lim_{k \rightarrow \infty} \left[\alpha \cdot 2^{-t-1} \cdot \frac{(1-2^{-k-t})^{\alpha \cdot 2^k}}{(1-2^{-k-t})} \right] = \sum_{t=-\infty}^{\infty} \alpha \cdot 2^{-t-1} \cdot e^{-\alpha 2^{-t}} = g(\alpha) \end{aligned}$$

Se puede ver que $g(\alpha)$ varía según el α , aunque varía tan poco que aparenta ser constante. Análisis numérico nos revela que para $\alpha \in (1, 2)$, g alcanza un máximo de 0.721354 cerca de $\alpha = 1.81$ y un mínimo de 0.721340 cerca de $\alpha = 1.28$. \square

David Speyer ofrece el siguiente argumento heurístico de porque es razonable esperar que $p_{MD}(n)$ no tenga probabilidad límite [38]: Supongamos que empezamos con dos maratones, una con 2^n personas y otra con $1.5 \cdot 2^n$. Tras la primera ronda, la primera maratón tendrá alrededor de 2^{n-1} personas y la segunda alrededor de $1.5 \cdot 2^{n-1}$. Repitiendo el proceso suficientes veces, esperaríamos que en algún momento la primera maratón tenga alrededor de 8 personas y la segunda alrededor de $1.5 \cdot 8 = 12$ y no hay razón para esperar que $p_{MD}(8) = p_{MD}(12)$.

Ahora vamos a definir una variante del problema anterior, que llamaremos la maratón de danza acotada.

Definición 4.5. La **maratón de danza acotada** cuenta con participantes $1, \dots, n$ y dura una cantidad entera de rondas. La dinámica es igual a la anterior, pero con la condición extra de que el i -ésimo participante abandona el concurso en la ronda i con probabilidad 1, si es que llegó tan lejos. Notemos que se sigue que este concurso puede tener como mucho n rondas. Denotemos con $p_{MDA}(n)$ la probabilidad de que la maratón de danza acotada de n participantes tenga un ganador.

Proposición 4.6. Sea $p_{MDA}(n)$ la probabilidad de que la maratón de danza acotada tenga un ganador. Luego vale:

$$\lim_{n \rightarrow \infty} [p_{MD}(n) - p_{MDA}(n)] = 0$$

Demostración. Empecemos demostrando que la probabilidad de que en una maratón de danza de n participantes termine en menos de \sqrt{n} rondas tiende a 1 cuando $n \rightarrow \infty$. Eso es equivalente a decir que todos los participantes abandonan la competición antes de la ronda \sqrt{n} , evento que tiene probabilidad:

$$(1 - 2^{-\sqrt{n}})^n = \left(1 - \frac{2^{-\sqrt{n}} \cdot n}{n}\right)^n \rightarrow e^{-2^{-\sqrt{n}} \cdot n}$$

Como $\lim_{n \rightarrow \infty} 2^{-\sqrt{n}} \cdot n = 0$, se sigue que $\lim_{n \rightarrow \infty} e^{-2^{-\sqrt{n}} \cdot n} = 1$. Luego, podemos suponer sin pérdida de generalidad que estamos en una maratón de danza de n participantes que termina en menos de \sqrt{n} rondas. La única diferencia entre la maratón estándar y la maratón acotada es la dinámica para los primeros \sqrt{n} participantes (ya que en la maratón acotada el participante i está obligado a irse en la ronda i), pero por simetría, la probabilidad de que uno de estos sea ganador es de $\frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$, que tiende a 0 cuando $n \rightarrow \infty$, luego con probabilidad 1 el resultado de la maratón acotada y la maratón estándar será el mismo. \square

Corolario 4.7. El límite $\lim_{n \rightarrow \infty} p_{MDA}(n)$ no existe.

Veamos ahora que la maratón de danza acotada se puede expresar como una fórmula de primer orden sobre las estructuras analizadas anteriormente:

Proposición 4.8. Hay una oración ϕ en lenguaje de grafos ordenados que codifica la maratón de danza acotada, es decir, tal que $Pr(G(n, \leq) \models \phi) = p_{MDA}(n)$.

Demostración. Sea ϕ la oración dada por:

$$\phi = [\exists x : \exists y : (\forall w < y \ x \sim w) \wedge (\forall z(\forall w < y \ z \sim w) \rightarrow (z = x))]$$

Dicha oración captura exactamente la dinámica de la maratón de danza acotada: el evento "el participante i pasa la ronda j ", con $i < j$ se corresponde con el evento $i \sim j$, que tienen probabilidad $\frac{1}{2}$ y son independientes entre sí, y el hecho que el participante i abandona siempre en la ronda i se corresponde con el hecho de que en los grafos vale siempre que $i \not\sim i$. La fórmula expresa la existencia de un x que se mantiene en la maratón más rondas que cualquier otro y , que es exactamente el evento "La competencia tiene un ganador" \square

Corolario 4.9. Existe una oración ϕ sobre el lenguaje de grafos ordenados tal que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi)$ no existe.

4.1.3. El método de los cerebros de Boltzmann

A principios del siglo pasado, antes del surgimiento de la teoría del Big Bang y de la cosmología moderna, los físicos se encontraban ante el problema de creer simultáneamente tres proposiciones aparentemente contradictorias. La primera, que el universo era temporalmente ilimitado y existía desde siempre, la segunda, que el universo se encontraba en un estado de baja entropía y la tercera, que la entropía tiende siempre a aumentar. Se sigue de la primera y tercera proposición que el universo debería estar en un estado de entropía máximo, que contradice explícitamente la segunda.

Boltzmann, padre de la mecánica estadística y uno de los físicos más brillantes de su época, propuso una atrevida solución. Recordó que la segunda ley de la termodinámica era una ley estadística, no absoluta, y que aunque la entropía casi siempre aumentaba puede reducirse por razones puramente estocásticas. Luego, es posible que el universo haya existido desde siempre en un estado de muy alta entropía y solo recientemente nuestra región del universo haya experimentado una súbita caída de entropía por razones puramente azarosas.

Dicha solución no convenció a demasiada gente. Primero, porque las probabilidades de que algo así ocurra son tan astronómicamente bajas que la palabra milagro no es capaz de hacerle justicia. Pero además, porque razonamientos estadísticos de esa naturaleza pueden llevar a conclusiones absurdas y que ponen incluso en cuestionamiento la capacidad de hacer ciencia empírica. Ya que antes que pensar que regiones enteras del universo experimentaron una súbita caída de entropía, deberíamos considerar un "milagro" de carácter mucho menor: que del caos cósmico surja fugazmente una configuración de moléculas consciente, que crea (¡incorrectamente!) vivir en un mundo de baja entropía. Esos ficticios observadores fueron llamados **Cerebros de Boltzmann**.

El método de aritmetización de Shelah y Compton, que nosotros decidimos rebautizar como el método de los cerebros de Boltzmann, funciona bajo una idea muy parecida. Vamos a ver que en grafos aleatorios suficientemente grandes aparecen por puro azar estructuras complejas (nuestros "cerebros de Boltzmann") y que la lógica de primer orden con orden es capaz de detectarlas, lo que nos permitirá crear oraciones sin probabilidad límite, entre otras cosas. Como mencionamos antes, Shelah y Compton trabajaron con cuerpos finitos, pero para obtener resultados más generales nosotros codificaremos la estructura más general posible: Máquinas de Turing probabilísticas.

Empecemos con algunas definiciones:

Definición 4.10. Una **Máquina de Turing probabilística** M es una tupla $(Q, \Sigma, \sigma, q_0, b, F)$ donde Q es el conjunto de estados de la máquina, Σ es el alfabeto, $q_0 \in Q$ es el estado inicial, $b \in \Sigma$ es el símbolo inicial (blanco de la cinta), $F \subset Q$ es el conjunto de estados de aceptación y $\sigma : Q \times \Sigma \times \{0, 1\} \rightarrow Q \times \Sigma \times \{L, R\}$ es la función de transición. El comportamiento de una máquina de Turing probabilística es análogo al de una máquina de Turing pero en vez de estar fijo se convierte en una **variable aleatoria**: En cada paso la máquina recibe un bit aleatorio $b \in \{0, 1\}$ elegido con una distribución Bernoulli independiente a los bits anteriores, y lo usa en conjunto de su estado y el símbolo de la cinta para determinar su comportamiento.

Definición 4.11. Dada una máquina de Turing $M = (Q, \Sigma, \sigma, q_0, b, F)$ decimos que $f : \{1, \dots, n\}^2 \rightarrow \Sigma \times (Q \cup \{*\})$ simula a M si, considerando el desarrollo de la máquina M a partir de la cinta vacía se cumple que:

- $f(t, i) = (s, *)$ si y solo si en el paso t , la celda i tiene el símbolo s y no contiene a la cabecera de la máquina.
- $f(t, i) = (s, q)$ si y solo si en el paso t , la celda i tiene el símbolo s y contiene a la cabecera de la máquina en estado q .

De la misma manera que la tesis de Church-Turing nos dice que una maquina de Turing puede simular cualquier proceso computable, la tesis de Church-Turing extendida nos dice que una maquina de Turing probabilística puede simular cualquier proceso computable estocástico. Un ejemplo de eso, que usaremos mas adelante, son las siguientes proposiciones:

Proposición 4.12. *Sea $\alpha \in (0, 1)$ un real computable. Luego hay una Maquina de Turing probabilística que empezando en la cinta vacía llega al estado de aceptación con probabilidad α .*

Demostración. La maquina M calculara por un lado un numero aleatorio x entre 0 y 1, con acceso a los bits aleatorios, y por otro lado los dígitos de α , que puede hacer que es computable. Si en algún momento de la computación determina que $x < \alpha$, acepta. \square

Proposición 4.13. *Sean α, β reales computables tal que $0 \leq \alpha \leq \beta \leq 1$. Luego que existe una maquina de Turing probabilística M y un estado $q \in Q$ tal que si denotamos con $p(n)$ la probabilidad de que la máquina M este en esta q tras n pasos, se cumple que $\liminf p(n) = \alpha$ y $\limsup p(n) = \beta$.*

Demostración. La maquina M realizara el siguiente proceso infinito, dividido en turnos indexados por los naturales. En el turno $2n$, calculara n dígitos de α , que denotaremos α_n y un numero aleatorio x de n dígitos, y llegara al estado q si y solo si $\alpha_n > x$. En el turno $2n+1$, calculara n dígitos de β , que denotaremos β_n y un numero aleatorio x de n dígitos, y llegara al estado q si y solo si $\beta_n > x$.

Es fácil ver que la probabilidad de que la máquina M alcance el estado q en el turno $2n + 1$ es $\lfloor \alpha \cdot 2^n \rfloor / 2^n$ y la probabilidad de que alcance el estado q en el turno $2n + 1$ es $\lfloor \beta \cdot 2^n \rfloor / 2^n$, luego el limite inferior de $p(n)$ es α y el limite superior de $p(n)$ es β . \square

Definición 4.14. *Sea $M = (Q, \Sigma, \sigma, q_0, b, F)$ una maquina de Turing probabilística. Vamos a definir una fórmula $\phi_M(a, b, c, d, e)$ en la lógica de primer orden de grafos ordenados, que significara que la tupla $(a, b, c, d, e) \in G^5$ **codifica el comportamiento de la maquina de Turing M por $a - 1$ pasos.** La fórmula ϕ_M sera la conjunción de lo siguiente:*

1. Vale que $2 < a < b < c < d < e < n - s$, donde $s = |\Sigma| \cdot (|Q| + 1)$
2. El grafo induce una isomorfismo entre los intervalos $(2, a)$ y (b, c) que respeta el orden. Es decir, vale:

$$(\forall x \in (2, a) \exists! y \in (b, c) : x \sim y) \wedge (\forall y \in (b, c) \exists! x \in (2, a) : x \sim y) \wedge (\forall x, x' \in (2, a) \forall y, y' \in (b, c) : x \sim y \wedge x' \sim y' \wedge x < x' \implies y < y')$$
3. El grafo induce un isomorfismo entre $(2, a) \times (b, c)$ y (d, e) que preserva el orden, dotando a $(2, a) \times (b, c)$ del orden lexicográfico . Lo primero se escribe como

$$(\forall x \in (2, a) \forall y \in (b, c) \exists! z \in (d, e) : (x \sim z \wedge y \sim z)) \wedge (\forall z \in (d, e) \exists! x \in (2, a) \exists! y \in (b, c) : x \sim z \wedge y \sim z)$$

Y lo segundo como:

$$(\forall x, x' \in (2, a) \forall y, y' \in (b, c) \forall z, z' \in (d, e) ((x \sim z \wedge y \sim z \wedge x' \sim z' \wedge y' \sim z' \wedge z' < z) \Rightarrow (x' < x \wedge y' < y)))$$

Sergio: No es muy fuerte esto último? No podría ser $x' = x \wedge y' < y$?

4. El grafo induce una función de (d, e) a $(n - s, n)$. Es decir, vale que:

$$\forall x \in (d, e) \exists! y \in (n - s, n) : (x \sim y)$$

5. Por los items (ii) y (iii) anteriores, podemos ver a (d, e) como $(2, a) \times (b, c) = (2, a)^2$, es decir, una cuadrícula de lado $a - 1$. Por el item (iv), cada cuadrícula tiene asignado un único elemento de $(n, n - s) \equiv \Sigma \times (Q \cup \{*\})$. La última condición es que esa cuadrícula simule la máquina M empezando desde el estado vacío, en el sentido de la definición 4.11, con el bit aleatorio en cada paso viniendo del vértice 1, que en toda la construcción anterior fue dejado de lado y es por tanto independiente al resto.

Observación 4.15. No es realmente necesario pedir en el punto (ii) que la biyección respete el orden para que la codificación funcione correctamente, pero nos facilitará las cuentas más adelante.

Lo interesante es que una vez que tenemos codificada una máquina de Turing probabilística, podemos expresar su comportamiento con fórmulas de primer orden.

Proposición 4.16. Sea $M = (Q, \Sigma, \sigma, q_0, b, F)$ una máquina de Turing probabilística, $S \subset Q$ un subconjunto de estados, G un grafo ordenado y $a, b, c, d, e \in G$. Son expresables en lógica de primer orden de grafos:

- La tupla (a, b, c, d, e) codifica a la máquina M y en la simulación dada por (a, b, c, d, e) la máquina llega en algún momento a un estado de S .
- La tupla (a, b, c, d, e) codifica a la máquina M y en la simulación dada por (a, b, c, d, e) termina en un estado de S .

Definición 4.17. En particular, hay una fórmula que expresa que $G \models \phi_M(a, b, c, d, e)$ y que en la simulación dada por (a, b, c, d, e) la máquina M llega a un estado de aceptación en a pasos. Vamos a denotar a esa fórmula $\theta_M(a, b, c, d, e)$.

Observación 4.18. Si la máquina de Turing es determinística, su comportamiento no depende de la simulación, luego es posible expresar "La tupla (a, b, c, d, e) codifica a la máquina M y la máquina M llega a un estado de S en $a - 1$ pasos"

Definición 4.19. Sea M una máquina de Turing probabilística, $n > 0$ y G un grafo ordenado. Decimos que G **codifica a M por n pasos** si que existen $b, c, d, e \in G$ tal que la tupla $(n + 1, b, c, d, e)$ codifica a M , es decir, si $G \models \exists x \exists y \exists z \exists w : \phi_M(n + 1, x, y, z, w)$.

El siguiente lema es fundamental porque nos da una cota muy precisa sobre cuantos pasos de una máquina M simula un grafo aleatorio de tamaño dado y en particular, demuestra que un grafo aleatorio suficientemente grande simula una cantidad arbitraria de pasos de cualquier máquina.

Lema 4.20. Sea M una maquina de Turing probabilística y sea $(G, <)$ un grafo aleatorio ordenado de $n = 2^{k^2+k^{1.5}}$ vértices. Cuando $k \rightarrow \infty$, casi seguro G codifica a M por k pasos y casi seguro G **no** codifica M por $k + 1$ pasos

Demostración. Fijemos un a y sea X la variable aleatoria que cuenta la cantidad de tuplas (x, y, z, w) tal que $G \models \phi_M(a + 1, x, y, z, w)$. Vamos a calcular la esperanza y varianza de X y eso nos permitira demostrar el enunciado.

Fijados $b, c, d, e \in G$, ¿Cual es la probabilidad de que $G \models \phi_M(a, b, c, d, e)$? Como las condiciones exigen que exista una biyección entre $(2, a + 1)$ y (b, c) , y entre $(2, a + 1) \times (b, c)$ y (d, e) , es 0 a no ser que $|(b, c)| = a$ y $|(d, e)| = a^2$. Si se cumplen las condiciones anteriores, podemos calcular la probabilidad de que se cumplan todas las condiciones necesarias. La probabilidad de que el grafo induzca una biyección que preserve el orden entre $(2, a + 1)$ y (b, c) es 2^{-a^2} , ya que entre las 2^{a^2} maneras que el grafico puede asignar los vértices entre $(2, a + 1)$ y (b, c) hay solo una valida. Similarmente, la probabilidad de que el grafo induzca una biyección entre $(2, a + 1) \times (b, c)$ y (d, e) es 2^{-a^4} y es independiente a la anterior. La probabilidad de que el grafo induzca una función entre (d, e) y $(n - s, n)$ es 2^{-sa^2} y la probabilidad de que en cada elemento de (d, e) simule correctamente la transición de M es 2^{-a^2} y son independientes a las anteriores.

Juntando todo, la probabilidad de que $G \models \phi_M(a, b, c, d, e)$ es $2^{-a^2} \cdot 2^{-a^4} \cdot 2^{-sa^2} \cdot 2^{-a^2} = 2^{-a^4 - O(a^2)}$

Enumeremos a todos las tuplas $(x, y, z, w) \in G^4$ tal que $|(x, y)| = a$ y $|(z, w)| = a^2$ y sea (S_i) el evento de que la i -ésima tupla cumple que $G \models \phi(a, x, y, z, w)$. tenemos que $X = \sum 1_{S_i}$, luego:

$$\mathbb{E}(X) = \mathbb{E}(\sum 1_{S_i}) = \sum \mathbb{E}(1_{S_i}) = \sum Pr(s_i)$$

Como hay aproximadamente $\frac{n^2}{2}$ tuplas de tales características nos queda:

$$\mathbb{E}(X) = \frac{n^2}{2} \cdot Pr(S_i) = \frac{n^2}{2} \cdot 2^{-a^4 - O(a^2)}$$

Calculemos ahora la varianza:

$$\mathbb{E}(X^2) = \mathbb{E}((\sum 1_{S_i})^2) = \mathbb{E}(\sum 1_{S_i}) + \mathbb{E}(\sum_{i \neq j} 1_{S_i} \cdot 1_{S_j}) = \mathbb{E}(X) + \mathbb{E}(\sum_{i \neq j} 1_{S_i} \cdot 1_{S_j})$$

Se puede ver que los eventos S_i y S_j son independientes si los intervalos de sus respectivas tuplas son disjuntos, mientras que si se intersecan la probabilidad de $S_i \cap S_j$ es 0

Nos queda entonces:

$$\mathbb{E}(X^2) = \mathbb{E}(X) + \mathbb{E}(\sum_{i, j \text{ disjuntos}} 1_{S_i} \cdot 1_{S_j}) = \frac{n^2}{2} \cdot 2^{-a^4 + O(a^2)} + (\frac{n^2}{2} \cdot 2^{-a^4 + O(a^2)})^2$$

Notemos que el término $O(a^2)$ es el mismo en ambos exponentes, luego podemos escribirlo como $\lambda \cdot a^2$.

$$Var(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{n^2}{2} \cdot 2^{-a^4 + \lambda a^2} + (\frac{n^2}{2} \cdot 2^{-a^4 + \lambda a^2})^2 - (\frac{n^2}{2} \cdot 2^{-a^4 + \lambda a^2})^2 = \frac{n^2}{2} \cdot 2^{-a^4 + \lambda a^2}$$

Por tanto:

$$\sigma(X) = \sqrt{\text{Var}(X)} = \frac{n}{\sqrt{2}} \cdot 2^{-\frac{a^4}{2} + O(a^2)}$$

Ahora podemos probar el enunciado:

Tomando $n = 2^{k^2+k^{1.5}}$ y $a = k + 1$ la esperanza nos queda:

$$\mathbb{E}(X) = 2^{k^4+k^3-1} \cdot 2^{-(k+1)^4+O((k+1)^2)} = 2^{k^4+k^3-k^4-4k^3-O(k^2)} = 2^{-3k^3-O(k^2)}$$

Y por lo tanto casi seguramente el grafo G **no** codifica a M por $k + 1$ pasos, ya que por la desigualdad de Markov

$$\lim_{k \rightarrow \infty} \Pr(X \geq 1) \leq \lim_{k \rightarrow \infty} \mathbb{E}(X) = \lim_{k \rightarrow \infty} 2^{-3k^3-O(k^2)} = 0$$

Por otro lado, tomando $n = 2^{k^2+k^{1.5}}$ y $a = k$ la esperanza nos queda:

$$\mathbb{E}(X) = 2^{k^4+k^3-1} \cdot 2^{-k^4+O(k^2)} = 2^{k^3+O(k^2)}$$

Y la desviación estandar:

$$\sigma(X) = \frac{2^{k^2+k^{1.5}}}{\sqrt{2}}$$

Por la desigualdad de Chebyshev:

$$\Pr(X = 0) \leq \Pr(|X - 2^{k^3+O(k^2)}| > 2^{k^3+O(k^2)}) = \Pr(|X - \mu| > 2^{k+o(k)} \cdot \sigma) < \frac{1}{(2^{k+o(k)})^2}$$

Como este ultimo término tiende a 0 cuando $k \rightarrow \infty$, se sigue que casi seguro G codifica a M por k pasos. □

Con ese teorema, podemos demostrar una serie de resultados negativos, empezando por resultados de indecidibilidad:

Proposición 4.21. [33] *No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de los grafos ordenados, si vale que $\lim_{n \rightarrow \infty} \Pr(G_{<}(n) \models \phi) = 1$.*

Demostración. Sea M una maquina de Turing deterministica y sea $\phi = \exists x \exists y \exists z \exists w \exists u : \theta_M(x, y, z, w, u)$, es decir, $G \models \phi$ si y solo si la G simula la maquina M hasta que llega al estado de aceptación. Si la maquina M nunca llega al estado de aceptación, se sigue que para todo n vale que $\Pr(G_{<}(n) \models \phi) = 0$. Por otro lado, si la maquina M llega al estado de aceptación en a pasos, por el corolario 4.20 se sigue que $\lim_{n \rightarrow \infty} \Pr(G_{<}(n) \models \phi) = 1$. Luego, un algoritmo para determinar dada ϕ si $\lim_{n \rightarrow \infty} \Pr(G_{<}(n) \models \phi)$ tiende a 1 nos permitira determinar si una maquina dada llega al estado de aceptación. Como sabemos que ese problema es indecidible, se sigue que no puede existir dicho algoritmo. □

Proposición 4.22. [33] *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de los grafos ordenados, si el límite $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi)$ existe.*

Demostración. Sea $M = (Q, \Sigma, \sigma, q_0, b, F)$ una maquina de Turing dada. Vamos a construir una maquina de turing $M^* = (Q \times \{0, 1\}, \Sigma, \sigma^*, q_0, b, F)$ con función de transición dada por:

$$\sigma^*((q, n), a) = \begin{cases} ((q', 1 - n), b) & \text{si } \sigma(q, a) = (q', b) \text{ y } q' \notin F \\ ((q', 0), b) & \text{si } \sigma(q, a) = (q', b) \text{ y } q' \in F \end{cases}$$

En lenguaje coloquial, la maquina M^* tiene dos estados por cada estado q de M , los estados $(q, 0)$ y $(q, 1)$. En cada paso, respeta las transiciones de M pero alterna entre los estados de forma $(q, 0)$ y los de $(q, 1)$, excepto cuando llega a un estado de aceptación, en cuyo caso se mantiene en los estados de la forma $(q, 0)$. Por la proposición ?, podemos construir una oración ϕ que exprese la propiedad "Si el grafo G codifica n pasos de M^* , la maquina M^* esta en un estado de la forma $(q, 0)$ tras n pasos".

Si la maquina M llega a un estado de aceptación, como por la proposición 4.20 sabemos que $G_{<}(n)$ simula una cantidad arbitraria de pasos de M^* , se sigue que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = 1$. Por otro lado, si la maquina M **no** llega a un estado de aceptación, se sigue que la maquina M^* alternara indefinidamente entre los estados de la forma $(q, 0)$ y $(q, 1)$, luego por el teorema 4.20 se sigue que:

$$\lim_{k \rightarrow \infty} Pr(G_{<}(2^{(2k)^2 + (2k)^{1.5}}) \models \phi) = 1$$

$$\lim_{k \rightarrow \infty} Pr(G_{<}(2^{(2k+1)^2 + (2k+1)^{1.5}}) \models \phi) = 0$$

luego $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi)$ no existe. Por tanto, de existir un algoritmo capaz de determinar cuando una oración en la lógica de grafos ordenados dada tiene probabilidad limite podríamos construir un algoritmo para determinar cuando una maquina de Turing M llega a un estado de aceptación, pero sabemos que eso es imposible, luego no hay tal algoritmo. □

El siguiente teorema es un caso particular del teorema de Trakhtenbrot, que mencionamos en el capítulo 1:

Proposición 4.23. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de los grafos ordenados, si hay existe un grafo ordenado finito G tal que $G \models \phi$.*

Demostración. Muy similar a la proposición ??. Sea M una maquina de Turing determinística y sea $\phi = \exists x \exists y \exists z \exists w \exists u : \theta_M(x, y, z, w, u)$. Si la maquina M nunca llega a un estado de aceptación, ningún grafo G cumple ϕ . Por otro lado, si la maquina M llega a un estado de aceptación, por el teorema 4.20 sabemos que debe existir un grafo ordenado de $2^{k^2 + k^{1.5}}$ que cumple ϕ . Se sigue que de existir un algoritmo que pueda determinar si una oración ϕ en el lenguaje de los grafos ordenados tiene modelo finito, podríamos construir un algoritmo que decida si una maquina de Turing dada llega a un estado de aceptación, pero sabemos que eso es imposible, luego no existe tal algoritmo. □

Una pregunta que no encontramos mencionada en la literatura es la de caracterizar que valores puede adoptar el límite $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi)$ donde ϕ es una oración de primer orden de grafos ordenados. Usando la codificación de maquinas de Turing probabilísticas, vamos a mostrar que incluye a todos los reales computables entre 0 y 1.

Proposición 4.24. *Sea $\alpha \in (0, 1)$ un real computable. Luego existe una oración ϕ en la lógica de primer orden de grafos ordenados tal que:*

$$\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = \alpha$$

Demostración. Sabemos por la proposición 4.12 que dado $\alpha \in (0, 1)$ computable, existe una maquina de Turing probabilística M que empezando con la cinta vacía llega al estado de aceptación con probabilidad α y rechaza con probabilidad $1 - \alpha$. Sea ϕ la oración que expresa la siguiente propiedad: "Si el grafo G simula a M por $a - 1$ pasos y no por a pasos, y (b, c, d, e) es la menor tupla en el orden lexicográfico tal que $G \models \phi_M(a, b, c, d, e)$, en la simulación dada por (a, b, c, d, e) la máquina M llega al estado de aceptación". En lenguaje formal:

$$\phi = [\forall a \forall b \forall c \forall d \forall e (\phi_M(a, b, c, d, e) \wedge \neg(\exists a' \exists b' \exists c' \exists d' \exists e' : a' > a \wedge \phi_M(a', b', c', d', e'))$$

$$\wedge \neg(\exists b' \exists c' \exists d' \exists e' : \phi_M(a, b', c', d', e') \wedge (b', c', d', e') <_L (b, c, d, e))) \implies \theta_M(a, b, c, d, e)]$$

Donde $\theta_M(a, b, c, d, e)$ es la fórmula que representa que (a, b, c, d, e) simula a la maquina y que dicha simulación llega al estado de aceptación, y $(b, c, d, e) >_L (b, c, d, e)$ es una abreviación del orden lexicográfico. Como sabemos por la proposición 4.20 que cuando n tiende a infinito la maquina simula una cantidad arbitrariamente grande de pasos, y la simulación minimal respecto al orden lexicográfico es elegida de manera uniforme entre todos los comportamientos de la maquina, se sigue que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = \alpha$. \square

De hecho, vale un resultado mas en general:

Proposición 4.25. *Sean $\alpha, \beta \in (0, 1)$ reales computables tal que $\alpha < \beta$. Luego existe una oración ϕ en la lógica de primer orden de grafos ordenados tal que $\lim inf_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = \alpha$ y $\lim sup_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = \beta$*

Demostración. análoga a la demostración anterior. Sabemos, por la proposición 4.13 que dados reales computables $0 \leq \alpha \leq \beta \leq 1$, existe una máquina de Turing probabilística M y un estado $q \in Q_M$ tal que si $p(n)$ denota la probabilidad de que M este en estado q tras n pasos vale que $\lim inf p(n) = \alpha$ y $\lim sup p(n) = \beta$. Sea ϕ la oración que expresa la propiedad "Si el grafo G simula a M por $a - 1$ pasos y no por a pasos, y (b, c, d, e) es la menor tupla en el orden lexicográfico tal que $G \models \phi_M(a, b, c, d, e)$, en la simulación dada por (a, b, c, d, e) la máquina M termina en estado q ".

Como sabemos por la proposición 4.20 que cuando n tiende a infinito la maquina simula una cantidad arbitrariamente grande de pasos, y la simulación minimal respecto al orden lexicografico es elegida de manera uniforme entre todos los comportamientos de la maquina, se sigue que $\lim inf Pr(G_{<}(n) \models \phi) = \alpha$ y $\lim sup Pr(G_{<}(n) \models \phi) = \beta$. \square

Por ultimo, el siguiente resultado muestra que incluso para las oraciones ϕ sobre grafos ordenados cuya probabilidad converge, es imposible acotar uniformemente cuanto tardan en alcanzar su probabilidad asintótica en términos de su rango de cuantificador $qr(\phi)$, como hicimos para grafos aleatorios.

Proposición 4.26. *No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración ϕ en el lenguaje de grafos ordenados valga que si $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = 1$ se sigue que $Pr(G_{<}(N) \models \phi) > 0.95$ para todo $N > f(qr(\phi))$.*

Demostración. Vamos a suponer que existe tal función f y mostraremos que se sigue que hay un algoritmo para determinar si una oración en el lenguaje de grafos ordenados tiene un modelo finito, lo que contradice la proposición 4.23.

Dada oración ϕ en el lenguaje de los grafos ordenados, sea ϕ^* la oración dada por $\exists x \exists y : x < y \wedge \phi_{(x,y)}$. Si ϕ no es satisfacible, es evidente que ningún grafo ordenado cumple ϕ^* . Por otro lado, si ϕ es satisfacible, es fácil ver que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi^*) = 1$. Luego, por hipótesis, para ver si ϕ es satisfacible basta ver si hay algún grafo de tamaño $n = f(qr(\phi))$ vértices que satisface ϕ^* . □

4.1.4. Ley de poca variación

En vista de los resultados anteriores, es fácil concluir que la lógica de primer orden sobre grafos ordenados es simplemente demasiado expresiva para poder obtener ningún resultado interesante, pero de hecho, no es el caso. En [34] Shelah demostró un resultado que bautizó como "ley 0-1 debil" de grafos ordenados, aunque sería más apropiado llamarla "Ley de poca variación". Su herramienta principal es un resultado bastante profundo de teoría de complejidad de circuitos. Antes de poder enunciarla, necesitamos algunas definiciones preliminares:

Definición 4.27. Un **circuito booleano** es una función $\mathcal{C} : \{0, 1\}^n \rightarrow \{0, 1\}$ definida a partir de la composición de las funciones negación y disjunción y conjunción de tamaño arbitrario $\vee : \{0, 1\}^k \rightarrow \{0, 1\}$, $\wedge : \{0, 1\}^k \rightarrow \{0, 1\}$.

Definición 4.28. La profundidad $depth(\mathcal{C})$ de un circuito se define recursivamente como:

- $depth(\neg \mathcal{C}) = depth(\mathcal{C}) + 1$
- $depth(\vee \mathcal{C}_i) = \max(depth(\mathcal{C}_i)) + 1$
- $depth(\wedge \mathcal{C}_i) = \max(depth(\mathcal{C}_i)) + 1$

Definición 4.29. El tamaño $size(\mathcal{C})$ de un circuito se define recursivamente como:

- $size(\neg \mathcal{C}) = size(\mathcal{C}) + 1$
- $size(\vee_i \mathcal{C}_i) = \sum_i (size(\mathcal{C}_i)) + 1$
- $size(\wedge_i \mathcal{C}_i) = \sum_i (size(\mathcal{C}_i)) + 1$

Observación 4.30. Es fácil ver que toda función $f : \{0, 1\}^n \rightarrow \{0, 1\}$ puede expresarse como un circuito booleano, en particular por uno de la forma

$$\mathcal{C}(x) = \bigvee_{S \in \Omega} \left[\left(\bigwedge_{i \in S} x_i \right) \wedge \left(\bigwedge_{i \notin S} \neg x_i \right) \right]$$

para algún $\Omega \subset \mathcal{P}(\{1, \dots, n\})$. Por otro lado, un argumento combinatorio originalmente de Shanon[?] muestra que casi todas las funciones $f : \{0, 1\}^n \rightarrow \{0, 1\}$ requieren circuitos de tamaño mayor a $O(\frac{2^n}{n})$ para ser computadas. Dicha observación dio lugar a la **teoría de complejidad de circuitos**, que estudia que clase de funciones pueden ser computados con circuitos de tamaño y profundidad acotados.

Definición 4.31. Sea \mathcal{C} un circuito de input n . Para $m \leq n$, denotamos con $p_{\mathcal{C}}(m)$ la probabilidad de que dada un input $s \in \{0, 1\}^n$ con m 1s elegida de manera uniforme, se cumpla que $\mathcal{C}(s) = 1$.

El resultado clave de la demostración de Shelah es el siguiente teorema:

Teorema 4.32. [34] Sea $\varepsilon > 0$. Luego, para todo d y todo k existe un n_0 tal que para todo circuito \mathcal{C} de input $2n > n_0$, con profundidad d y tamaño menor a n^k se cumple que:

$$|p_{\mathcal{C}}(n) - p_{\mathcal{C}}(n+1)| < \varepsilon$$

El teorema anterior implica uno de los resultados mas celebres de teoría de circuitos, que no existe ninguna familia de circuitos de profundidad constante y tamaño polinómico que compute PARITY, la función que computa la suma de los bits de input modulo 2. En términos técnicos, PARITY no pertenece a la clase AC_0 .

Proposición 4.33. (PARITY no pertenece a AC_0) No existe ninguna familia de circuitos $\mathcal{C}_n : \{0, 1\}^n \rightarrow \{0, 1\}$ que compute PARITY tal que existen d y k tal que para todo n se cumple que $size(\mathcal{C}_n) < n^k$ y $depth(\mathcal{C}_n) < d$.

Demostración. Si existieran, por el teorema anterior para n suficientemente grande se cumpliría que:

$$|p_{\mathcal{C}_{\varepsilon \setminus \infty}}(n) - p_{\mathcal{C}_{\varepsilon \setminus +\infty}}(n)| < \frac{1}{2}$$

Pero como los \mathcal{C}_n computan PARITY esa diferencia debería ser siempre 1, absurdo. \square

Para continuar la demostración, introduciremos una familia de funciones que evalúan la validez de formulas en grafos ordenados y mostraremos como representarlas con circuitos.

Definición 4.34. Sea $n > 0$ y sea ϕ una oración en el lenguaje de los grafos ordenados. Definimos la función booleana $f_{\phi} : \{0, 1\}^{\frac{n(n-1)}{2}} \times \{0, 1\}^n \rightarrow \{0, 1\}$ de la siguiente manera: interpretando la primer parte del input como un grafo G sobre $\{1, \dots, n\}$ y la segunda parte como un subconjunto $S \subset \{1, \dots, n\}$, $f_{\phi}(G, S) = 1$ si y solo si $G \models \phi_S$. Si fijamos el grafo G de antemano, podemos definir la función $f_{\phi, G} : \{0, 1\}^n \rightarrow \{0, 1\}$ como $f_{\phi, G}(S) = f_{\phi}(G, S)$

Lema 4.35. Sea ϕ una oración en el lenguaje de los grafos ordenados. Luego, existe un $d > 0$ y un polinomio P tal que para todo n existe un circuito $\mathcal{C}_\setminus : \{0, 1\}^{\frac{n(n-1)}{2}} \times \{0, 1\}^n \rightarrow \{0, 1\}$ que computa la función f_ϕ definida anteriormente, y mas aun, vale que $\text{depth}(\mathcal{C}_\setminus) \leq d$ y $\text{size}(\mathcal{C}_\setminus) < P(n)$.

Demostración. Por inducción en la estructura de ϕ . Para el caso base, es evidente que las oraciones atómicas $x < y, x \sim y, x \in S$ se pueden computar con circuitos de tamaño $O(1)$. Para el paso inductivo, las únicas oraciones que hace falta analizar son las de forma $\exists : \phi(x)$ y para eso basta notar que $\exists x \phi(x) = \bigvee_x \phi(x)$, luego como por hipótesis para cada x el valor de $\phi(x)$ se puede computar con un circuito de profundidad d y tamaño $P(n)$, se sigue que $\exists x : \phi(x)$ se puede computar con un circuito de profundidad $d + 1$ y tamaño $n \cdot P(n)$. \square

Ejemplo 4.36. Sea ϕ la oración $[\exists x : \forall y x \sim y]$. Luego la familia de circuitos del teorema anterior esta dada por:

$$\bigvee_x [(x \in S) \wedge \bigwedge_y ((y \notin S) \vee (x \sim y))]$$

Con profundidad 3 y tamaño $n \cdot (1 + n \cdot (1 + 1)) = 2n^2 + n$.

Fijando el grafo G primero, obtenemos la siguiente variante del lema anterior:

Lema 4.37. Sea ϕ una oración en el lenguaje de los grafos ordenados. Luego, existe $d > 0$ y un polinomio P tal que para todo grafo G de n vértices, hay un circuito booleano $\mathcal{C} : \{0, 1\}^n \rightarrow \{0, 1\}$ que computa la función $f_{\phi, G}$ definida anteriormente y mas aun, $\text{depth}(\mathcal{C}_\setminus) < d$ y $\text{size}(\mathcal{C}) < P(n)$.

Teorema 4.38. [34](*Ley de poca variación de grafos ordenados*) Sea ϕ una oración en el lenguaje de los grafos ordenados. Luego, vale que:

$$\lim_{n \rightarrow \infty} [Pr(G_{<}(n) \models \phi) - Pr(G_{<}(n+1) \models \phi)] = 0$$

Demostración. Sea ϕ una oración en el lenguaje de los grafos ordenados. Por el lema 4.37 existen $d > 0$ y polinomio P tal que para todo grafo G sobre $\{1, \dots, n\}$ existe un circuito \mathcal{C} que computa $f_{\phi, G}$ y tal que $\text{depth}(\mathcal{C}) < d$ y $\text{size}(\mathcal{C}) < P(n)$. Ahora, por el teorema 4.32, para todo $\varepsilon > 0$ existe un n_0 tal que para todo grafo G de $2n > n_0$ vértices, se cumple que:

$$|[p_{\mathcal{C}}(n) - p_{\mathcal{C}}(n+1)]| < \varepsilon$$

Es decir, para todo grafo G de $2n > n_0$ vértices se cumple que

$$|Pr(G \models \phi_S : |S| = n) - Pr(G \models \phi_S : |S| = n+1)| < \varepsilon$$

Como eso se cumple para **todo** grafo G de $2n > n_0$, en particular sera cierto si elegimos G de forma aleatoria entre todos los grafos de $2n$ vértices. Pero el grafo inducido por un subconjunto de tamaño k elegido de forma uniforme de un grafo elegido de forma uniforme de n vértices no es mas que un grafo de k vértices elegido de forma uniforme. Luego, nos queda que para todo n tal que $2n > n_0$:

$$|Pr(G_{<}(n) \models \phi) - Pr(G_{<}(n+1) \models \phi)| < \varepsilon$$

De lo que se sigue el teorema. \square

Observación 4.39. Notemos que el teorema anterior es consistente con todo lo que observamos hasta ahora sobre probabilidades de oraciones sobre grafos ordenados: El primer ejemplo de oración sin probabilidad limite que vimos es el que codifica la maratón de danza, ya que la probabilidad de que hubiera un ganador en la maratón es asintóticamente distinta si hay 2^n o $1.5 \cdot 2^n$ participantes, pero es fácil ver que la no hay diferencia asintótica entre una maratón de n participantes y una de $n + 1$ participantes. Por otro lado, es fácil ver por la proposición 4.20 que lo mismo ocurre con las oraciones que hablan sobre propiedades de las maquinas de Turing codificadas en un grafo ordenado.

El método anterior es extremadamente poderoso y de hecho, permite demostrar un enunciado bastante mas general que el que enuncia Shelah, considerando no solo grafos aleatorios ordenados también a todos los grafos ordenados sparse:

Teorema 4.40. (Ley de poca variación de grafos ordenados sparse) Sea $c > 0$ y $\alpha \geq 0$. ϕ una oración en el lenguaje de los grafos ordenados. Luego, vale que:

$$\lim_{n \rightarrow \infty} [Pr(G_{<}(n, c \cdot n^{-\alpha}) \models \phi) - Pr(G_{<}(n + 1, c \cdot n^{-\alpha}) \models \phi)] = 0$$

Demostración. En la demostración anterior vimos que dada oración ϕ y $\varepsilon > 0$ existe un n_0 tal que para todo grafo ordenado G de $2n > n_0$ vértices se cumple que:

$$|Pr(G \models \phi_S : |S| = n) - Pr(G \models \phi_S : |S| = n + 1)| < \varepsilon$$

Como eso se cumple para todo grafo, en particular se cumple si elegimos G de forma aleatoria entre los grafos sobre $\{1, \dots, n\}$. Elegimos $G \sim G(2n, c \cdot n^{-\alpha})$. Ahora, el grafo inducido por un subconjunto de n vértices de G tiene distribución $G(n, c \cdot n^{-\alpha})$ y el inducido por uno de $n + 1$ tiene distribución $G(n + 1, c \cdot n^{-\alpha})$. Luego nos queda que :

$$|Pr(G_{<}(n, c \cdot n^{-\alpha}) \models \phi) - Pr(G_{<}(n + 1, c \cdot n^{-\alpha}) \models \phi)| < \varepsilon$$

□

El teorema de Shelah nos permita probar algunos resultados de inexpresabilidad asintótica, como demostramos anteriormente para grafos no ordenados:

Proposición 4.41. No hay ninguna oración en el lenguaje de los grafos ordenados que exprese de forma asintótica la propiedad "El grafo tiene una cantidad par de vértices".

Demostración. De existir tal proposición, valdria que $\lim_{n \rightarrow \infty} [Pr(G_{<}(2n) \models \phi) - Pr(G_{<}(2n + 1) \models \phi)] = 1$, lo que contradice el teorema anterior. □

4.2. Otras estructuras ordenadas

4.2.1. Conjuntos con múltiples ordenes

En esta sección estudiaremos el comportamiento asintótico de conjuntos equipados con multiples ordenes lineales, que denotaremos $(A, <_1, \dots, <_n)$, y demostraremos resultados analogos a los de Shelah y Compton para grafos aleatorios ordenados. Aunque este tipo de estructura no aparece frecuentemente en matemática, los resultados que demostraremos en esta sección nos sirvan para demostrar fácilmente resultados análogos para otros tipos de estructuras mas naturales y no analizadas previamente en la literatura, como grafos sparse ordenados y conjuntos equipados de un orden y una función.

Definición 4.42. Denotaremos $A(n, <_1, \dots, <_k)$ la variable aleatoria con una distribución uniforme entre todas las formas de asignar k ordenes lineales al conjunto $\{1, \dots, n\}$.

De manera análoga a lo que sucedía con grafos ordenados aleatorios, es fácil ver que $A(n, <_1, \dots, <_k)$ no cumple una ley 0 – 1 para $k \geq 2$, como muestra el siguiente ejemplo: Sea ϕ la siguiente oración sobre conjuntos con dos ordenes

$$\phi = [\forall x \forall y (\forall z (x <_1 z) \wedge (z <_1 x)) \rightarrow (x <_2 y)]$$

que afirma que si x e y son respectivamente el elemento minimal y maximal respecto a $<_1$, entonces $x <_2 y$. Se sigue que $Pr(A(n, <_1, <_2) \models \phi) = \frac{1}{2}$.

Vamos a demostrar ahora que en general, la probabilidad de que $A(n, <_1, <_2)$ cumpla una oración ϕ no converge a ningún valor cuando n tiende a infinito, de la misma manera que lo que hicimos para grafos ordenados:

Proposición 4.43. Hay una oración ϕ en lenguaje de conjuntos con dos ordenes que simula la maratón de danza acotada, es decir, tal que:

$$Pr(A(n, <_1, <_2) \models \phi) = p_{MDA}(n)$$

Demostración. Sea ϕ la oración dada por:

$$\phi = [\exists x : \exists y : (\forall w w <_1 y \rightarrow x <_2 w) \wedge (\forall z : (\forall w w <_1 y \rightarrow z <_2 w)) \rightarrow (z = x)]$$

Esta oración codifica la maratón de danza de una forma muy similar a la oración vista en anterior: Sea $S = \{1, \dots, n\}$ ordenados por $<_1$. El evento "el participante i pasa la ronda j " se corresponde con el evento $i <_2 j$, que tienen probabilidad $\frac{1}{2}$ y son independientes entre si, y el hecho de que el participante i abandone siempre en la ronda i se corresponde con el hecho de que las relaciones de orden son antisimétricas, es decir, que para todo i se cumple que $i \not<_2 i$. \square

Corolario 4.44. Existe una oración ϕ en el lenguaje de conjuntos con tres ordenes tal que el limite $\lim_{n \rightarrow \infty} Pr(A(n, <_1, <_2, <_3) \models \phi)$ **no** existe.

Antes de seguir, necesitamos la siguiente definición:

Definición 4.45. Sea A un conjunto equipado de un orden $<$ y sea $S \subset A$. Decimos que S induce la función $f : A \rightarrow A$ dada por

$$f(a) = \begin{cases} \text{mín } \{s \in S : s < a\} & \text{si } \text{mín } \{s \in S : s < a\} \neq \emptyset \\ \text{mín } \{s \in S\} & \text{si } \text{mín } \{s \in S : s < a\} = \emptyset \end{cases}$$

Al igual que con los grafos ordenados, vamos a demostrar que se puede codificar el comportamiento de una máquina de Turing probabilística en un conjunto con múltiples ordenes.

Definición 4.46. Sea $M = (Q, \Sigma, \sigma, q_0, b, F)$ una máquina de Turing probabilística. Vamos a definir una fórmula $\phi_M(a, b, c)$ sobre el lenguaje de los conjuntos con tres ordenes lineales, que significara que la tupla (a, b, c) codifica el comportamiento de M por $a - 1$ pasos. La fórmula ϕ_M sera la conjunción de lo siguiente:

1. Vale que $2 <_1 a <_1 b <_1 c <_1 n - s$, donde $s = |\Sigma| \cdot (|Q| + 1)$
2. Los ordenes $<_2$ y $<_3$ inducen una biyección entre $(b, c)_1$ y $(2, a)_1 \times (2, a)_1$, y mas aun, si dotamos a $(2, a)_1^2$ del orden lexicográfico esa biyección preserva el orden. Es decir, vale:

$$\forall x \in (2, a)_1 \quad \forall x' \in (2, a)_1 \quad \exists! z \in (b, c)_1 : \forall y \in (2, a)_1 ((y <_2 x) \leftrightarrow (y <_2 z)) \wedge ((y <_3 x') \leftrightarrow (y <_3 z))$$

3. Por el item ii), podemos ver a $(b, c)_1$ como $(2, a)_1 \times (2, a)_1$, es decir, una cuadrícula de lado $a - 1$. Además, por el morfismo inducido por $<_2$, cada elemento de la cuadrícula tiene asignado un elemento de $(n, n - s) \equiv \Sigma \times (Q \cup \{*\})$. La última condición es que esa cuadrícula codifique a la máquina M en el sentido de la definición 4.11, con el bit aleatorio dado por el orden $<_3$ respecto al elemento 1, que fue dejado de lado en toda la construcción y es por tanto independiente de los eventos anteriores.

Definición 4.47. Sea M una máquina de Turing probabilística y $(A, <_1, <_2, <_3)$ un conjunto con tres ordenes lineales. Decimos que A codifica a M por n pasos si $A \models \exists x \exists y : \phi_M(n+1, x, y)$, es decir, existen $b, c \in A$ tal que $A \models \phi_M(n+1, b, c)$

El siguiente lema, análogo del lema 4.20, nos permitirá acotar con bastante precisión cuantos pasos de una máquina de Turing dada codifica un conjunto con tres ordenes aleatorios. En particular, demuestra que fijada una máquina M y un $k > 0$, la probabilidad de que $A(n, <_1, <_2, <_3)$ codifique a M por mas de k pasos tiende a 1 cuando n tiende a infinito.

Lema 4.48. Sea $c > 0$, M una máquina de Turing probabilística y $n = c \cdot (k + 1)^{k^2} \cdot (\ln k)^k$. Luego, cuando k tiende infinito, casi seguro $A(n, <_1, <_2, <_3)$ simula a M por k pasos y casi seguro **no** simula a M por $k + 1$ pasos.

Demostración. Dado k , sea X_k la variable aleatoria que cuenta la cantidad de tuplas $(b, c) \in A$ tal que $A \models \phi_M(k + 1, b, c)$. Vamos a calcular la esperanza y varianza de X_k y eso nos permitirá demostrar el enunciado.

Dados b, c en A ¿Cual es la probabilidad de que $A \models \phi_M(k + 1, b, c)$? Como la segunda condición de ϕ_M exige que haya una biyección entre $(2, k + 1)_1^2$ y $(b, c)_1$, la probabilidad es 0 si $|(b, c)_1| \neq k^2$. Si $|(b, c)_1| = k^2$, la probabilidad de que $<_2$ y $<_3$ induzcan, entre todos los morfismos posibles, la biyección que preserva el orden lexicográfico es $((k^2 + 1)^{k^2})^{-1}$. Por otro lado, la probabilidad de que $(b, c)_1$ simule correctamente todos los pasos de la máquina de Turing M esta acotada por $2^{-\lambda k^2}$ para algún $\lambda > 0$, que depende de la máquina M .

Enumeremos todas las tuplas (b, c) tal que $k + 1 <_1 b <_1 c <_1 n - s$ y que $|(b, c)_1| = k^2$ y sea S_i el evento "La i -ésima tupla cumple ϕ_M ". Calculemos ahora la esperanza:

$$\mathbb{E}(X_k) = \mathbb{E}\left(\sum 1_{S_i}\right) = \sum Pr(S_i) = (n - k^2) \cdot (k^2 + 1)^{-k^2} \cdot 2^{-ck^2}$$

Calculemos ahora la varianza:

$$\mathbb{E}(X_k^2) = \mathbb{E}\left(\left(\sum_i 1_{S_i}\right)^2\right) = \mathbb{E}\left(\sum_i 1_{S_i}\right) + \mathbb{E}\left(\sum_{i, j \text{ distintos}} 1_{S_i} 1_{S_j}\right) = \mathbb{E}(X_k) + \mathbb{E}\left(\sum_{i, j \text{ distintos}} 1_{S_i} 1_{S_j}\right)$$

Ahora, es fácil ver que los eventos S_i y S_j son independientes si los correspondientes intervalos son disjuntos, e incompatibles si no son disjuntos. Además, hay como mucho $(n - k)^2$ pares (S_i, S_j) con S_i y S_j disjuntos. Por tanto

$$\mathbb{E}\left(\sum_{i, j \text{ distintos}} 1_{S_i} 1_{S_j}\right) = \mathbb{E}\left(\sum_{i, j \text{ disjuntos}} Pr(S_i) \cdot Pr(S_j)\right) < (n-k)^2 \cdot [(k^2+1)^{-k^2} \cdot 2^{\lambda \cdot k^2}]^2$$

Pero notemos que ese último término no es otra cosa que $\mathbb{E}(X_k)^2$. Luego nos queda:

$$\mathbb{E}(X_k^2) < \mathbb{E}(X_k)^2 + \mathbb{E}(X_k)$$

Y por tanto:

$$Var(X_k) = \mathbb{E}(X_k^2) - \mathbb{E}(X_k)^2 < \mathbb{E}(X_k)$$

$$\sigma(X_k) < \sqrt{\mathbb{E}(X_k)}$$

Ahora podemos probar el enunciado. Supongamos $n = c \cdot (k+1)^{2k^2} \cdot (\ln k)^k$.

$$\mathbb{E}(X_k) = c \cdot (k+1)^{2k^2} \cdot (\ln k)^k - k^2 \cdot (k^2+1)^{-k^2} \cdot 2^{-ck^2} > c \cdot (\ln k)^k \cdot 2^{-ck^2}$$

Luego tenemos, por la desigualdad de Chebyshev:

$$Pr(X_k = 0) \leq Pr(|X_k - \mu| \geq \mu) = Pr(|X_k - \mu| < \sigma) < \frac{\sigma^2}{\mu^2} = \frac{1}{\mu} \longrightarrow 0$$

Luego, casi seguro $A(n, <_1, <_2, <_3)$ codifica a M por k pasos. Por otro lado:

$$\mathbb{E}(X_{k+1}) = n \cdot (k^2+2)^{-(k+1)^2} \cdot 2^{-c(k+1)^2} > n \cdot (k+1)^{-(k+1)^2} \cdot 2^{-ck^2}$$

$$\mathbb{E}(X_{k+1}) < c \cdot (k+1)^{k^2} \cdot (\ln k)^k \cdot (k+1)^{-(k+1)^2} \cdot 2^{-ck^2} = c \cdot (\ln k)^k \cdot (k+1)^{k^2-(k+1)^2} \cdot 2^{-ck^2}$$

$$\mathbb{E}(X_{k+1}) < c \cdot (\ln k)^k \cdot (k+1)^{-4k} \cdot 2^{-ck^2}$$

Ahora, es claro que el término crece mucho $(k+1)^{4k}$ más rápido que $(\ln k)^k$, luego, el término anterior tiende a 0 cuando k tiende a infinito. Entonces por la desigualdad de Markov tenemos:

$$Pr(X_{k+1} > 0) < \mathbb{E}(X_{k+1}) \longrightarrow 0$$

Y por tanto, tenemos que casi seguro $A(n, <_1, <_2, <_3)$ no codifica a M por $k+1$ pasos. \square

Proposición 4.49. *No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de conjuntos con tres ordenes, si vale que $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi) = 1$.*

Demostración. análoga a la demostración 4.21. Dada maquina de Turing M , sea ϕ la oración $\exists x \exists y \exists z \exists w \exists u : \theta_M(x, y, z, w, u)$. Si la maquina no llega a un estado de aceptación, se sigue que $Pr(G_{<}(n) \models \phi) = 0$ para todo n . Por otro lado, como por el lema 4.48 para n suficientemente grande A codifica una cantidad de pasos arbitraria de M , si la maquina M llega a un estado de aceptación se sigue que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = 1$. Luego, de existir un algoritmo para determinar que oraciones tienen probabilidad asintótica 1, existiría un algoritmo para determinar que maquinas de Turing llegan a un estado de aceptación, pero sabemos que eso es imposible, luego no existe tal algoritmo. \square

Proposición 4.50. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de conjuntos con tres ordenes, si vale que el limite $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi)$ existe*

Demostración. análoga a la demostración 4.22. Dada maquina de Turing M , construimos maquina de Turing M^* con un subconjunto de estados S tal que si M nunca para, M^* esta en S en los pasos pares, pero si M para, M^* para en un estado fuera de S . Sea ϕ la oración que expresa "Si n es la mayor cantidad de pasos que codifica A , entonces M^* esta en un estado de S tras n pasos".

Si M para, se sigue que $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi) = 0$. Por otro lado, si M no para, se sigue por el lema 4.48 que:

$$\lim_{k \rightarrow \infty} Pr(\mathbf{A}((2k+1)^{(2k)^2} \cdot \ln(2k)^{2k}, <_1, <_2, <_3) \models \phi) = 1$$

$$\lim_{k \rightarrow \infty} Pr(\mathbf{A}((2k+2)^{(2k+1)^2} \cdot \ln(2k+1)^{2k+1}, <_1, <_2, <_3) \models \phi) = 0$$

Luego, el limite $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi)$ no existe. Es decir, si existiera un algoritmo para determinar cuando una oración en lenguaje de conjuntos con tres ordenes tiene una probabilidad limite existiría un algoritmo para determinar cuando una maquina de Turing para, pero eso es imposible. \square

Proposición 4.51. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de los conjuntos con tres ordenes, si hay existe un conjunto finito $(A, <_1, <_2, <_3)$ tal que $A \models \phi$.*

Demostración. Análoga a la demostración 4.23. \square

Proposición 4.52. *Sean $\alpha, \beta \in (0, 1)$ reales computables tal que $\alpha < \beta$. Luego, existe una oración ϕ en en lenguaje de conjuntos con tres ordenes tal que $\liminf Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi) = \alpha$ y $\limsup Pr(\mathbf{A}(n, <_1, <_2, <_3) \models \phi) = \beta$.*

Demostración. Análoga a la demostración 4.25. Sea \square

Proposición 4.53. *No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración ϕ en el lenguaje de conjuntos con tres ordenes valga que si $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(N, <_1, <_2, <_3) \models \phi) = 1$ se sigue que $Pr(\mathbf{A}(N, <_1, <_2, <_3) \models \phi) > 0.95$ para todo $N > f(qr(\phi))$.*

Demostración. Análoga a la demostración 4.26. De existir, tendríamos un algoritmo para determinar cuando una oración dada sobre el lenguaje de conjuntos con tres ordenes tiene modelo finito, pero demostramos anteriormente que tal algoritmo no existe. \square

La codificación que hicimos de una máquina de Turing requiere tres ordenes distintos, pero notamos anteriormente que solo dos ordenes bastan para codificar la maratón de danza acotado, lo que muestra que es un lenguaje con alta capacidad expresiva. Eso nos lleva a conjeturar lo siguiente:

Conjetura 4.54. Los teoremas 4.49-4.53 se pueden extender a conjuntos aleatorios con dos ordenes $A(n, <_1, <_2)$

4.2.2. Conjuntos ordenados con permutaciones

Definición 4.55. Denotamos $A(n, <, \sigma)$ a la variable aleatoria entre todos los conjuntos ordenados sobre $\{1, \dots, n\}$ con una permutación $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

Definición 4.56. Sea A un conjunto equipado de un orden total $<_1$ y una permutación $\sigma : A \rightarrow A$. Definimos los ordenes $<_2$ y $<_3$ sobre A como:

$$x <_2 y \text{ si y solo si } \sigma(x) <_1 \sigma(y)$$

$$x <_3 y \text{ si y solo si } \sigma(\sigma(x)) <_1 \sigma(\sigma(y))$$

Ahora que tenemos definidos dos nuevos ordenes en cualquier conjunto ordenado con una permutación, resulta tentador trasladar los resultados de la sección anterior sobre conjuntos aleatorios con tres ordenes a conjuntos aleatorios con un orden y una permutación. Pero el traslado no puede ser inmediato por el siguiente problema: Si elegimos $\sigma \in S_n$ de manera uniforme, el orden $<_3$ **no** esta elegido uniformemente entre todos los ordenes posibles, ni es elegido de forma independiente a $<_2$. Es fácil ver este problema para conjuntos pequeños: Como toda permutación de un conjunto de dos elementos cumple que $\sigma^2 = Id$, se sigue que si $n = 2$ siempre sera el caso que $<_3 = <_1$.

Sin embargo, vamos a demostrar que los ordenes $<_2$ y $<_3$ si se comportan **localmente** como si fueran elegidos de forma uniforme e independiente entre todos los ordenes, lo que sera suficiente para trasladar los resultados de la sección anterior.

Lema 4.57. Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación elegida de forma uniforme y $B \subset \{1, \dots, n\}$ un conjunto de tamaño $o(\sqrt{n})$. Luego la probabilidad de que $B, \sigma(B)$ y $\sigma(B^2)$ sean disjuntos tiende a 1 cuando n tiende a infinito.

Demostración. Veamos que la probabilidad de que esos conjuntos no sean disjuntos, que denotaremos $p_{n,B}$, tiende a 0 cuando n tiende a infinito.

$$\begin{aligned} p_{n,B} &= Pr\left(\bigcup_{x,y \in B} \sigma(x) = y \vee \sigma^2(x) = y \vee \sigma^2(x) = \sigma(y)\right) \\ &\leq \sum_{x,y \in B} Pr(\sigma(x) = y) + Pr(\sigma^2(x) = y) + Pr(\sigma^2(x) = \sigma(y)) \end{aligned}$$

Luego:

$$P_{n,B} \leq 3 \cdot |B|^2 \cdot \frac{1}{n} \rightarrow 0$$

□

Lema 4.58. Sea $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una permutación elegida de forma uniforme y $B \subset \{1, \dots, n\}$ un conjunto de tamaño $o(\sqrt{n})$. Luego, cuando n tiende a infinito, las distribuciones de los ordenes \langle_2, \langle_3 **sobre** B tienden a la distribución de tres ordenes elegidos de forma independiente y uniforme.

Demostración. Sea $\mathcal{B} = \{\sigma \in S_n : B, \sigma(B), \sigma^2(B)\}$. Como para todo $\sigma \in \mathcal{B}$ los ordenes $\langle_1, \langle_2, \langle_3$ sobre B están controlados por conjuntos disjuntos, es fácil ver que tienen distribución uniforme e independiente entre si. Por otro lado, sabemos por el lema anterior que cuando n tiende a infinito, la probabilidad de que un $\sigma \in S_n$ elegido de forma uniforme este en \mathcal{B} tiende a 1. Luego, las distribuciones de $\langle_1, \langle_2, \langle_3$ **sobre** B tienden a la distribución de tres ordenes elegidos de forma independiente y uniforme. □

Definición 4.59. Sea M una maquina de Turing y sea (A, \langle_1, σ) un conjunto equipado con un orden y una permutación σ . Decimos que A simula a M por k pasos si $(A, \langle_1, \langle_2, \langle_3)$ simula a M por k pasos.

Lema 4.60. Sea $c > 0$, M una maquina de Turing y sea $n = c \cdot (k+1)^{k^2} \cdot \ln k^k$. Luego, cuando k tiende a infinito, la probabilidad de que $A(n, \langle, \sigma)$ simule a M por k pasos tiende a 1 y la probabilidad de que $A(n, \langle, \sigma)$ simule a M por $k+1$ pasos tiende a 0.

Demostración. La idea es usar los lemas anteriores para emular la demostración 4.48. Dado $k > 0$, sea X_k la variable aleatoria que cuenta la cantidad de tuplas (b, c) tal que $A \models \phi_M(k+1, b, c)$. $X_k = \sum 1_{A_i}$, donde A_i es una enumeración de todas las tuplas (b, c) tal que $|(b, c)| = k^2$ y A_i es el evento $(A, \langle, \sigma) \models \phi_M(k+1, b, c)$. Sabemos, por la demostración de 4.48 que si los ordenes $\langle_1, \langle_2, \langle_3$ son elegidos de forma uniforme e independiente la probabilidad de todo A_i es $(k^2+1)^{-k^2} \cdot 2^{-\lambda k^2}$. Por otro lado, como $c \cdot (k+1)^{k^2} \cdot \ln k^k \gg k^2$, por el lema 4.58 cuando k tiende a infinito la probabilidad de que \langle_2, \langle_3 se distribuyan como ordenes elegidos de manera uniforme tiende a 1. Luego, para todo k suficientemente grande la probabilidad de A_i es mayor a $\frac{1}{2} \cdot (k^2+1)^{-k^2} \cdot 2^{-\lambda k^2}$. En particular, si $n =$ tenemos:

$$\mathbb{E}(X_k) = (n - k^2) \frac{1}{2} \cdot (k^2+1)^{-k^2} \cdot 2^{-\lambda k^2} = (n = c \cdot (k+1)^{k^2} \cdot \ln k^k) \frac{1}{2} \cdot (k^2+1)^{-k^2} \cdot 2^{-\lambda k^2}$$

Por otro lado, por la misma argumentación que en 4.20 llegamos a que:

$$\mathbb{E}(X_k^2) \leq \mathbb{E}(X_k)^2 + \mathbb{E}(X_k)$$

Y por tanto:

$$\sigma(X_k) \leq \sqrt{\mathbb{E}(X_k)}$$

Luego, si $n =$ tenemos por Chebyshev:

$$Pr(X_k = 0) \leq Pr(|X_k - \mu| \geq \mu) \leq \frac{\sigma(X_k)^2}{\mu(X_k)^2} \leq \frac{1}{\mu(X_k)} \rightarrow 0$$

Por otro lado, por Markov:

$$Pr(X_{k+1} > 0) \leq \mathbb{E}(X_{k+1}) \longrightarrow 0$$

□

Proposición 4.61. *No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de conjuntos con un orden y una función, si vale que $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <, \sigma) \models \phi) = 1$.*

Demostración. análoga a la demostración 4.21. Dada maquina de Turing M , podemos construir una oración ϕ_M tal que $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <, \sigma) \models \phi_M) = 1$ si y solo si la maquina M llega al estado de aceptación empezando con la cinta vacía. □

Proposición 4.62. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de conjuntos con un orden y una función, si vale que el limite $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(n, <, \sigma) \models \phi)$ existe.*

Demostración. análoga a la demostración 4.22. □

Proposición 4.63. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de los conjuntos con tres ordenes, si hay existe un conjunto finito $(A, <_1, <_2, <_3)$ tal que $A \models \phi$.*

Demostración. análoga a la demostración 4.23. □

Proposición 4.64. *Sean $\alpha, \beta \in (0, 1)$ reales computables tal que $\alpha < \beta$. Luego, existe una oración ϕ en lenguaje de conjuntos con un orden y una función. $\liminf Pr(\mathbf{A}(n, <, \sigma) \models \phi) = \alpha$ y $\liminf Pr(\mathbf{A}(n, <, \sigma)) \models \phi = \beta$.*

Demostración. Análogo a la demostración 4.25. □

Proposición 4.65. *No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración ϕ en el lenguaje de conjuntos con un orden y una función valga que si $\lim_{n \rightarrow \infty} Pr(\mathbf{A}(N, <, \sigma) \models \phi) = 1$ se sigue que $Pr(\mathbf{A}(N, <, \sigma) \models \phi) > 0.95$ para todo $N > f(qr(\phi))$.*

Demostración. Análogo a la demostración 4.26. De existir, tendríamos un algoritmo para determinar cuando una oración dada sobre el lenguaje de conjuntos con tres ordenes tiene modelo finito, pero demostramos anteriormente que tal algoritmo no existe. □

4.2.3. Conjuntos ordenados con funciones

Vamos a considerar ahora conjuntos equipados con un orden y una función, que denotaremos $(A, <, f)$. Denotamos con $\mathbf{A}(n, <, f)$ la variable aleatoria que toma valores de manera uniforme entre conjuntos ordenados sobre $\{1, \dots, n\}$ equipados con una función $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$.

Al igual que con los conjuntos ordenados equipados con una permutación, podemos definir en $\mathbf{A}(n, <, f)$ las relaciones $<_2$ y $<_3$, pero ahora serán ordenes parciales, no totales. Esto se debe a que las función f es casi seguro no inyectiva y si $f(x) = f(y)$ se sigue que x e y son incomparables por el orden $<_2$.

Definición 4.66. Sea A un conjunto equipado de un orden total $<_1$ y una función $f : A \rightarrow A$. Definimos los ordenes parciales $<_2$ y $<_3$ sobre A como:

$$x <_2 y \text{ si y solo si } f(x) <_1 f(y)$$

$$x <_3 y \text{ si y solo si } f(f(x)) <_1 f(f(y))$$

Al igual que con conjuntos ordenados con una permutación mostraremos que aunque $<_2$ y $<_3$ no son ordenes totales, localmente se comportan como ordenes totales elegidos de manera independiente y uniforme, lo que nos permitirá trasladar los resultados anteriores de conjuntos con tres ordenes totales.

Lema 4.67. Sea $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una función elegida de forma uniforme y $B \subset \{1, \dots, n\}$ un conjunto de tamaño $o(\sqrt{n})$. Luego la probabilidad de que B , $f(B)$ y $f^2(B)$ sean disjuntos tiende a 1 cuando n tiende a infinito.

Demostración. Veamos que la probabilidad de que esos conjuntos no sean disjuntos, que denotaremos $p_{n,B}$, tiende a 0 cuando n tiende a infinito.

$$\begin{aligned} p_{n,B} &= Pr\left(\bigcup_{x,y \in B} f(x) = y \vee f^2(x) = y \vee f^2(x) = f(y)\right) \\ &\leq \sum_{x,y \in B} Pr(f(x) = y) + Pr(f^2(x) = y) + Pr(f^2(x) = f(y)) \end{aligned}$$

Fijados x, y , la probabilidad de que $f(x) = y$, $f(f(x)) = y$ o $f(f(x)) = f(y)$ es $\frac{1}{n}$, luego nos queda:

$$P_{n,B} \leq 3 \cdot |B|^2 \cdot \frac{1}{n} \rightarrow 0$$

□

Lema 4.68. Sea $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una función elegida de forma uniforme y $B \subset \{1, \dots, n\}$ un conjunto de tamaño $o(\sqrt{n})$. Luego, cuando n tiende a infinito, la probabilidad de que la función f restringida a B sea inyectiva tiende a 1.

Demostración. Denotemos con p la probabilidad de que f restringida a B sea inyectiva. Tenemos que:

$$p = Pr\left(\bigcup_{x,y \in B} f(x) = f(y)\right) \leq \sum_{x,y \in B} Pr(f(x) = f(y)) = \frac{|B|^2}{n} \rightarrow 0$$

□

Observación 4.69. El lema anterior es el mejor posible. Es fácil ver, usando la desigualdad de Chebyshev, que si B es de tamaño $\Omega(\sqrt{n})$ entonces la probabilidad de que f restringida a B sea inyectiva tiende a 0 cuando n tiende a infinito.

Lema 4.70. Sea $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ una función elegida de forma uniforme y $B \subset \{1, \dots, n\}$ un conjunto de tamaño $o(\sqrt{n})$. Luego, cuando n tiende a infinito, las distribuciones de los ordenes parciales $<_1, <_2, <_3$ sobre B tienden a la distribución de tres ordenes elegidos de forma independiente y uniforme.

Demostración. Sea \mathcal{B} el conjunto de todas las funciones $f : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ tal que $B, f(B)$ y $f^2(B)$ son disjuntos entre si y tal que f restringida a B y a $f(B)$ es inyectiva. Por los dos lemas anteriores se sigue que cuando n tiende a infinito la probabilidad de que una función f elegida de forma uniforme este en B tiende a 1.

Para todo $f \in \mathcal{B}$, por la inyectividad de f , los ordenes parciales $<_2$ y $<_3$ son ordenes totales sobre B . Además, como $<_1, <_2$ y $<_3$ sobre B estan determinados por los conjuntos disjuntos $B, f(B)$ y $f^2(B)$, es fácil ver que $<_1, <_2, <_3$ tienen una distribución uniforme e independiente entre todos los ordenes totales de B □

Proposición 4.71. *No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de conjuntos equipados con un orden y una función, si vale que $\lim_{n \rightarrow \infty} Pr(A(n, <, f) \models \phi) = 1$*

Demostración. análoga a la demostración 4.21. Dada maquina de Turing M , podemos construir una oración ϕ_M tal que $\lim_{n \rightarrow \infty} Pr(A(n, <, f) \models \phi_M) = 1$ si y solo si la maquina M llega al estado de aceptación empezando con la cinta vacia. □

Proposición 4.72. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de conjuntos equipados con un orden y una función, si vale que el limite $\lim_{n \rightarrow \infty} Pr(A(n, <, f) \models \phi)$ existe*

Demostración. Análoga a la demostración 4.22 □

Proposición 4.73. *No hay ningún algoritmo que pueda determinar, dada fórmula ϕ en el lenguaje de conjuntos equipados con un orden y una función, si hay existe un conjunto finito $A(n, <, f)$ tal que $A \models \phi$*

Demostración. Análoga a la demostración 4.23. □

Proposición 4.74. *Sean $\alpha, \beta \in (0, 1)$ reales computables tal que $\alpha < \beta$. Luego, existe una oración ϕ en en lenguaje de conjuntos equipados con un orden y una función tal que $\lim inf Pr(A(n, <, f) \models \phi) = \alpha$ y $\lim inf Pr(A(n, <, f) \models \phi) = \beta$.*

Demostración. Análoga a la demostración 4.25 □

Proposición 4.75. *No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración ϕ en el lenguaje de conjuntos lenguaje de conjuntos equipados con un orden y una función valga que si $\lim_{n \rightarrow \infty} Pr(A(n, f, <) \models \phi) = 1$ se sigue que $Pr(A(n, f, <) \models \phi) > 0.95$ para todo $N > f(qr(\phi))$.*

Demostración. Análoga a la demostración 4.26. De existir, tendríamos un algoritmo para determinar cuando una oración dada sobre el lenguaje de conjuntos con tres ordenes tiene modelo finito, pero demostramos anteriormente que tal algoritmo no existe. □

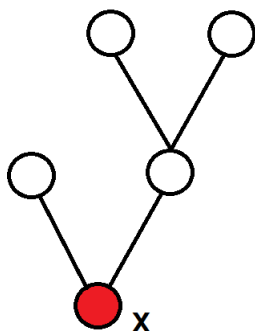
4.2.4. Grafos sparse ordenados

En esta sección estudiaremos las propiedades de grafos sparse aleatorios equipados con un orden, que denotaremos $G_{<}(n, \frac{c}{n})$. Demostramos en 4.40 que cumplen una ley de poca variación, veremos ahora que cumplen las mismas propiedades de no convergencia y no decidibilidad que los grafos ordenados.

Empecemos notando que todas las demostraciones que hicimos para grafos aleatorios ordenados **no** se pueden trasladar a los grafos sparse: Por ejemplo, si ϕ es la oración que codificaba la maratón de danza acotada, aunque vimos que el límite $Pr(G_{<}(n) \models \phi)$ no existía, es fácil ver que $\lim_{n \rightarrow \infty} Pr(G_{<}(n, \frac{c}{n}) \models \phi) = e^{-c} \cdot c$. De forma similar, se puede ver que las oraciones que usamos para codificar el comportamiento de máquinas de Turing en grafos aleatorios tampoco funcionan para grafos sparse.

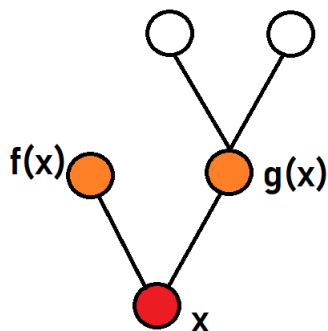
Lo que haremos para demostrar los resultados, será dado un grafo aleatorio sparse definir otros dos ordenes sobre cierto subconjunto del grafo y veremos que es posible trasladar los resultados de la sección anterior sobre conjuntos con tres ordenes a este nuevo contexto.

Definición 4.76. Sea G un grafo. Decimos que un vértice $x \in G$ es **bueno** si $B_4(x)$ es el siguiente árbol:



Observación 4.77. La propiedad de ser un vértice bueno es definible por una oración en la lógica de primer orden de grafos ordenados.

Definición 4.78. Sea G un grafo ordenado y sea $\mathcal{B} \subset G$ el conjunto de los vértices buenos. Cada vértice bueno x tiene un solo vecino sin ningún otro vecino, que denotaremos $f(x)$ y un vecino con un solo otro vecino, que denotaremos $g(x)$.



Definimos los ordenes \leq_2 y \leq_3 sobre \mathcal{B} de la siguiente manera:

$$x \leq_2 x' \text{ si y solo si } f(x) \leq f(x')$$

$$x \leq_3 x' \text{ si y solo si } g(x) \leq g(x')$$

Observación 4.79. Los ordenes \leq_2 y \leq_3 definidos anteriormente se pueden definir como fórmulas de primer orden en el lenguaje de los grafos ordenados. Explícitamente:

$$x \leq_2 x' = [\exists y \exists y' : (\forall z \ y \sim z \Leftrightarrow z = x) \wedge (\forall z \ y' \sim z \Leftrightarrow z = x') \wedge (y \leq y')]$$

$$x \leq_3 x' = [\exists y \exists y' : (\exists z \neq x : y \sim z \wedge y \sim x) \wedge (\exists z \neq x' : y' \sim z \wedge y' \sim x') \wedge (y \leq y')]$$

Observación 4.80. Si $G \sim G_{<}(n, \frac{c}{n})$ es un grafo sparse ordenado aleatorio y $\mathcal{B} \subset G$ es el conjunto de los vértices buenos, $(\mathcal{B}, \leq, \leq_2, \leq_3)$ tiene distribución uniforme entre los conjuntos con tres ordenes.

Proposición 4.81. Sea $c > 0$ y sea $G \sim G(n, \frac{c}{n})$ un grafo aleatorio sparse. Sea $\mathcal{B} \subset G$ el subconjunto de los vértices buenos. Luego, cuando n tiende a infinito, casi seguro G se cumple que:

$$\frac{1}{2} \cdot c^4 \cdot e^{-5c} \cdot n < |\mathcal{B}| < \frac{3}{2} \cdot c^4 \cdot e^{-5c} \cdot n$$

Demostración. Consecuencia directa de la proposición 3.46 del capítulo anterior, □

Definición 4.82. Sea M una maquina de Turing probabilística y G un grafo ordenado. Decimos que sus vértices buenos codifican M por k pasos si $(\mathcal{B}, \leq, \leq_2, \leq_3)$ codifica M por k pasos.

Lema 4.83. Sea M una maquina de Turing probabilística, sea $n = (k+1)^{k^2} \cdot (\ln k)^k$ y sea $G \sim G_{<}(n, \frac{c}{n})$ un grafo sparse ordenado aleatorio. Luego, cuando k tiende infinito, casi seguro los vértices buenos de G codifican a M por k pasos y casi seguro los vértices buenos de G no codifican a M por $k+1$ pasos.

Demostración. Por el lema 3.46, cuando n tiende a infinito casi seguro la cantidad de vértices buenos de G esta entre $\frac{1}{2} \cdot c^4 \cdot e^{-5c} \cdot n$ y $\frac{3}{2} \cdot c^4 \cdot e^{-5c} \cdot n$. Por otro lado, por el lema 4.48 para todo $\lambda > 0$, cuando k tiende a infinito casi seguro $A((k+1)^{k^2} \cdot (\ln k)^k, <_1, <_2, <_3)$ codifica a M por $k+1$ pasos y no codifica a M por $k+1$ pasos. □

Con ese ultimo lema, podemos demostrar analogos de todos los teoremas para conjuntos con tres ordenes:

Proposición 4.84. Sea $c > 0$. No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de grafos ordenados, si vale que $\lim_{n \rightarrow \infty} Pr(G_{<}(n, \frac{c}{n}) \models \phi) = 1$

Demostración. análoga a la demostración 4.21. □

Proposición 4.85. *Sea $c > 0$. No hay ningún algoritmo que pueda determinar, dada oración ϕ en el lenguaje de grafos ordenados, si vale que el limite $\lim_{n \rightarrow \infty} Pr(G_{<}(n, \frac{c}{n}) \models \phi)$ existe.*

Demostración. análoga a la demostración 4.22. □

Proposición 4.86. *Sea $c > 0$ y sean $\alpha, \beta \in (0, 1)$ reales computables tal que $\alpha < \beta$. Luego, existe una oración ϕ en lenguaje de grafos ordenados tal que $\liminf Pr(G_{<}(n, \frac{c}{n}) \models \phi) = \alpha$ y $\limsup Pr(G_{<}(n, \frac{c}{n}) \models \phi) = \beta$.*

Demostración. análoga a la demostración 4.25 □

Proposición 4.87. *Sea $c > 0$. No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración ϕ en el lenguaje de grafos ordenados valga que si $\lim_{n \rightarrow \infty} Pr(G_{<}(n, \frac{c}{n}) \models \phi) = 1$ se sigue que $Pr(G_{<}(N, \frac{c}{N}) \models \phi) > 0.95$ para todo $N > f(qr(\phi))$.*

Demostración. análoga a la demostración 4.26. □

4.3. Lógica de segundo orden monádica sobre grafos

En el capítulo 2 demostramos la ley 0-1 de Fagin para la lógica de primer orden de grafos y mas en general, para las lógicas infinitarias $L_{\infty, w}^k$. En ese mismo capítulo mostramos que la ley 0-1 que **no** se puede extender a la lógica de segundo orden de grafos, ya que la lógica de segundo orden puede expresar la propiedad "El grafo tiene una cantidad par de aristas", que tiene probabilidad asintótica $\frac{1}{2}$. Mas aun, como la lógica de segundo orden es capaz de cuantificar sobre ordenes de grafos, se sigue de los resultados de la primer sección de este capítulo que ni siquiera seguirá una ley de convergencia.

Sin embargo, queda abierta la pregunta de si vale una ley 0-1 o de convergencia para fragmentos de la lógica de segundo orden, como la lógica de segundo orden monádica. La respuesta es negativa. En [25] Shelah y Kaufmann demostraron que la ley 0-1 no vale para lógica de segundo monádica sobre grafos dirigidos. En [24] Kaufmann mostró que ni siquiera vale para la lógica de segundo orden monádica existencial (EMSO), aunque su demostración era sobre estructuras con cuatro relaciones binarias. Por ultimo, en [27] Le Bars refinó la demostración de Kaufmann y demostró que la ley 0-1 no se cumple para la lógica EMSO sobre grafos.

Presentaremos a continuación la demostración de Le Bars, con ciertos cambios de notación. Vamos a demostrar que hay una oración en la lógica EMSO de grafos que define, con probabilidad que tiende a 1 cuando el tamaño del grafo tiende a infinito, un orden total sobre el grafo y eso nos permitirá trasladar los resultados anteriores sobre no convergencia y no decibilidad sobre grafos ordenados a la lógica EMSO.

Empecemos con algunas definiciones.

Definición 4.88. *Sea G un grafo, $A \subset G$ y $b \in G$. Definimos el subconjunto de A codificado por b , denotado A_b , como $A_b = \{a \in A : a \sim b\}$*

Definición 4.89. Sea G un grafo y $A, B \subset G$ subconjuntos disjuntos. Decimos que B **codifica subconjuntos distintos** de A si vale que:

- A_b no es un singleton para ningún $b \in B$
- $A_b \neq A_{b'}$ para todo par de elementos distintos $b, b' \in B$

Definición 4.90. Sea G un grafo y $A, B \subset G$ subconjuntos disjuntos. Decimos que B **codifica el powerset** de A si codifica conjuntos distintos de A y para todo $S \subset A$ no singleton existe un $b \in B$ tal que $A_b = S$

Observación 4.91. Las propiedades " B codifica conjuntos distintos de A " y " B codifica el powerset de A " pueden ser expresadas por formulas en lógica de segundo orden monádica de grafos.

Definición 4.92. Sea G un grafo, $A, B \subset G$ subconjuntos disjuntos y sea R una relación en A . Para cada elemento $x \in A \cup B$ podemos asociar un subconjunto de A , denotaremos A_x . Para los $b \in B$ este sera el subconjunto de A codificado por B , mientras que para $a \in A$ no sera mas que el singleton $\{a\}$. Definimos la **extensión** de R a $A \cup B$, denotado R' como la siguiente relación en $A \cup B$:

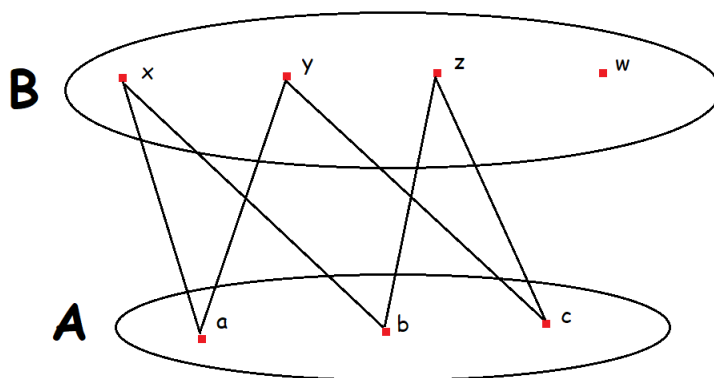
$R'(x, y)$ si y solo si existe $z \in A_x$ tal que $z \notin A_y$ y para todo $w \in A_y$ si $R(w, z)$ se sigue que $w \in A_x$.

La importancia de la definición anterior radica en la siguiente propiedad:

Proposición 4.93. Sea G un grafo, $A, B \subset G$ tal que B codifica subconjuntos distintos de A y R un orden total en A . Luego, su extensión R' es un orden total en $A \cup B$ que extiende al orden de A .

Demostración. El orden R' no es mas que el orden lexicografico sobre subconjuntos de A , que extiende el orden R . □

Ejemplo 4.94. Consideremos un grafo con los subconjuntos A y B como muestra la figura



Es fácil ver que B codifica subconjuntos distintos de A . Supongamos que tenemos un orden total en A dado por:

$$a < b < c$$

Luego se extiende a un orden total en $A \cup B$ dado por:

$$w < a < b < x < c < y < z$$

Observación 4.95. Sea G un grafo y sean $A, B \subset G$ subconjuntos disjuntos. Si un orden R sobre A es definible por una fórmula de primer orden de grafos, entonces su extensión R' a $A \cup B$ también es definible por una fórmula de primer orden.

Veamos ahora una serie de lemas técnicos que usaremos para demostrar el enunciado principal.

Lema 4.96. Sea G un grafo aleatorio, $A, B \subset G$ subconjuntos disjuntos tal que $|A| = k$ y $|B| = 2^{k/3}$. Luego, cuando $k \rightarrow \infty$, la probabilidad de que B codifique conjuntos distintos de A tiende a 1.

Demostración. Sea S el evento " B no codifica subconjuntos distintos de A ". Veamos que la probabilidad de S tiende a 0 cuando k tiende a infinito. Por definición:

$$Pr(S) = Pr\left(\bigcup_{b, b' \in B} A_b = A_{b'} \cup \bigcup_{b \in B, a \in A} A_b = \{a\}\right) \leq \sum_{b, b' \in B} Pr(A_b = A_{b'}) + \sum_{b \in B, a \in A} Pr(A_b = \{a\})$$

Es fácil ver que los eventos $A_b = A_{b'}$ y $A_b = \{a\}$ tienen probabilidad 2^{-k} luego:

$$Pr(S) \leq \binom{2^{k/3}}{2} \cdot 2^{-k} + 2^{k/3} \cdot 2^{-k} < 2^{\frac{2k}{3}} \cdot 2^{-k} = 2^{-\frac{k}{3}} \rightarrow 0$$

□

Lema 4.97. Sea G un grafo aleatorio de mas de 3^k elementos y $A \subset G$ un subconjunto de k elementos. Luego, cuando $k \rightarrow \infty$ la probabilidad de que exista un $B \subset G$ que codifique el powerset de A tiende a 1.

Demostración. Basta demostrar que cuando $k \rightarrow \infty$ la probabilidad de que para todo subconjunto no singleton $S \subset A$ exista un $x \in G - A$ que lo codifica tiende a 1, luego el conjunto B que codifica el powerset de A sera la unión de dichos elementos.

Fijemos un $S \subset A$, no singleton. La probabilidad de que un $x \in G$ codifique a S es 2^{-k} y esos eventos son independientes entre si, luego la probabilidad de que ningún x codifique a S es $(1 - 2^{-k})^{3^k - k}$. Luego tenemos:

$$Pr(\text{ningún } x \text{ codifica a } S) = (1 - 2^{-k})^{3^k - k} = (1 - 2^{-k})^{\frac{3^k - k}{3^k - k}} \rightarrow e^{-\frac{3^k - k}{2^k}}$$

$$Pr\left(\bigcup_{S \subset A} \text{ningún } x \text{ codifica a } S\right) \leq \sum_{S \subset A} Pr(\text{ningún } x \text{ codifica a } S) = 2^k \cdot e^{-\frac{3^k - k}{2^k}} = e^{(-\frac{3}{2})^k} \cdot 2^k \cdot e^{\frac{k}{2^k}}$$

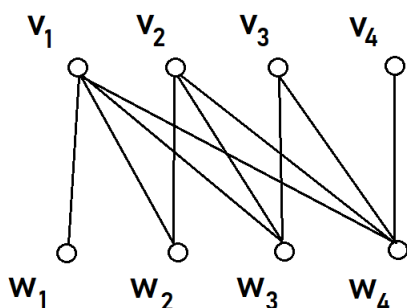
Y es fácil ver que este último término tiende a 0 cuando k tiende a infinito, ya que $e^{(-\frac{3}{2})^k}$ decrece superexponencialmente.

□

Lema 4.98. Existe una fórmula $\phi_1(x, y, z)$ en la lógica de primer de grafos tal que para todo k la probabilidad de que en un grafo aleatorio $G(n)$ haya un subconjunto $A \subset G$ de tamaño k y un $a \in A$ tal que la relación binaria $R(y, z) = \phi_1(a, y, z)$ sea un orden total en A tiende a 1 cuando $n \rightarrow \infty$.

Demostración. Vamos a ver que existe una fórmula $\phi(x, y, z)$ en la lógica de primer orden de grafos, una familia de grafos G_k de tamaño $2k$ y un $g_k \in G_k$ tal que para todo k se cumple que la relación binaria $R(y, z) = \phi(g_k, y, z)$ define un orden total en G_k . Como sabemos que asintóticamente todo grafo G_k pertenece a $G(n)$, se sigue el enunciado.

La fórmula $\phi(x, y, z)$ sera $[\neg(x \sim y) \wedge (x \sim z)] \vee [\forall w (w \sim x) \implies (w \sim y)]$, es decir, se cumple si z es vecino de x e y no; o si todos los vecinos de y son vecinos de z . Los grafos G_k seran grafos bipartitos, con vértices $\{v_1, \dots, v_k, w_1, \dots, w_k\}$ y aristas $v_i \sim w_j$ si y solo si $i \geq j$. Es fácil ver que la relación binaria $\phi(v_1, y, z)$ induce en G_k el orden total dado por $v_1 > v_2 > (\dots) > v_k > w_k > (\dots) > w_2 > w_1$



□

Proposición 4.99. [27] Hay una fórmula $\psi(X, Y, x, y, z)$ en la lógica EMSO de grafos tal que si $G \sim G(n)$ la probabilidad de que existan subconjuntos $A, B \subset G$ y vértice $a \in A$ tal que $R''(x, y) = \psi(A, B, a, x, y)$ defina un orden total en G tiende a 1 cuando $n \rightarrow \infty$.

Demostración. Sea $R(x, y)$ la relación en A dada por $\phi_1(a, x, y)$, donde ϕ_1 es la fórmula definida en el lema 4.98. Sea $R'(x, y)$ la extensión de R a $A \cup B$ y sea $R''(x, y)$ la extensión de R' a todo G . Como observamos anteriormente, R'' es definible por una fórmula EMSO. Veamos que cuando n tiende a infinito, casi seguro existen $A, B \subset G$ y $a \in G$ tal que R'' define un orden total sobre G .

Por el lema 4.98, casi seguro hay un subconjunto A de tamaño $\log_2(n)$ y un $a \in A$ en le que $R(x, y)$ es un orden total. En particular, casi seguro hay un conjunto A de tamaño $\frac{\log_2(n)}{2}$ en el que R es un orden total. Luego, como $3^{\frac{\log_2(n)}{2}} < n$ por el lema 4.97, casi seguro hay un conjunto B que codifica el powerset de A . Como B codifica subconjuntos distintos de A , se sigue que R' define un orden total sobre $A \cup B$. Por ultimo, como B tiene $\sim 2^{\frac{\log_2(n)}{2}} = \sqrt{n}$ elementos, por el lema 4.96 casi seguro $G - (A \cup B)$ codifica subconjuntos distintos de B , luego R'' define un orden total sobre G . □

Ahora que tenemos una fórmula que define casi seguro un orden sobre el grafo, es muy fácil imitar los resultados de indecidibilidad y no convergencia que teníamos para grafos ordenados.

Proposición 4.100. [27] Hay una oración ϕ en la lógica EMSO de grafos que expresa asintóticamente la propiedad "Hay una cantidad par de vértices".

Demostración. Sea $\psi(X, Y, x, y, z)$ la fórmula definida en la proposición 4.99 y sea $R(x, y) = \psi(A, B, a, x, y)$. Definamos ϕ como la oración:

$$\exists A \exists B \exists a \exists C : (R \text{ define orden total}) \wedge (\forall x (\forall y R(x, y)) \rightarrow (x \in C))$$

$$\wedge (\forall x (\forall y R(y, x)) \rightarrow (x \notin C)) \wedge (\forall x (x \in C) \leftrightarrow (x^+ \notin C))$$

Es decir, postula la existencia de subconjuntos $A, B, C \subset G$ y $a \in G$ tal que $R(x, y) = \psi(A, B, a, x, y)$ define un orden total, C contiene al elementos mínimos y no al máximo respecto a ese orden, y para todo par de sucesores inmediatos respecto al orden exactamente uno esta en C . Es fácil ver que, dado un orden total, podemos encontrar un conjunto C de dichas características si y solo si el tamaño del conjunto es par: si numeramos al conjunto ordenado como $\{1, \dots, 2n\}$ el conjunto C sera $\{1, 3, \dots, 2n - 1\}$.

Por otro lado, sabemos por la proposición 4.99 que cuando $n \rightarrow \infty$ casi seguro existen $A, B \subset G$ tal que R define un orden total. Luego tenemos que $\lim_{n \rightarrow \infty} Pr(G(2n) \models \phi) = 1$ y $\lim_{n \rightarrow \infty} Pr(G(2n + 1) \models \phi) = 0$, es decir, ϕ codifica asintóticamente la propiedad " G tiene una cantidad par de vértices". \square

Corolario 4.101. Existe una oración en la lógica EMSO de grafos tal que el límite $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi)$ **no** existe.

Teorema 4.102. [27] No hay ningún algoritmo que pueda determinar, dada oración ϕ en la lógica EMSO de grafos, si vale que $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi) = 1$.

Demostración. Vamos a ver que de existir tal algoritmo, existirá un algoritmo para determinar si una oración en la lógica de primer orden de grafos ordenados tiene modelo finito, pero como demostramos en 4.23 que eso es imposible, se sigue que no existe tal algoritmo.

Sea ϕ una oración en la lógica de primer orden de grafos ordenados. Sea $\psi(X, Y, x, y, z)$ la oración definida en 4.99 y sea $R(x, y) = \psi(A, B, a, x, y)$. Sea ϕ^* la oración en lógica EMSO de grafos dada por:

$$\phi^* = \exists A \exists B \exists a \exists Z : (R \text{ define un orden total}) \wedge \phi_Z$$

Donde ϕ_Z es la oración ϕ restringida al conjunto Z y usando R como orden. Si la oración ϕ no tiene modelo finito, ϕ_Z sera siempre falsa y por tanto $Pr(G(n) \models \phi^*) = 0$ para todo n . Por otro lado, supongamos que la oración ϕ tiene un modelo finito H . Por la proposición 4.99 sabemos que casi seguro en todo grafo suficientemente grande existen A, B y elemento a tal que R es un orden. Además, sabemos que casi seguro todo grafo suficientemente grande tiene a H como subconjunto. Se sigue que entonces $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi^*) = 1$. \square

Teorema 4.103. No hay ningún algoritmo que pueda determinar, dada oración ϕ en la lógica EMSO de grafos, si el límite $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi)$ existe.

Demostración. Vamos a ver que de existir tal algoritmo, existirá un algoritmo para determinar si una oración en la lógica de primer orden de grafos ordenados tiene modelo finito, pero como demostramos en 4.23 que eso es imposible, se sigue que no existe tal algoritmo.

En la demostración anterior construimos, dada oración ϕ en la lógica de primer orden de grafos ordenados, una oración ϕ^* en la lógica EMSO de grafos tal que $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi^*) = 1$ si y solo si ϕ tiene modelo finito. Además, en la proposición ? construimos una oración θ en la lógica EMSO de grafos que codifica asintóticamente la propiedad "El grafo tiene una cantidad par de vértices". Se sigue que $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi^* \vee \theta)$ existe si y solo si ϕ tiene modelo finito. \square

Proposición 4.104. *No existe ninguna función computable $f : \mathbb{N} \rightarrow \mathbb{N}$ tal que toda oración EMSO ϕ en el lenguaje de grafos ordenados valga que si $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = 1$ se sigue que $Pr(G_{<}(N) \models \phi) > 0.95$, para $N > f(qr(\phi))$*

Demostración. Vamos a ver que de existir tal f , existirá un algoritmo para determinar si una oración ϕ dada en la lógica de primer orden de grafos ordenados tiene modelo finito, que demostramos anteriormente que es un problema indecidible. Vimos en la demostración 4.102 que dada oración ϕ en la lógica de primer orden de grafos ordenados, podemos construir una oración ϕ^* en la lógica EMSO de grafos tal que si ϕ tiene modelo finito entonces $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi^*) = 1$ pero si ϕ no tiene modelo finito entonces para todo n vale que $Pr(G(n) \models \phi^*) = 0$. Luego, para determinar si ϕ tiene modelo finito bastaría con calcular $Pr(G(f(N)) \models \phi^*)$ con $N = qr(\phi)$, que se puede hacer en tiempo finito. \square

Un resultado que no podemos trasladar directamente a la lógica EMSO de grafos, sin embargo, es la proposición 4.24, que afirmaba que para cualquier real computable $0 < \alpha < 1$ existe una oración ϕ en el lenguaje de los grafos ordenados tal que $\lim_{n \rightarrow \infty} Pr(G_{<}(n) \models \phi) = \alpha$. La razón es que aunque podemos trasladar dicha oración a una oración EMSO del estilo $\phi^* = \exists A \exists B \exists a : (R \text{ define un orden } \wedge \phi)$, su probabilidad límite será 1, no α , ya que en todo grafo suficientemente grande hay **muchos** pares (A, B, a) que hacen que la relación sea un orden, luego casi seguro existe uno que induzca un orden en el que se cumpla ϕ . Mas aun, hasta ahora no encontramos ninguna oración en la lógica EMSO de grafos con probabilidad límite distinta de 0 o 1, ni hay ninguna mencionada en la literatura. Eso nos lleva a la siguiente conjetura:

Conjetura 4.105. Para toda oración ϕ en la lógica EMSO de grafos se cumple una de las siguientes:

- $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi) = 0$
- $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi) = 1$
- $\lim_{n \rightarrow \infty} Pr(G(n) \models \phi)$ no existe.

Conclusiones

Este trabajo presenta un recorrido amplio por el t3pico de las propiedades l3gicas de estructuras aleatorias, iniciado por el descubrimiento de la ley 0-1 de Fagin. Uno de los pocos resultados importantes que por falta de tiempo no cubrimos son las investigaciones de Shelah y Spencer[37] sobre las propiedades l3gicas de strings aleatorias, que demostraron que cumplen una ley de convergencia para l3gica de primer orden, una ley de convergencia mas d3bil para l3gica de segundo orden mon3dica y tienen interesantes conexiones con teor3a de aut3matas. Incluimos una tabla de las contribuciones originales:

Contribuci3n	Referencia
Cotas para la convergencia de la ley 0-1 de Fagin	Secci3n 2.1.2.
Cotas para la convergencia de la ley 0-1 de Obserchelp	Secci3n 2.2.3.
Nueva demostraci3n de la ley de convergencia de Lynch	Teorema 3.55.
Cota inferior para la velocidad de convergencia de la ley de Lynch	Teorema 3.59.
Cota inferior para la complejidad de la ley de Lynch	Teorema 3.61.
Ley de no convergencia para conjuntos con orden y permutaci3n	Secci3n 4.2.2.
Ley de no convergencia para conjuntos con orden y funci3n	Secci3n 4.2.3.
Ley de no convergencia para grafos ordenados muy sparse	Secci3n 4.2.4.

Quedan todav3a varios interrogantes sobre los temas tratados, como el de encontrar cotas superiores para la velocidad de convergencia y la complejidad computacional del teorema de Lynch sobre grafos aleatorios muy sparse (y ver si se corresponden con nuestras cotas inferiores) y las conjeturas [4.54](#) y [4.105](#) que enunciamos en el cap3tulo 4 sobre el comportamiento de conjuntos con dos ordenes y el espectro de la l3gica EMSO sobre grafos, respectivamente.

Sin embargo, la pregunta abierta mas interesante que nos dejan las investigaciones es de car3cter mas general: Vimos que algunas l3gicas, como la l3gica de primer orden de grafos, son **sorprendentemente inexpresivas**, mientras que otras, como la l3gica de primer orden sobre grafos ordenados o la l3gica EMSO sobre grafos son **sorprendentemente expresivas**, permitiendo codificar todo tipo de comportamiento indecidible. ¿Ser3 posible encontrar una l3gica de expresividad intermedia, capaz de expresar los resultados fascinantes de Erdos y Renyi sobre los treshhold de conexi3n y de aparici3n del componente gigante manteniendo alg3n an3logo de la ley de Fagin?. En otras palabras **¿Ser3 posible encontrar una l3gica que capture mejor como se practica la combinatoria y teor3a de probabilidad?**

Apéndice A

Algunas definiciones elementales

En este apéndice daremos las definiciones elementales de lógica que usaremos a lo largo del trabajo. Estas definiciones están cubiertas en cualquier libro de lógica elemental. En particular, nosotros seguiremos a [36]. Empezaremos definiendo la lógica de primer orden.

Definición A.1. Un lenguaje \mathcal{L} consiste en:

- Un conjunto \mathcal{L}_c que llamaremos el conjunto de **símbolos de constante**.
- Para todo $k \geq 1$, un conjunto \mathcal{L}_f^k que llamaremos el conjunto de **símbolos de función** de aridad k .
- Para todo $k > 1$, un conjunto \mathcal{L}_r^k que llamaremos el conjunto de **símbolos de relación** de aridad k .

Definición A.2. Un lenguaje \mathcal{L} se dice **relacional** si no contiene símbolos de función.

Definición A.3. Sea \mathcal{L} un lenguaje. Una \mathcal{L} -**estructura** consiste en un conjunto M (que llamaremos el dominio) equipado con lo siguiente:

- Un elemento $c^M \in M$ por cada símbolo de constante $c \in \mathcal{L}_c$
- Una función $f^M : M^k \rightarrow M$ por cada función $f \in \mathcal{L}_f$ de aridad k
- Un conjunto $R^M \subset M^K$ por cada relación $R \in \mathcal{L}_r$ de aridad k

Observación A.4. Como un elemento destacado de M no es más que una función $f : \{*\} \rightarrow M$ y para todo conjunto $M^0 = \{*\}$, algunos autores deciden ser más económicos y no definir los símbolos de constante, en su lugar tratándolos como funciones de aridad 0. Nosotros decidimos seguir el *approach* clásico para evitar confusiones.

Ejemplo A.5. Si \mathcal{L} es el lenguaje con una única función binaria, las \mathcal{L} -estructuras son conjuntos equipados con una relación binaria, es decir, **magmas**. En particular, esta clase engloba a todos los grupos.

Ejemplo A.6. Si \mathcal{L} es el lenguaje con un único símbolo de constante, las \mathcal{L} -estructuras son conjuntos con un elemento destacado, es decir, **conjuntos punteados**.

Definición A.7. Sea \mathcal{L} un lenguaje y M una \mathcal{L} -estructura. Un subconjunto $N \subset M$ es llamado una **subestructura** si para todo símbolo de constante $c \in \mathcal{L}$, vale que $c^M \in N$ y para todo símbolo de función $f \in M$ vale que N es cerrado por f^M .

Definición A.8. Sea \mathcal{L} un lenguaje y sean M y N dos \mathcal{L} -estructuras. Una función $\pi : M \rightarrow N$ es un **morfismo** de \mathcal{L} -estructuras si:

- Para toda constante $c \in \mathcal{L}$, vale que $\pi(c^M) = c^N$.
- Para toda función $f \in \mathcal{L}$ de aridad k y toda tupla $\bar{m} \in M^k$, vale que $\pi(f^M(\bar{m})) = f^N(\pi(\bar{m}))$.
- Para toda relación $R \in \mathcal{L}$ de aridad k y toda tupla $\bar{m} \in M^k$, vale que $M \models R^M(\bar{m})$ implica $N \models R^N(\pi(\bar{m}))$.

Observación A.9. La definición anterior generaliza las nociones de morfismo entre estructuras algebraicas que conocemos, como el morfismo de grupos y de anillos.

Definición A.10. Sea \mathcal{L} un lenguaje. Dos \mathcal{L} -estructuras M y N son **isomorfas** (denotado $M \cong N$) si existen morfismos de \mathcal{L} -estructuras $\pi : M \rightarrow N$ y $\theta : N \rightarrow M$ tal que $\pi \circ \theta = id_M$ y $\theta \circ \pi = id_N$.

Ahora tenemos que definir la parte sintáctica, las fórmulas de primer orden sobre \mathcal{L} . Las fórmulas sobre \mathcal{L} son una combinación formal de los símbolos de \mathcal{L} , los símbolos lógicos $=, \neg, \wedge, \exists$ y un conjunto infinito $\{x, y, z, \dots\}$ de variables. Empezamos definiendo los términos y luego definiremos las fórmulas.

Definición A.11. Sea \mathcal{L} un lenguaje. Los **términos** sobre \mathcal{L} se definen recursivamente:

- Si x es una variable, es un término.
- Si $c \in \mathcal{L}$ es un símbolo de constante, c es un término.
- Si $f \in \mathcal{L}$ es una función de aridad k y t_1, \dots, t_k son términos, $f(t_1, \dots, t_k)$ es un término.

Definición A.12. Sea \mathcal{L} un lenguaje. Las **fórmulas** sobre \mathcal{L} se definen recursivamente:

- Si t_1, t_2 son términos, $t_1 = t_2$ es una fórmula.
- Si $R \in \mathcal{L}$ es una relación de aridad k y t_1, \dots, t_k son términos, $R(t_1, \dots, t_k)$ es una fórmula.
- Si ϕ es una fórmula, $\neg\phi$ es una fórmula.
- Si ϕ_1 y ϕ_2 son fórmulas, $\phi_1 \wedge \phi_2$ es una fórmula.

- Si ϕ es una fórmula y x es una variable, $\exists x : \phi$ es una variable.

Definición A.13. Una fórmula formada de forma i) o ii) es llamada **atómica**.

Definición A.14. Una variable que aparece en una fórmula bajo el dominio de un cuantificador se dice que esta **ligada**, mientras que si no aparece bajo el dominio de ningún cuantificador se dice que esta **libre**.

Definición A.15. Una fórmula sin variables libres es llamada una **oración**.

Ejemplo A.16. Sea $\mathcal{L} = \{f\}$ el lenguaje con una única función unaria. $\phi_1 = [\forall x : f(f(x)) = x]$ es una oración, mientras que $\phi_2 = [\exists y : f(y) = x]$ no, ya que tiene a la variable x libre.

Ahora viene la definición que conecta el mundo la semántico con el sintáctico, de cuando una fórmula es cierta en una \mathcal{L} -estructura dada. En general, el valor de una fórmula depende del valor que le demos a las variables libres, concepto que se captura mediante la definición de **valuación**.

Definición A.17. Sea \mathcal{L} un lenguaje, M una \mathcal{L} -estructura. Una valuación sobre M es una función $v : VAR \rightarrow M$.

Definición A.18. Dada valuación v sobre M , variable x y elemento $m \in M$, definimos la valuación $v(x = m)$ como la valuación sobre M que coincide con v en toda variable salvo en x , donde vale m .

Definición A.19. Sea \mathcal{L} un lenguaje, M una \mathcal{L} -estructura y v una valuación sobre M . Definimos $\bar{v} : TERM \rightarrow M$ de forma recursiva

- $\bar{v}(x) = v(x)$
- $\bar{v}(c) = c^M$
- $\bar{v}(f(t_1, \dots, t_k)) = f^M(v(t_1), \dots, v(t_k))$

Definición A.20. Sea \mathcal{L} un lenguaje y v una valuación.

- $M[v] \models t_1 = t_2$ si y solo si $\bar{v}(t_1) = \bar{v}(t_2)$
- $M[v] \models R(t_1, \dots, t_k)$ si y solo si $(\bar{v}(t_1), \dots, \bar{v}(t_k)) \in R^M$
- $M[v] \models \neg\phi$ si y solo si $M[v] \not\models \phi$
- $M[v] \models \phi_1 \wedge \phi_2$ si y solo si $M[v] \models \phi_1$ y $M[v] \models \phi_2$
- $M[v] \models \exists x \phi$ si y solo si hay un $m \in M$ tal que $M[v(x = m)] \models \phi$

A partir de ahora denotaremos a las oraciones como ϕ y a las fórmulas con variables libres como $\phi(x_1, \dots, x_k) = \phi(\bar{x})$. Además denotaremos $M[(x_1 = m_1, \dots, x_k = m_k)] \models \phi(x_1, \dots, x_k)$ como $M \models \phi(m_1, \dots, m_k)$.

Observación A.21. El lector atento habrá notado que en nuestra definición de lógica de primer orden no incluimos ni el signo \vee ni el cuantificador \forall . La razón es puramente económica: como muchas demostraciones son por inducción estructural, es conveniente tener la definición más corta posible. En su lugar, definimos $\phi_1 \vee \phi_2$ como $\neg(\neg\phi_1 \wedge \neg\phi_2)$ y $\forall x \phi$ como $\neg(\exists x \neg\phi)$. Por las leyes de Morgan, se cumple la semántica que esperamos:

- $M[v] \models \phi_1 \vee \phi_2$ si y solo si $M[v] \models \phi_1$ o $M[v] \models \phi_2$
- $M[v] \models \forall x \phi$ si y solo si hay para todo $m \in M$ vale que $M[v(x = m)] \models \phi$

Observación A.22. Sea ϕ fórmula en lenguaje \mathcal{L} , es fácil ver que el valor de $M[v] \models \phi$ depende solo de los valores que toma v en las variables libres de ϕ , en particular, si ϕ es una oración el valor de $M[v] \models \phi$ es el mismo para toda valuación v , luego lo denotaremos simplemente como $M \models \phi$.

Definición A.23. Definimos las siguientes abreviaciones:

- $\exists!x : \phi(x)$ es abreviación de $\exists x : \phi(x) \wedge (\forall y \phi(y) \rightarrow y = x)$
- $\forall x_1 (\dots) \forall x_k \text{ distintos } \phi(x_1, \dots, x_k)$ es abreviación de $\forall x_1 (\dots) \forall x_k [(\bigwedge_{i \neq j} (x_i \neq x_j)) \rightarrow \phi(x_1, \dots, x_k)]$

La **lógica de segundo orden** (SO) extiende a la lógica de primer orden con la capacidad de cuantificar sobre relaciones sobre una estructura dada. Formalmente, extiende el lenguaje con, para todo $k > 0$, una colección de variables X_1^k, X_2^k, \dots que representaran relaciones de aridad k , y la capacidad de formar formulas de la forma $(t_1, \dots, t_k) \in X_i^k$ para todo conjunto de términos t_1, \dots, t_k . La definición de valuación y de satisfacibilidad se extiende de manera análoga.

Definición A.24. La **lógica de segundo orden monádica** (MSO) es el fragmento de la lógica de segundo orden en la que solo se cuantifica sobre conjuntos, es decir, sobre relaciones de aridad 1.

A diferencia de la la lógica de primer orden, la lógica de segundo orden depende de la elección de un universo de conjuntos. Sin embargo, esas sutilezas técnicas no son de importancia para esta tesis, en la que solo aparecerán estructuras finitas y contables.

Bibliografía

- [1] Sanjev and Arora. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2012.
- [2] A D Barbour, Lars Holst, and Svante Janson. *Poisson Approximation*. Oxford University Press, 1992.
- [3] Patrick Billingsley. *Probability and measure*. Wiley, 1995.
- [4] Béla Bollobás. *Random Graphs*. Cambridge University Press, 2011.
- [5] Heinz Dieter Ebbinghaus. *Finite Model Theory*. Springer, 1995.
- [6] Yuri Gurevich Egon Börger, Erich Grädel. *The Classical Decision Problem*. Springer, 1997.
- [7] Andrzej Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 1961.
- [8] Paul Erdos and Alfred Renyi. On random graphs i. *Publicationes Mathematicae (Debrecen)*, 6, 1959.
- [9] Paul Erdos and Alfred Renyi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5, 1960.
- [10] Paul Erdős and Alfréd Rényi. Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungaricae*, 14, 1963.
- [11] Paul Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53, 1947.
- [12] Ronald Fagin. Probabilities on finite models. *THE JOURNAL OF SYMBOLIC LOGIC*, 41, 1976.
- [13] Ronald Fagin. Finite model theory: a personal perspective. *Theoretical Computer Science*, 116, 1993.
- [14] Ronald Fagin. Easier ways to win logical games. *Descriptive Complexity and Finite Models*, 31, 1996.
- [15] Roland Fraisse. *Theory of Relations*. North-Holland, 1959.
- [16] Haim Gaifman. On local and non-local properties. In J. Stern, editor, *Proceedings of the Herbrand Symposium, Logic Colloquium '81*, 1982.
- [17] Eric Grädel. *Automata, Logics, and Infinite Games*. Springer, 2002.

- [18] William Hanf. Model-theoretic methods in the study of elementary logic. In J. W. Addison, Leon Henkin, and Alfred Tarski, editors, *The Theory of Models*. Addison–Wesley, 1965.
- [19] Frank Harary. *Graphical Enumeration*. Springer, 1973.
- [20] William Hodges. *Model theory*. Cambridge University Press, 1993.
- [21] Neil Immerman. *Descriptive Complexity*. Springer, 1999.
- [22] Svante Janson and Tomasz Łuczak. *Random Graphs*. Wiley, 2000.
- [23] Carol Karp. Finite quantifier equivalence. *Journal of Symbolic Logic*, 36(1):407–412, 1965.
- [24] Matt Kaufmann. A counterexample to the 0-1 law for existential monadic second-order logic. In *CLI Internal Note 32*. Computational Logic Inc., 1987.
- [25] Matt Kaufmann and Saharon Shelah. On random models of finite power and monadic logic. *Discret. Math.*, 54, 1985.
- [26] Phokion G. Kolaitis and Moshe Y. Vardi. Infinitary logics and 0–1 laws. *Information and Computation*, 98(2):258–294, 1992.
- [27] Jean Marie Le Bars. The 0–1 law fails for monadic existential second-order logic on undirected graphs. *Information processing letters*, 77, 2001.
- [28] James Lynch. Probabilities of first order sentences about unary functions. *Transactions of the American Mathematical Society*, 287, 1985.
- [29] James Lynch. Probabilities of sentences about very sparse random graphs. *Random Structures and Algorithms*, 3, 1992.
- [30] Walter Oberschelp. Asymptotic 0-1 laws in combinatorics. In Dieter Jungnickel and Klaus Vedder, editors, *Combinatorial Theory*. Springer Berlin Heidelberg, 1982.
- [31] Elena Pezzoli. Computational complexity of ehrenfeucht-fraïssé games on finite structures. In Georg Gottlob, Etienne Grandjean, and Katrin Seyr, editors, *Computer Science Logic*, pages 159–170. Springer, 1999.
- [32] Saharon Shelah. The monadic theory of order. *Annals of Mathematics*, 102(3):379–419, 1975.
- [33] Saharon Shelah. Nonconvergence, undecidability, and intractability in asymptotic problems. *THE JOURNAL OF PURE AND APPLIED LOGIC*, 36, 1987.
- [34] Saharon Shelah. Very weak zero one law for random graphs with order and random binary functions. *Random Structures & Algorithms*, 9(4):351–358, 1996.
- [35] Saharon Shelah and Joel Spencer. Zero-one laws for sparse random graphs. *Journal of the American Mathematical Society*, (1):97–115, 1988.
- [36] Raymond Smullyan. *A Beginner’s Guide to Mathematical Logic*. Courier Corporation, 2014.
- [37] Joel Spencer. *The Strange Logic of Random Graphs*. Springer, 2001.
- [38] Speyer. Answer to “The dance marathon problem”. MathOverflow, 2017. Accessed: 2026-01-12.