Quadrati Sudoku e Polinomi Cromatici*

Agnes M. Herzberg e M. Ram Murty†

giugno 2007

Il Sudoku è diventato un puzzle molto popolare che molti giornali propongono come appuntamento giornaliero. Il gioco consiste in una griglia di 9 x 9 celle, in ognuna delle quali si deve collocare un numero da 1 a 9, in modo tale che in ognuna delle 9 righe, colonne o sottogriglie 3×3, compaiano tutte e sole le cifre da 1 a 9.

Figura 1. La griglia del Sudoku.

Ricordiamo che un quadrato Latino di rango n è un ordinamento $n \times n$ di numeri, tale che ogni riga e colonna contenga tutti i numeri da *1* a n. In particolare, si può affermare che ogni Sudoku è un quadrato Latino di rango 9, ma non viceversa, a causa della condizione relativa alle sottogriglie 3×3.

							1	
4								
	2							
				5		**4**		**7**
		8				**3**		
		1		**9**				
3			**4**			**2**		
	5		**1**					
			8		**6**			

Figura 2. Un Sudoku con 17 dati specifici.

Coloro che tentano di risolvere un Sudoku si pongono naturalmente diversi interrogativi, quali ad esempio:

1. per un dato Sudoku esiste una soluzione?
2. se la soluzione esiste è unica?
3. se la soluzione non è unica quante soluzioni esistono?

*Articolo pubblicato dal Notiziario della Società Americana di Matematica (http://www.ams.org/notices/200706/tx070600708p.pdf) nel numero di giugno 2007 e tradotto da arch. FEDERICO ALVINO (05/2010), alvino.federico@yahoo.it, la cui pubblicazione solo su questo sito è stata autorizzata dal dr. Erin M. Buck, per conto della AMS, unica titolare dei diritti di copyright, in data 07/06/2010.

†Agnes M. Herzberg (herzberg@post.queensu.ca) e M. Ram Murty (murty@mast.queensu.ca) sono professori emeriti di matematica presso la Queen's University, Canada. La loro ricerca è stata parzialmente finanziata dal NSERC, Natural Science and Engineering Research Council.

4. esiste una procedura sistematica per determinarne le soluzioni?

5. quanti sono i Sudoku con soluzione unica?

6. qual'è il numero minimo di dati specifici che garantisce l'unicità della soluzione?

Per esempio, in figura 2 si mostra che il numero minimo di dati specifici è al massimo 17. (Lasciamo al lettore la verifica dell'unicità della soluzione del Sudoku di fig. 2).

Al momento non si sa se un Sudoku con 16 dati specifici ha una soluzione unica. Gordon Royle ha collezionato 36628 distinti Sudoku con 17 dati specifici.[1]

Noi riformuleremo molti di questi interrogativi in un contesto matematico per tentare di trovare le risposte. Più precisamente, noi interpretiamo il Sudoku come un problema di colorazione dei vertici all'interno della teoria dei grafi. Ciò rende possibile generalizzare le questioni e inquadrarle in una struttura più ampia. Noi discuteremo, inoltre, la relazione tra quadrati Latini e Sudoku per dimostrare che l'insieme dei quadrati Sudoku è sostanzialmente più piccolo di quello dei quadrati Latini.

1 Polinomi cromatici.

Per convenienza del lettore, ricordiamo la nozione di $grafo$ *colorato completo*. Un $\lambda -$ *colorato* di un grafo G è una mappa f dall'insieme dei vertici di G a $\{1, 2, ...\lambda\}$. Pertanto una mappa si definisce *colorata completa* se si verifica che *f(x)* $\neq$ *f(y)* con x e y giacenti in G. Il numero minimo di colori richiesti per colorare *completamente* i nodi di un grafo G è chiamato *numero cromatico* di G e viene denotato con $\chi(G)$. Pertanto non è difficile vedere che, in realtà, il Sudoku diventa un problema di colorazione dei grafi. Infatti, noi associamo un grafo alla griglia *9 x 9* del Sudoku nel modo seguente. Il grafo avrà 81 vertici, ognuno corrispondente a una cella della griglia. Due distinti vertici saranno considerati *adiacenti se e solo se le celle corrispondenti della griglia si trovano o sulla stessa riga, o sulla stessa colonna o nella stessa sottogriglia.* Quindi, un Sudoku esatto corrisponde a un *colorato completo* di questo grafo. Poniamo ciò in un contesto un po più generale. Consideriamo una griglia n^2 *x* n^2. Associamo a ogni cella della griglia un vertice etichettato da (i, j) con $1 \leq i, j \leq n^2$. Noi diremo che (i, j) e (i', j') sono adiacenti se $i = i'$ o $j = j'$oppure se $\lceil i/n \rceil = \lceil i'/n \rceil$ e $\lceil j/n \rceil = \lceil j'/n \rceil$[2]. Indicheremo questo grafo con X_n e lo chiameremo *grafo del Sudoku di rango n.* Un quadrato Sudoku di rango n sarà un *colorato completo* di questo grafo usando n^2 colori. Un puzzle Sudoku corrisponde a un *colorato parziale* e il problema è di vedere se questo *colorato parziale* può essere completato in un *colorato completo* del grafo. Sottolineiamo che, in genere, è più conveniente etichettare i vertici di un *grafo Sudoku di rango n* usando *(i,j)* con $0 \leq i, j \leq n^2 - 1$. Quindi *(i,j)* e *(i',j')* sono adiacenti se $i = i'$ o $j = j'$ oppure se $\lfloor i/n \rfloor = \lfloor i'/n \rfloor$ e $\lfloor j/n \rfloor = \lfloor j'/n \rfloor$[3]. Il che significa che $\lfloor \chi \rfloor$ è l'intero più grande minore o uguale a χ. Ricordiamo che un grafo è definito *regolare* se ogni vertice ha lo stesso grado. Semplici calcoli mostrano che X_n è un grafo regolare quando ogni vertice ha grado pari a $3n^2 - 2n - 1 = (3n + 1)(n - 1)$. Nel caso di $n = 3$, X_3 è un grafo $20 -$ *regolare* e nel caso $n = 2$, X_2 è un grafo $7 -$ *regolare*. Il numero di modi per colorare un grafo G con λ colori è ben noto ed espresso da un polinomio in λ di grado eguale al numero dei vertici di G. Il nostro primo teorema dimostra che dato un *parziale colorato* C di G, il numero di modi per completare la colorazione per ottenere un *colorato completo* usando λ colori è ancora un polinomio in λ, a condizione che λ sia maggiore o uguale al numero dei colori usati in C. Più precisamente tanto è indicato nel Teorema n° 1.

1.1 Teorema n° 1.

Sia G un grafo finito di v vertici. Sia C un parziale colorato di G di t vertici formato con d_0 colori. Sia $p_{G,C}(\lambda)$ il numero di modi per completare la colorazione usando λ colori per ottenere un colorato completo di G. Allora $p_{G,C}(\lambda)$ è un polinomio a una variabile (λ) con coefficienti interi di grado *v-t* per $\lambda \geq d_0$. Daremo due dimostrazioni di questo teorema.

La prova più diretta utilizza la teoria degli insiemi parzialmente ordinati e le funzioni di Möbius che ricapitoliamo brevemente. Un insieme parzialmente ordinato (acronimo *poset*), è un insieme P di elementi uniti con un ordinamento parziale indicato dalla relazione $\leq$ che soddisfa le seguenti condizioni: (a) $x \leq x$ per ogni $x \in P$; (b) $x \leq y$ e $y \leq x$ implica che $x = y$; (c) $x \leq y$ e $y \leq z$ implica $x \leq z$. Noi considereremo solo *posets* finiti. Esempi comuni di *posets* comprendono la collezione dei sottogruppi di un gruppo finito parzialmente ordinato dall'insieme inclusione e la collezione dei divisori positivi di un numero naturale n parzialmente ordinato dal criterio di divisibilità. Un piccolo esempio è dato dalla seguente costruzione.

[1] vedi http://www.csse.uwa.edu.au/~gordon/sudokumin.php

[2] Qui la notazione $\lceil \bullet \rceil$ significa che arrotondiamo all'intero maggiore più vicino.

[3] Con $\lfloor \bullet \rfloor$ che indica la funzione intera più grande.

Sia G un grafo finito ed e un bordo di G. Il grafo ottenuto da G identificando i due vertici uniti da e (con la rimozione dei conseguenti lati multipli) è denotato da G/e ed è chiamato *contrazione di G da e.* In generale, noi diciamo che il grafo *G'* è una contrazione di G se *G'* è ottenuto da G con una serie di contrazioni. L'insieme di tutte le contrazioni di un grafo finito G può ora essere parzialmente ordinato definendo che $A \leq B$ se A è una contrazione di B. Dato un *poset* finito parzialmente ordinato da $\leq$, noi definiamo la funzione di Möbius $\mu : P \times P \to \mathbb{Z}$ ricorsivamente, impostando: $\mu(x,x) = 1$; $\sum_{x \leq y \leq z} \mu(x,y) = 0$; se $x \neq z$.

Il teorema principale della teoria delle funzioni di Möbius è il seguente.

Se $f : P \to \mathbb{C}$ è una funzione a valori complessi e definiamo

$$g(y) = \sum_{x \leq y} f(x)$$

allora

$$f(y) = \sum_{x \leq y} \mu(x,y) g(x)$$

e viceversa.

Rimandiamo il lettore a (vedi nota)[4] per ampliare queste idee.

1.2 Dimostrazione del Teorema n° 1.

Useremo la richiamata teoria delle funzioni di Möbius per provare il Teorema n° 1. Sia dato (G,C), con G grafo e C colorato completo di alcuni vertici. Diremo che (G',C) è un sottografo di (G,C) se *G'* è ottenuto contraendo qualche bordo di G con al più un punto finale in C. In questo modo otteniamo un insieme parzialmente ordinato. La contrazione minima sarebbero i vertici di C esclusi gli adiacenti. Useremo la lettera C per indicare il grafo minimo nel nostro *poset*. Per ogni sottografo(G',C) di questo *poset*, sia $p_{G',C}(\lambda)$ il numero minimo di modi per colorare completamente *G'* con λ colori tra quelli specificati per C. Sia $q_{G',C}(\lambda)$ il numero di modi per colorare (non necessariamente in modo completo) i vertici di *G*' usando λ colori tra quelli specificati per C. Chiaramente $q_{G',C}(\lambda) = \lambda^{v'-t}$, dove v' è il numero di vertici di *G'* e t è il numero dei vertici di C. Se $\lambda \geq d_0$, allora, dato un $\lambda -$ *colorato* di (G,C), possiamo derivare un colorato completo di un unico sottografo(G',C) semplicemente contraendo i bordi, ogni volta che due vertici adiacenti di (G,C) hanno lo stesso colore assegnato. In questo modo otteniamo la relazione

$$q_{G,C}(\lambda) = \lambda^{v-t} = \sum_{C \leq G'} p_{G'C}(\lambda)$$

Dalla relazione inversa del Teorema di Möbius, deduciamo che

$$p_{G,C}(\lambda) = \sum_{C \leq G'} \mu(C,G') \lambda^{v'-t}$$

dove il secondo membro è proprio un polinomio a una variabile a coefficienti interi di grado *v-t*, come enunciato dal teorema.

Possiamo anche dimostrare il Teorema n° 1 senza l'uso delle funzioni di Möbius. Applichiamo il principio di induzione sui numeri dei bordi del grafo(G,C). Consideriamo tre casi:

(1) Supponiamo che e sia un bordo che unisce due vertici di G, con al più uno contenuto in C. Useremo la seguente notazione $G - e$ che denoterà il grafo ottenuto da G eliminando il bordo e, ma non i suoi punti finali. Il grafo ottenuto da G, identificante i due vertici uniti da e (e rimuovendo ogni bordo multiplo) sarà denotato con G/e. Con questa notazione, abbiamo:

$$p_{G,C}(\lambda) = p_{G-e,C}(\lambda) - p_{G/e,C}(\lambda)$$

perché ogni colorato completo di G è anche un colorato completo di $G - e$ e un colorato completo di $G - e$ genera un colorato completo di G se e solo se assegna colori distinti ai punti finali x, y di e. Perciò il numero di colorati completi $p_{G,C}(\lambda)$ può essere ottenuto da $p_{G-e,C}(\lambda)$ sottraendo quelle colorazioni che assegnano lo stesso colore sia a x che a y, e questo numero corrisponde a $p_{G/e,C}(\lambda)$. Ogni *G-e* e G/e hanno meno bordi di G. Perciò possiamo applicare il principio di induzione e completare, in questo caso, la dimostrazione.

(2) Ora supponiamo che G ha un solo vertice v_0 non contenuto in C. Se questo vertice non è adiacente ad ogni vertice di C, allora $G = C \cup v_0$, che è l'unione disgiunta di C e del vertice v_0, per cui possiamo colorare v_0 usando ognuno dei λ colori. Quindi, in questo caso, $p_{G,C}(\lambda) = \lambda$. Se il vertice è adiacente a d vertici di C, e questi vertici utilizzano d_0 colori, allora $p_{G,C}(\lambda) = max(\lambda - d_0, 0)$.

[4]J.H. Van Lint e R.M.Wilson, A Course in Combinatorics, Cambridge University Press, 1992.

(3) Ogni vertice di G è contenuto in C. In questo caso, abbiamo già un colorato di G e $p_{G,C}(\lambda) = 1$. Perciò, con il principio di induzione sul numero dei bordi del grafo, il teorema è dimostrato.

In una sezione successiva esamineremo le implicazioni di questo teorema per il puzzle Sudoku. Adesso basta evidenziare che, dato un puzzle Sudoku, il numero di modi per completare il grafo è dato da $p_{x_3,c}(9) = 1$. Sarebbe estremamente interessante determinare sotto quali condizioni un colorato parziale può essere esteso a un colorato completo. In questa direzione abbiamo il seguente enunciato generale.

1.3 Teorema n° 2.

Sia G un grafo con numero cromatico $\chi(G)$ e C un parziale colorato di G che utilizza solo $\chi(G) - 2$ colori. Se il parziale colorato può essere completato in colorato completo di G, allora ci sono al più due modi per estendere la colorazione.

1.4 Dimostrazione del Teorema n° 2.

Dal momento che due colori non sono stati usati nel parziale colorato iniziale, questi due colori possono essere scambiati nel colorato completo finale per ottenere un'altra estensione completa.

Il Teorema n° 2 implica che se C è un parziale colorato di G che può essere completato *univocamente* in un colorato completo di G, allora C deve utilizzare almeno $\chi(G) - 1$ colori. In particolare, in ogni Sudoku puzzle, almeno 8 colori devono essere usati nelle celle date. In generale, per puzzles Sudoku n^2 *x* n^2, almeno $n^2 - 1$ colori devono essere utilizzati, nel parziale colorato dato, affinché il puzzle abbia una soluzione unica.

2 Programmazione e parziali colorati

Dato un grafo G, con uno specifico *parziale colorato*, possiamo chiederci se questo può essere esteso a un *colorato completo* del grafo. E' ben noto che i problemi di colorazione dei grafi codificano problemi di programmazione nella vita reale. L'estensione da un *parziale colorato* a un *colorato completo* corrisponde a un problema di programmazione con vincoli addizionali, come, per esempio, quando vogliamo programmare incontri di varie commissioni in certi intervalli di tempo, con alcune di queste già impegnate in altri determinati intervalli. Naturalmente, la corrispondente relazione di adiacenza è che due commissioni sono unite da un bordo se hanno un membro in comune. Questo è un problema di rilevante interesse e sembra difficile determinare i criteri per dire quando un *parziale colorato* può essere esteso a un *colorato completo* dell'intero grafo. Una situazione simile nasce per l'assegnazione delle frequenze. Supponiamo che, in una data regione, vi siano trasmittenti televisive che hanno bisogno di una frequenza per poter trasmettere. Due trasmittenti, entro 100 miglia l'una dall'altra, avranno assegnate frequenze diverse per evitare problemi di interferenze. Spesso alcune trasmittenti hanno già frequenze assegnate, e alle nuove trasmittenti devono essere assegnate le frequenze tenendo presenti questi vincoli. E', di nuovo, un problema di completamento da un *parziale colorato* a un *colorato completo* di un grafo. Infatti, noi associamo un vertice a ogni trasmittente e uniamo due di loro se si trovano entro 100 miglia l'una dall'altra. L'assegnazione di una frequenza corrisponde a un "*colore*" assegnato a quel vertice. Possiamo evidenziare molteplici esempi relativi a differenti contesti della "vita reale". In ogni caso il problema del passaggio da un *parziale colorato* a un *colorato completo* emerge come archetipo. Quadrati Latini e Sudoku sono solo casi specifici di questo tema. Tuttavia, essi possono anche essere studiati indipendente da questo contesto. L'esplicita costruzione di quadrati Latini è ben nota per le applicazioni nella progettazione di esperimenti statistici. Negli studi agricoli, per esempio, per poter piantare ν varietà di piante in ν file e ν colonne, in modo tale che le peculiarità del suolo in cui sono state piantate non abbiano influenza sull'esperimento. Gli agricoltori hanno già usato quadrati Latini ν *x* ν per progettare un tale esperimento. Questo serve per bilanciare i trattamenti dell'esperimento prima della randomizzazione. In questo contesto, se interessa anche testare il ruolo dei vari fertilizzanti sulla crescita di queste piante, si potrebbe utilizzare un quadrato Sudoku, dove ad ogni sottogriglia (o banda) viene applicato un diverso fertilizzante in modo da averne uno per ogni trattamento.

3 Colorazione esplicita per X_n

In questa sezione, mostreremo come si può colorare completamente il grafo Sudoku X_n. Ricordiamo che il *numero cromatico* di un grafo è il numero minimo di colori necessari per colorare completamente i suoi vertici. Così, il grafo completo K_n, composto da n vertici, in cui ogni vertice è adiacente a ogni altro vertice, ha numero cromatico n.

3.1 Teorema n° 3.

Per ogni numero naturale n, c'è un colorato completo del grafo Sudoku X_n usando n^2 colori. Il numero cromatico di X_n è n^2.

3.2 Dimostrazione del Teorema n° 3.

Notiamo subito che tutte le celle nell'angolo superiore sinistro della griglia $n\ x\ n$ sono adiacenti a tutte le altre, e questo forma un grafo isomorfo a K_{n^2}. Il numero cromatico di K_{n^2} è n^2 e così, X_n necessita di almeno n^2 colori per essere un colorato completo. Come rimarcato in precedenza, è conveniente etichettare i vertici (i, j) con $0 \leq i, j \leq n^2 - 1$. Consideriamo le classi di modulo residuo n^2. Per $0 \leq i, j \leq n^2 - 1$, scriviamo $i = t_i n + d_i$ con $0 \leq d_i \leq n-1$ e $0 \leq t_i \leq n-1$ e similarmente per $0 \leq j \leq n^2-1$. Assegniamo il "*colore*" $c(i, j) = d_i n + t_i + n t_j + d_j$, modulo ridotto n^2, alla (i, j)esima posizione nella griglia $n^2\ x\ n^2$. Affermiamo che in questo caso si ha un colorato completo. Per vedere ciò, mostreremo che ogni due coordinate adiacenti (i, j) e (i', j') hanno colori distinti. Infatti, se $i = i'$, allora dobbiamo mostrare che $c(i, j) \neq c(i, j')$ a meno che $j = j'$. Se $c(i, j) = c(i, j')$, allora $n t_j + d_j = n t_{j'} + d_{j'}$ che significa $j = j'$. Allo stesso modo, se $j = j'$, allora $c(i, j) \neq c(i', j)$ a meno che $i = i'$. Se ora, $[i/n] = [i'/n]$ e $[j/n] = [j'/n]$, allora $d_i = d_{i'}$ e $d_j = d_{j'}$. Se $c(i, j) = c(i', j')$ allora

$$t_i + n t_j = t_{j'} + n t_{j'}.$$

Riducendo questo modulo n abbiamo $t_i = t_{i'}$. Quindi $t_j = t_{j'}$, per cui $(i, j) = (i', j')$ anche in questo caso. Pertanto, questo è un colorato completo.

4 Conteggio delle soluzioni.

Vogliamo affrontare brevemente la questione dell'unicità della soluzione per un dato puzzle Sudoku. Non è sempre chiaro, in via preliminare, se un dato puzzle ha una soluzione. In questa sezione, determineremo qualche condizione necessaria perché vi sia una soluzione unica.

- La figura 3 mostra un esempio di puzzle Sudoku che offre precisamente due soluzioni.

Figura 3. Sudoku con due soluzioni.

9		6		7		4		3
			4			2		
	7			2	3		1	
5						1		
	4		2		8		6	
		3						5
	3		7				5	
		7			5			
4		5		1		7		8

- Al lettore il compito di mostrare come si passa dal puzzle della figura 3 alla configurazione della figura 4.

Figura 4. Soluzione del Sudoku di fig. 3.

9	2	6	5	7	1	4	8	3
3	5	1	4	8	6	2	7	9
8	7	4	9	2	3	5	1	6
5	8	2	3	6	7	1	9	4
1	4	9	2	5	8	3	6	7
7	6	3	1			8	2	5
2	3	8	7			6	5	1
6	1	7	8	3	5	9	4	2
4	9	5	6	1	2	7	3	8

- Chiaramente si può inserire l'una o l'altra delle soluzioni della figura 5 per completare la griglia. Così, abbiamo due soluzioni.

Figura 5. I due modi di completare il Sudoku.

9	4
4	9

4	9
9	4

- Questa osservazione porta alla seguente considerazione. Se nella soluzione di un puzzle Sudoku, abbiamo una configurazione del tipo indicato nella figura 6 nella stessa pila verticale, allora almeno una di queste voci deve essere inclusa come “dato” nel puzzle iniziale, altrimenti avremo due possibili soluzioni semplicemente scambiando a e b nella configurazione.

Figura 6. Una configurazione che porta a due soluzioni.

a	b
b	a

- Come rimarcato in precedenza, se il numero di “colori” utilizzati in un dato puzzle Sudoku è al più 7, ci sono almeno due soluzioni. Tanto perché potremo scambiare i due colori residui e trovare ancora una soluzione valida. La molteplicità delle soluzioni può anche essere spiegata con i polinomi cromatici. Se d_0 è il numero dei colori usati, abbiamo visto che $p_{x_3,c}(\lambda)$ è un polinomio in λ purché $\lambda \geq d_0$. Poiché il numero cromatico di X_3 è 9, dobbiamo avere $p_{x_3,c}(\lambda) = 0$ per $\lambda = d_0, d_0+1, ..., 8$. Siccome $p_{x_3,c}(\lambda)$ è un polinomio ad una variabile con coefficienti interi, possiamo scrivere

$$p_{x_3,c}(\lambda) = [(\lambda - d_0)(\lambda - d_0 + 1)....(\lambda - 8)]\, q(\lambda)$$

per ogni polinomio $q(\lambda)$ a coefficienti interi. Ponendo $\lambda = 9$ si ha

$$p_{x_3,c}(9) = (9 - d_0)!q(9)$$

con il secondo membro maggiore o uguale a 2 se $d_0 \leq 7$. Questo prova la condizione necessaria per l'unicità della soluzione, ammesso che ci sia una soluzione, che costituisce una tacita ammissione in ogni dato Sudoku. Nell'ultima sezione, mostriamo un Sudoku che utilizza solamente 8 colori e 17 dati specifici.

5 Conteggio dei Sudoku di rango 2.

In un recente articolo[5] Felgenhauer e Jarvis, con un metodo a forza bruta, hanno calcolato il numero dei quadrati Sudoku. Ci sono

6.670.903.752.021.072.936.930

quadrati Sudoku validi. Questo numero, approssimativamente pari a 6.671×10^{21}, è una piccola parte del numero totale dei quadrati Latini 9 x 9 pari a[6]:

$$5.524.751.496.156.892.842.531.225.600$$

$$\approxeq 5,525 \times 10^{27}.$$

Questo numero enorme di quadrati Sudoku può essere ridotto se facciamo alcune piccole osservazioni. Primo, iniziando con un quadrato Sudoku, possiamo creare $9! = 362.880$ nuovi Sudoku semplicemente rietichettando. Più precisamente, se $a_{i,j}$ rappresenta il (i, j)esimo dato di un Sudoku, e σ è una permutazione dell'insieme $\{1, 2, ...9\}$, allora il quadrato il cui (i, j)esimo dato è $\sigma(a_{ij})$ è ancora un Sudoku valido. Ci sono altre simmetrie da considerare. Per esempio, la trasposta di un Sudoku è ancora un Sudoku. Possiamo permutare ognuna delle tre fasce orizzontali, o delle tre pile verticali, oppure le righe all'interno delle fasce, oppure le colonne all'interno delle pile. Considerate tutte queste simmetrie, il numero dei Sudoku fondamentali diventa molto più maneggevole

5.472.730.538

approssimativamente $5,47 \times 10^9$, come è stato dimostrato[7].

[5] B. Fegelnhauer e F. Jarvis, Mathematics of Sudoku I, *Mathematical Spectrum* 39 (2006), 15-22.
[6] S. Bammel e J. Rothestein, The number of 9 x 9 Latin squares, *Discrete Math.* 11 (1975), 93-95.
[7] E. Russell e F. Jarvis, Mathematics of Sudoku II, *Mathematical Spectrum* 39 (2006), 54-58.

1	2		
3	4		

Figura 7. Una griglia Sudoku 4 x 4.

Questi calcoli possono essere ben compresi se consideriamo il caso dei Sudoku *4 x 4*.

Senza perdita di generalità, possiamo etichettare i dati nel primo blocco *2 x 2* come mostrato in figura 7. Non è difficile vedere che vi sono al più 2^4 modi per completare questa griglia. Tuttavia, possiamo essere più precisi. Si può mostrare che considerando le ovvie simmetrie già indicate ci sono solo *4 x 4* Sudoku! Infatti, dal momento che permutando le ultime due colonne abbiamo una simmetria ammissibile, possiamo supporre senza perdita di generalità che gli altri due dati nella prima riga sono (3,4) in quest'ordine. Allo stesso modo, dal momento che possiamo scambiare le ultime due righe, possiamo supporre che il nostro quadrato Sudoku si presenta come mostrato in figura 8.

Allora, 1 e 4 sono gli unici dati che possono essere inseriti nella posizione (3,3) ed è facilmente visibile che la scelta di 1 porta a una contraddizione.

1	2	3	4
3	4		
2			
4			

Figura 8. Un puzzle Sudoku.

Questo porta alle seguenti configurazioni.

1	2	3	4
3	4		
2		4	
4			1

1	2	3	4
3	4		
2		4	
4			2

1	2	3	4
3	4		
2		4	
4			3

Figura 9. Tre puzzle Sudoku 4 x 4.

I quadrati possono essere completati facilmente come mostrato in figura 10.

1	2	3	4
3	4	1	2
2	1	4	3
4	3	2	1

1	2	3	4
3	4	2	1
2	1	4	3
4	3	1	2

1	2	3	4
3	4	1	2
2	3	4	1
4	1	2	3

Fig. 10. I tre Sudoku determinati dalla figura 9.

Tuttavia, l'ultima configurazione è la trasposta della seconda con l'inversione di 2 e 3. Perciò, ci sono solo due soluzioni non equivalenti per il quadrato Sudoku 4 x 4. Da qui vediamo che, senza considerare le simmetrie, abbiamo un totale di $4! \times 2 \times 2 \times 3 = 288$ quadrati Sudoku di rango 2.

In questo contesto è interessante determinare il numero minimo di celle piene in un puzzle Sudoku 4 x 4, che porta ad un'unica soluzione. La figura 11 ne mostra uno con quattro dati. Ce n'è uno con tre? Diamo una prova che non esiste.

1			
			2
		4	
	3		

Figura 11. Un puzzle Sudoku 4 x 4 con quattro celle piene.

Come abbiamo dimostrato precedentemente, il numero dei "*colori*" da usare in ogni *parziale colorato* del grafo Sudoku di rango n deve essere almeno $n^2 - 1$, per avere una soluzione unica. Perciò nel puzzle di figura 11, noi dobbiamo utilizzare almeno 3 colori. Per provare che, per i grafi Sudoku di rango 2, il numero minimo è quattro, dobbiamo dimostrare che con tre colori compilati non si ha una soluzione unica. Il che si effettua facilmente osservando le due griglie inequivalenti della figura 10 e verificando i casi che si presentano ad uno ad uno.

6 Permanenti e Sistemi di Rappresentativi Distinti.

In questa sezione riprenderemo diversi teoremi su *permanenti* e *sistemi di rappresentativi distinti* che useremo nella prossima sezione per la nostra enumerazione dei quadrati Sudoku. Per maggiori dettagli rimandiamo il lettore alla nota (4).

Se A è una matrice n *x* n, con la (i,j)esima posizione data da $a_{i,j}$, il *permanente* di A, denotato con *per* A, è definito da

$$\sum_{\sigma \in \xi_n} a_{1\sigma(1)} a_{2\sigma(2)} \ldots\ldots\ldots\ldots a_{n\sigma(n)},$$

dove ξ_n (*vedi nota 9*) denota il gruppo simmetrico sugli n simboli $\{1, 2, \ldots\ldots\ldots\ldots n\}$.

La matrice A è detta *doppiamente stocastica* se entrambe le somme di righe e colonne sono uguali a 1.

Nel 1926, B. L. van der Waerden pose il problema della determinazione del *permanente minimo* tra tutte le matrici n *x* n doppiamente stocastiche.

Si è ritenuto che il minimo è raggiunto dalla matrice costante i cui elementi fossero uguali a $1/n$. Nel corso degli anni questa congettura è cambiata nell'altra:

$$per\ A \geq \frac{n!}{n^n}, \tag{1}$$

per ogni matrice doppiamente stocastica, ed è stata poi denominata come congettura van der Waerden. Dal 1981 sono apparse due prove differenti della congettura van der Waerden, una di D. I. Falikman e l'altra di G. P. Egoritsjev. Come per i limiti superiori, nel 1967, H. Minc congetturò che se A è una matrice *(0,1)* con somme delle righe r_i, allora

$$per\ A \leq \prod_{i=1}^{n} r_i!^{1/n}. \tag{2}$$

Questa congettura è stata provata nel 1973 da L. M. Brégman (*vedi nota 4* - pag 82). Noi vogliamo utilizzare sia il limite inferiore che superiore per i *permanenti*, nella nostra enumerazione dei quadrati Sudoku. L'applicazione avverrà tramite il teorema di Phillip Hall (noto come *teorema del matrimonio*) che descriviamo. Supponiamo di avere sottoinsiemi A_1, A_2,.......A_n dell'insieme $\{1, 2, \ldots\ldots n\}$ e di voler selezionare distinti elementi $a_i \in A_i$. Una tale selezione è chiamata *sistema di rappresentativi distinti.* Il teorema fornisce le condizioni necessarie e sufficienti per dire quando ciò può essere fatto (*vedi nota 4*). Se per ogni sottoinsieme S di$\{1, 2, \ldots\ldots n\}$, poniamo

$$N(S) = \cup_{j \in S} A_j,$$

allora una semplice riflessione mostra che deve essere $\mid N(S) \mid \geq \mid S \mid$ affinché vi sia una selezione. Qui $\mid S \mid$ indica la cardinalità dell'insieme S. Il teorema di Hall stabilisce che questa condizione è anche sufficiente. Il numero di modi per fare ciò è dato dal *permanente* della matrice A, n *x* n, definita come segue. E' una matrice *(0,1)* il cui *(i,j)esimo* elemento è 1 se e solo se $i \in A_j$. Si farà riferimento ad A come alla *matrice di Hall* associata con gli insiemi A_1, A_2,......A_n. Questi risultati possono essere usati per ottenere i limiti superiori e inferiori del numero dei quadrati Latini di ordine n (*vedi nota 4*). Dal momento che nella prossima sezione avremo bisogno del limite inferiore, diamo i dettagli di questo calcolo.

Per un quadrato Latino *n x n*, il numero di modi per riempire la prima riga è chiaramente *n!*. Supponiamo di aver completato K righe del quadrato Latino e che vogliamo riempire la *(k+1)esima* riga. Per ogni cella i della *(k+1)esima* riga, sia A_i l'insieme dei numeri non utilizzati nella *i esima* colonna. La dimensione di A_i è, perciò, *n-k*. Riempire la *(k+1)esima* riga del nostro quadrato Latino equivale a trovare un insieme di *rappresentativi distinti* degli insiemi A_1, A_2,.......A_n. Il numero di modi in cui ciò si può fare è il *permanente* della corrispondente *matrice di Hall A*. Dal momento che la matrice $(n-k)^{-1}A$ è doppiamente stocastica, l'equazione (1) mostra che ci sono almeno

$$\frac{(n-k)^n n!}{n^n}$$

modi di farlo. Prendendo il prodotto per k che va da *0* a *n-1*, abbiamo:

- Lemma 4. *Il numero di quadrati Latini di ordine n è almeno*: $\frac{n!^{2n}}{n^{n^2}}$.
- Corollario 5. *Il numero dei quadrati Latini di ordine* n^2 *è almeno*: $n^{2n^4}e^{-2n^4+O(n^2 logn)}$.
- Dimostrazione. Dal Lemma 4, il numero dei quadrati Latini di ordine n^2 è almeno

$$\frac{n^2!^{2n^2}}{n^{2n^4}}.$$

Utilizzando la formula di Stirling,[8]

$$\log n! = n\log n - n + \frac{1}{2}\log n + O,$$

otteniamo il limite inferiore ricercato.

7 Quadrati Latini e Quadrati Sudoku.

Adesso possiamo provare che:

7.1 Teorema 6.

Il numero dei quadrati Sudoku di rango n, per n sufficientemente grande, è delimitato da

$$n^{2n^4}e^{-2{,}5n^4+O(n^3\log n)}.$$

7.2 Dimostrazione del Teorema 6.

Il quadrato Sudoku $n^2 \times n^2$ è composto di n^2 sottogriglie di dimensione $n \times n$. Il riempimento di ogni sottogriglia può essere visto come una permutazione di $\{1, 2, ..., n^2\}$. Perciò, un limite superiore grezzo per il numero dei quadrati Sudoku è dato da

$$\left[(n^2)!\right]^{n^2}.$$

Cominciamo osservando che ci sono n "fasce" nella griglia Sudoku. ("fasce" sono i gruppi di n righe successive). Il numero di modi per completare la prima fascia può essere stimato come segue. Abbiamo $n^2!$ scelte per la prima riga. Il numero di modi per completare la seconda riga è calcolato valutando il *permanente* della seguente matrice. Abbiamo una matrice $n^2 \times n^2$ le cui righe parametrizzano le celle della seconda riga. Notiamo che, per ogni cella, abbiamo $n^2 - n$ possibilità, dal momento che abbiamo già utilizzato n colori nella corrispondente sottogriglia $n \times n$. Il numero di modi per riempire la seconda riga è pari al numero dei corrispondenti possibili *sistemi rappresentativi distinti*. Più precisamente, consideriamo la matrice $n^2 \times n^2$ le cui righe parametrizzano le celle della seconda riga, e le cui colonne parametrizzano i numeri da 1 a n^2; poniamo 1 nell'ingresso *(i,j)esimo* se j è un valore ammissibile per la cella i, altrimenti poniamo 0. Questo ci dà una matrice *(0,1)* il cui *permanente* è il numero di modi per scegliere l'insieme dei *rappresentativi distinti* (*vedi nota 4*). Dall'equazione (2) questa quantità è delimitata da

$$(n^2-n)!^{\frac{n}{n-1}}.$$

Procedendo analogamente otteniamo il numero delle possibilità per la terza riga come

$$(n^2-n)!^{\frac{n}{n-2}}.$$

[8] vedi nota n° 6

In questo modo otteniamo una stima finale di

$$\prod_{k=0}^{n-1}(n^2-kn)!^{\frac{n}{n-k}}$$

per il numero dei modi di riempimento della prima fascia formata da n righe di un quadrato Sudoku di rango n. Supponiamo ora che *(i-1)* delle n fasce sono state completate. Calcoliamo il numero dei modi per completare la *i esima* fascia. In questo caso, daremo un limite superiore. Il numero di possibili ingressi per la prima cella della *i esima* fascia è

$$n^2-(i-1)n.$$

Così, come prima, il numero di modi per completare la prima riga della *i esima* fascia è delimitato da

$$(n^2-(i-1)n)!^{\frac{n^2}{n^2-(i-1)n}},$$

dal momento che per ogni cella, dobbiamo escludere i dati già inseriti nella colonna a cui la cella appartiene. Similarmente, il numero di modi per completare la seconda riga della *i esima* fascia è al massimo

$$(n^2-(i-1)n+1))!^{\frac{n^2}{n^2-(i-1)n+1)}}.$$

Procediamo in questo modo fino alla *i esima* riga della *i esima* fascia per ottenere una stima di

$$(n^2-((i-1)n+i))!^{\frac{n^2}{n^2-((i-1)n+i)}}.$$

Per la *(i+1) esima* riga, cambiamo la nostra strategia ed escludiamo i numeri già inseriti nella sottogriglia in cui cade la cella. Questo significa che abbiamo già inserito $i{\cdot}n$ elementi nella sottogriglia e dobbiamo escluderli per ottenere una stima di

$$(n^2-i\cdot n)!^{\frac{n^2}{n^2-in}},$$

per il numero di modi di riempire la *i esima* riga della *i esima* fascia. Procedendo in questa maniera, otteniamo che il numero S_n dei quadrati Sudoku di rango n è delimitato da

$$\prod_{i=1}^{n}(n^2-(i-1)n)!^{\frac{n^2}{n^2-(i-1)n}}$$

$$(n^2-[(i-1)n+1])!^{\frac{n^2}{n^2-[(i-1)n+1]}}.................(n^2-[(i-1)n+i])!^{\frac{n^2}{n^2-[(i-1)n+1]}}$$

$$(n^2-in)!^{\frac{n^2}{n^2-in}}...............(n^2-(n-1)n)!^{\frac{n^2}{n^2-(n-1)n}}.$$

Così

$$\frac{\log S_n}{n^2}\leq\sum_{i=1}^{n}\left(\sum_{j=0}^{i}\frac{\log(n^2-[(i-1)n+j])!}{n^2-(i-1)n+j}+\sum_{k=i}^{n-1}\frac{\log(n^2-kn}{(n^2-kn}\right).$$

Noi stimiamo la somma usando la formula di Stirling. Essa è

$$\leq\sum_{i=1}^{n}\left(\sum_{j=0}^{i}\log(n^2-[(i-1)n+j])+\sum_{k=i}^{n-1}\log(n^2-kn)\right)-n^2+O(\log^2 n).$$

Questo è facilmente visibile per essere

$$\leq\sum_{i=1}^{n}\left(i\log(n^2-(i-1)n)+\sum_{k=i}^{n-1}\log(n^2-kn)\right)-n^2+O(\log^2 n).$$

Così

$$\frac{\log(S_n)}{n^2}+n^2+O(n\log n)\leq\frac{1}{2}n^2\log n+\sum_{i=1}^{n}\left(i\log(n-i+1)+(n-i)\log n+\sum_{k=i}^{n-1}\log(n-k)\right).$$

La sommatoria su k, mediante la formula di Stirling, è

$$\log(n-i)!=(n-i)\log(n-i)-(n-i)+O\left(\log n\right).$$

Combinando tutto si ha:

$$\frac{\log S_n}{n^2}\leq 2n^2\log n-2,5n^2+O\left(n\log n\right),$$

da cui il teorema è facilmente deducibile.
Abbiamo anche il seguente corollario.

7.3 Corollario 7.

Sia p_n la probabilità che un quadrato Latino di ordine n^2, scelto casualmente, è anche un quadrato Sudoku. Allora

$$p_n \leq e^{-0{,}5n^4+O\,(n^3 \log n)}.$$

In particolare, $p_n \to 0$ se n tende all'infinito.

7.4 Dimostrazione del corollario 7.

Dal teorema 6 sappiamo che il numero dei quadrati Sudoku di rango n è al massimo

$$n^{2n^4} e^{-2{,}5n^4+O\,(n^3 \log n)}.$$

Il numero dei quadrati Latini di rango n^2, dal corollario 5, è almeno

$$n^{2n^4} e^{-2n^4+O\,(n^2 logn)}.$$

Pertanto, la probabilità che un quadrato Latino di ordine n^2, scelto a caso, è anche un quadrato Sudoku è inferiore a

$$e^{-0{,}5n^4+O\,(n^3 \log n)},$$

valore che tende a 0 per n che tende all'infinito.

Infatti, il teorema mostra che il numero dei quadrati Sudoku è sostanzialmente più piccolo del numero dei quadrati Latini.

8 Osservazioni conclusive.

E' interessante notare che il puzzle Sudoku è estremamente popolare per una varietà di ragioni. La prima è di essere sufficientemente difficile da rappresentare una seria sfida mentale per chiunque tenti di risolverlo. Poi perché, semplicemente facendo la scansione delle righe e delle colonne, è facile introdurre i "colori mancanti", e questo dà al solutore l'incoraggiamento a persistere. Il novizio, solitamente, resta perplesso dopo qualche tempo. Tuttavia, il gioco può essere risolto sistematicamente annotando i colori inutilizzati in ogni riga, colonna e sottogriglia. Un semplice processo di eliminazione spesso porta al completamento del gioco. Qualche puzzle classificato come "demoniaco" comporta una versione un po più raffinata di questo processo di eliminazione, ma la strategia generale resta la stessa. Si potrebbe arguire che il Sudoku sviluppi abilità logiche necessarie al pensiero matematico.

Ciò che è degno di nota è il fatto che questo semplice gioco ha suscitato diversi problemi matematici che devono ancora essere risolti. Abbiamo già menzionato il problema del "minimo puzzle Sudoku", dove ci chiediamo se c'è un Sudoku, con 16 o meno dati d'ingresso, che ammette una soluzione unica.

Abbiamo già commentato il fatto che se si usano solamente 7 o meno colori, il gioco non ammette una soluzione unica. Ci si può chiedere se esiste una configurazione con solo 8 colori e con un numero minimo di dati. In figura 12, proponiamo un tale puzzle che ha solamente 17 dati d'ingresso (tratto da *vedi nota 1*).

							1	2
				3	5			
			6				7	
7						3		
			4			8		
1								
			1	2				
	8						4	
	5					6		

Figura 12. Un Sudoku con 8 colori e 17 dati d'ingresso.

Questi interrogativi suggeriscono la domanda più generale della determinazione del "Sudoku minimo" per il puzzle di rango n. Sarebbe interessante determinare la crescita asintotica di questo minimo come funzione di n. La nostra discussione mostra che il minimo è almeno $n^2 - 1$. E' vero che il minimo è $o(n^4)$?

Un altro problema interessante è di determinare il numero di quadrati Sudoku di rango n. Più precisamente, se $p_{\chi_n}(\lambda)$ è il polinomio cromatico del grafo Sudoku di rango n, qual'è il comportamento asintotico di $p_{\chi_n}(n^2)$? Se S_n è il numero di quadrati Sudoku di rango n, sembra ragionevole congetturare che

$$\lim_{n\to\infty} \frac{\log S_n}{2n^4 \log n} = 1.$$

E' chiaro che questo limite, se esiste, è minore o uguale a 1. Abbiamo già notato le varie simmetrie dei quadrati Sudoku. Per esempio, applicando una permutazione agli elementi $\{1, 2, ...n^2\}$, ci costruiamo un nuovo quadrato Sudoku. Perciò, partendo da un tale quadrato, possiamo costruire $n^2!$ nuovi quadrati Sudoku. Ci sono anche le permutazioni delle fasce, pari a *n!*, come anche le permutazioni delle pile, anch'esse pari a *n!*. Possiamo permutare le righe all'interno delle fasce come le colonne all'interno delle pile e queste determinano $n!^n$ simmetrie. Inoltre, possiamo prendere la trasposta del quadrato. In totale, questo genera un gruppo di simmetrie, che può essere visto come un sottogruppo di ξ_{n^4}.[9] Sarebbe interessante determinare dimensione e struttura di questo sottogruppo. Se concordiamo che due quadrati Sudoku sono *equivalenti* se l'uno può essere trasformato nell'altro eseguendo un sottoinsieme di queste simmetrie, allora un quesito interessante riguarda il comportamento asintotico del numero S_n^*, del numero di quadrati Sudoku di rango n inequivalenti.

Il problema della determinazione del comportamento asintotico di S_n e di S_n^* sembra presentare difficoltà come l'enumerazione dei quadrati Latini di rango n. In questo contesto, ci sono dei risultati parziali. Un rettangolo Latino k *x* n è una matrice k *x* n, con elementi $\{1, 2, ...n\}$, tale che nessun elemento compaia più di una volta in ogni riga o colonna. Godsil e Mackay[10] hanno definito una formula asintotica per il calcolo dei rettangoli Latini k *x* n per $k = o(n^{6/7})$. Questo suggerisce di considerare la nozione del *rettangolo Sudoku* di rango *(k,n)* con n^2 colonne e *kn* righe con elementi $\{1, 2, ...n\}$ tale che nessun elemento compaia più di una volta in ogni riga, colonna o sottogriglia n *x* n. Sarebbe possibile estendere tale metodo[11] per studiare il comportamento asintotico dei rettangoli Sudoku di rango *(k,n)* per k in determinati intervalli.

Ringraziamenti. I nostri ringraziamenti vanno a Cameron Franc, David Gregory, e Akiko Manada per gli utili commenti sulla precedente versione di questo scritto. Ringraziamo anche Jessica Teves per l'assistenza nella ricerca letteraria.

[9] Il carattere tipografico nell'originale è una S gotica non disponibile nella mia versione LyX 1.6 con cui ho scritto questo testo.
[10] C. D. Godsil e B. D. Mackay, Asymptotic enumeretion of Latin rectangles, *J. Comb. Theory Ser.* B 48 (1990), no. 1, 19-44.
[11] vedi nota 10