ISKOI 2021 Cremona: Polimorfismo e conoscenza

Da Ranganathan ai lattice, sino al concetto di conoscenza

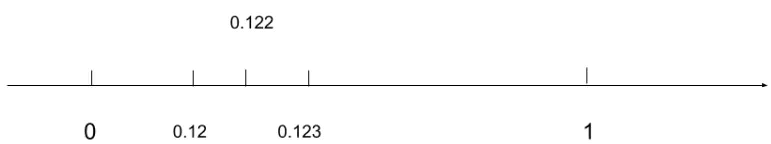
1. L'origine: da Dewey e Peano

Il 4 gennaio 1924 Ranganathan concorre a un bando per un posto da bibliotecario presso la *Madras University Library* e lo vince, lasciando il posto di insegnante. Ma non abbandona la matematica: entro il 1928 realizza venti pubblicazioni matematiche.

> I suoi successivi studi di classificazione a Londra sotto **W.C. Berwick Sayers** lo misero di fronte all'enorme numero e varietà dei soggetti presenti in uno schema di classificazione generale. Più significativamente, sembra avergli fornito il primo importante collegamento tra la grande espansione dei soggetti e i numeri ordinali nella forma della DDC. (miksa)

Dopo un primo momento di difficoltà, durante il suo viaggio a Londra, sembra che Ranganathan riesce a vedere una soluzione alternativa al problema della classificazione. Ma vediamo seguendo i vari passaggi quale potrebbe essere stato lo scenario.
SRR inizia ad apprendere la classificazione Dewey e scopre che ad ogni numero corrisponde una stringa di soggetto, che sono classi disgiunte che rappresentano la realtà attraverso la struttura di un albero, simile alla classificazione di Linneo. Ma scopre qualcosa di più affascinante, il numero 12 (filosofia, altri problemi speculativi) è più generale di 122 (filosofia, altri problemi speculativi, casualità). Inoltre, 122 si dispone su una retta subito prima di 123 (filosofia, altri problemi speculativi, libertà libero arbitrio). Allora perché non vedere la DDC come una successione di soggetti rappresentabili su una retta con valori compresi tra 0 e 1? Per Ranganathan la classificazione biblioteconomica diventa un linguaggio artificiale dei numeri ordinali per meccanizzare la disposizione dei soggetti.
Prima di tutto cosa sono i numeri ordinali. In matematica ci sono i numeri cardinali che rappresentano una magnitudine e poi i numeri ordinali che hanno una posizione sulla retta, sono appunto ordinati secondo la magnitudine. Ranganathan si era accorto che la *Decimal Dewey Classification* rappresentava soggetti di libri attraverso i numeri ordinali, che quindi tali soggetti potevano essere **non solo ordinati ma anche manipolati** da funzioni matematiche.
SRR, potrebbe aver conosciuto la teoria di Peano (Peano Sur une curve pag 157), per cui dati due numeri X e Y, compresi tra [0,1), di n/2 cifre, è possibile costruire un numero unico T, compreso tra [0,1) di n cifre. E da questa teoria potrebbe essere stato portato a vedere i concetti come uno spazio multidimensionale. La storia racconta che durante il suo viaggio a Londra vedendo il meccano, gli sia venuto in mente di rappresentare le stringhe di soggetto come composte da parti che si uniscono, come aspetti che si mettono in relazione, come dimensioni di uno spazio multidimensionale che nel punto trovano la loro essenza.

La DDC vista come la disposizione di numeri lungo una retta. I numeri reali del segmento [0,1)



In **topologia** tra i punti di uno spazio vettoriale si può calcolare il concetto di ***distanza***. Così come sulla retta il punto 3 ha distanza 2 da il punto 5, allo stesso modo, punti in due, tre ed n dimensioni hanno una distanza relativa da punti con la stessa dimensione. Allo stesso modo il numero 122 della DDC dista 1 dal numero 123. Il concetto di distanza unito a quello di ***outer product*** (che costruisce lo spazio vettoriale), di combinatoria e di trasformazione (l'idea di Peano può essere considerata una trasformazione da uno spazio a 2 dimensioni ad uno monodimensionale in cui i punti di una unità quadrata stanno su una linea di segmento) saranno le basi della teoria biblioteconomica di Ranganathan.

2. Prolegomena parte Q

Capitolo QA

E' utile ricordare cosa per Ranganathan (Ranganathan Prolegomena Cap.Q) fosse il parametro o la caratteristica. In matematica un *parametro* è una costante che compare nell'equazione di una curva o di una superficie al variare della quale si ottiene la famiglia di curve e di superfici. Allo stesso modo nella catalogazione, ciascuna delle successive *caratteristiche* che si usano per arrivare a quella che lui chiamava l'*idea isolata*, possono essere definiti i parametri della classificazione. Preso l'idea isolata di "professore" e le due caratteristiche "materia" e "capacità retorica", possiamo rappresentare il concetto con questi due insiemi: PM = {Pc,Pe,Pl} dove Pc è il professore di chimica, Pe è il professore di inglese e Pl il professore di diritto e l'insieme dei professori PR = {B,M,D} dove B è il professore brillante, M il mediocre e D il noioso (dull). Ritornando alla metafora del parametro è come se avessimo la funzione PM(x) in cui x può assumere i valori {Pc,Pe,Pl}. Se vogliamo rappresentare il concetto di professore dobbiamo creare uno spazio vettoriale a due dimensioni con nelle ordinate la caratteristica “materia” e nelle ascisse la caratteristica “capacità retorica”.

Capitolo QB

Scoperto che i concetti sono spazi multidimensionali dobbiamo trovare un modo per trasformarli in uno monodimensionale, in quanto i libri vengono esposti sugli scaffali e quindi su una successione lineare dove c’è la relazione di immediata vicinanza. Un aspetto che vale la pena sottolineare di questo processo è che l’ordine con cui si dispongono le categorie sugli assi dello spazio vettoriale cambierà l’ordine in cui i concetti appariranno nello spazio monodimensionale e quindi le relative relazioni di immediata vicinanza. Immaginiamo il nostro libro sullo scaffale, sarà circondato da altri libri e le distanze degli altri libri dal nostro creano quello che Ranganathan chiamava “relazione di immediata vicinanza”.
La nostra linea di scaffale sarà disposta in modo tale che il libro cercato sarà circondato da libri molto simili riguardo il soggetto trattato e, a mano a mano che ci si allontana, o verso destra o verso sinistra, i libri saranno sempre più alieni. Quello che Ranganathan chiamava la disposizione APUPA (Alieno, Penumbral, Umbral, Penumbral, Alieno).

Capitolo QC

Iniziamo partendo da un concetto definito da una sola caratteristica. Abbiamo il concetto professore definito dalla caratteristica “materia”. Quindi il nostro insieme PM = {Pc,Pe,Pl}. Se facciamo un mapping con l’insieme {01, 02, 03} avremo 6 possibili combinazioni (3!). Scelta la combinazione a noi più favorevole, avremo mappato il nostro concetto su una retta in cui i valori della caratteristica assumeranno un ordine. Per esempio Pc=01, Pe=02, Pl=03. Nella terminologia di Ranganathan il professore di chimica è la remove 1, quello di inglese, la remove 2 e quello di diritto la remove 3. La *remove* rappresenta la posizione sulla retta, quella che potremmo chiamare il numero ordinale del valore della caratteristica sulla retta. L’ordine delle remove è la relazione di immediata vicinanza. Come vedremo nel prossimo capitolo, aggiungendo caratteristiche si creano delle spazi vettoriali a più dimensioni dove l’ordinamento sugli assi dei concetti e l’attribuzione di una caratteristica ad un particolare asse crea delle invarianti sull’ordinamento finale.

Capitolo QD

Dato V uno spazio vettoriale a due dimensioni nel campo N dei numeri naturali, si possono rappresentare spazi di categorie multidimensionali. Preso l’insieme PM e PR (del sotto-capitolo QA) e fatto un mapping tra l’insieme PM e {01,02,03} e l’insieme Pr e {01,02,03} otteniamo così che Pc=01, Pe=02 , Pl=03, B=01, M=02 e D=03. Fatto l’outer product tra PM e PR otteniamo una matrice A di 3*3=9 elementi con tutte le possibili combinazioni di due vettori di tre elementi. Ora il prodotto vettoriale viene costruito in modo di avere il campo della matrice A_{ij} non come $PM_i \mid PR_j$ ma come:

$$charAt(0)_{PM_i} \mid charAt(0)_{PR_j} \mid charAt(1)_{PM_i} \mid charAt(1)_{PR_j} \qquad \text{(F1)}$$

dove charAt(n) è la funzione che restituisce il carattere n della stringa in pedice e l'operatore a|b concatena la stringa a con la stringa b. Nel nostro caso presi ad esempio Pc (01) e M (02) avremo la stringa 0012. (vedi (Bertani))
Nella figura 1 possiamo vedere cosa otteniamo.
Attraverso l'outer product dei due insiemi PM e PR abbiamo un latice con relativo meet e join, da cui si ricava un insieme ordinato monodimensionale. Il latice è una struttura geometrica regolare che si ripete, dotata di funzioni di ordinamento. Ogni suo membro è in relazione di $\leq$ (minore o uguale) o $\geq$ (maggiore uguale) rispetto agli altri membri. Dire che esiste un meet ed un join significa affermare che c'è un minimo assoluto e un massimo assoluto. Attraverso una trasformazione, abbiamo ottenuto un manifod 1 da uno spazio bidimensionale. Siamo riusciti a rappresentare un concetto multidimensionale lungo una linea, il nostro scaffale. Un aspetto importante è che invertendo gli assi, per esempio mettendo il parametro "materia" sull'asse x e quello di "retorica" sull'asse y, l'ordinamento finale cambia perchè l'invariante della relazione di immediata vicinanza diventa "retorica". Questa capacità delle basi di cambiare lo spazio vettoriale introduce il concetto di two-fold infinity. C'è l'ordinamento di chi dispone lo scaffale e l'ordinamento che il lettore si aspetta che ci sia sullo scaffale. Tanto più si discostano le due prospettive tanto più il lettore farà fatica a trovare quello che cerca.

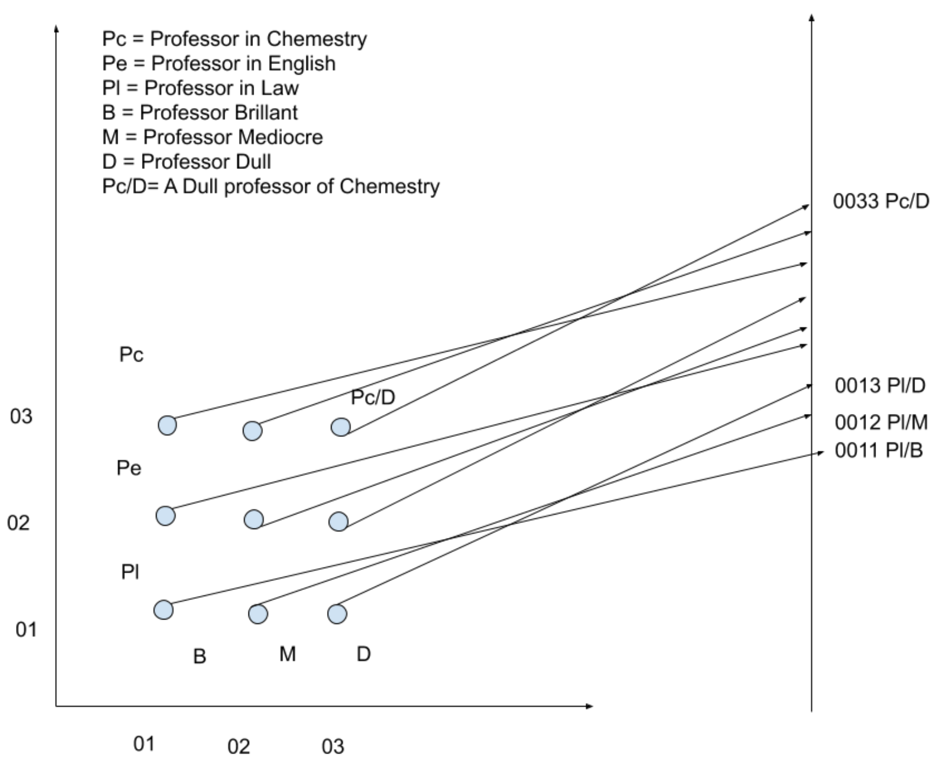


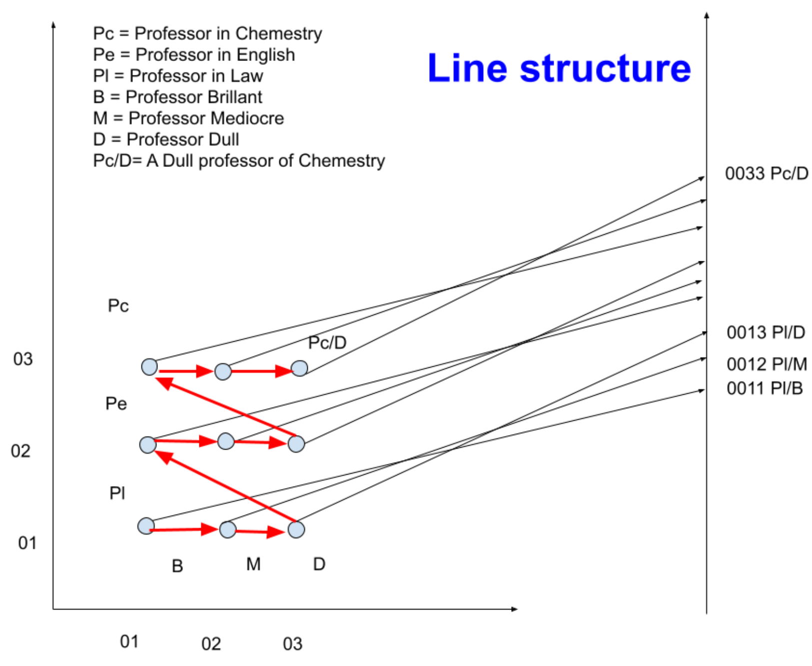
figura 1

Capitolo Q1

Se come abbiamo visto le linee sono insieme di punti, combinazione di aspetti del concetto, alcuni concetti possono non avere punti in comune e allora le linee andranno all'infinito senza mai incontrarsi, altri avranno punti in comune, in quei punti i concetti condividono combinazioni di aspetti. Per esempio immaginiamo una definizione di professore con caratteristica “grado di studio”, PG={Pg, Pu, Po} dove Po è il professore della scuola dell’obbligo, Pu è quello universitario e Pg il valore generico, ossia valore non riportato, sconosciuto. Costruita la matrice A = 3*3*3 = 27 unendo gli aspetti PM, PR e PG, fatto il mapping sulla retta come visto nel capitolo QC avremo che alcuni punti della linea trovata nel mapping PM,PR coincidono con la linea creata con mapping PM,PR,PG. Tali punti sono quelli in cui PG è uguale a Pg. Si crea una cosa molto simile ai sottospazi che potrebbe essere un esempio di relazione “is a”. Maggiori caratteristiche permettono di scendere nella specificità. Più è alta la dimensione dello spazio più è alta la specificità del concetto. Tenere una caratteristica non settata ci permette di risalire al valore più generico (nel nostro caso a tre dimensioni, da un piano si torna ad una linea).

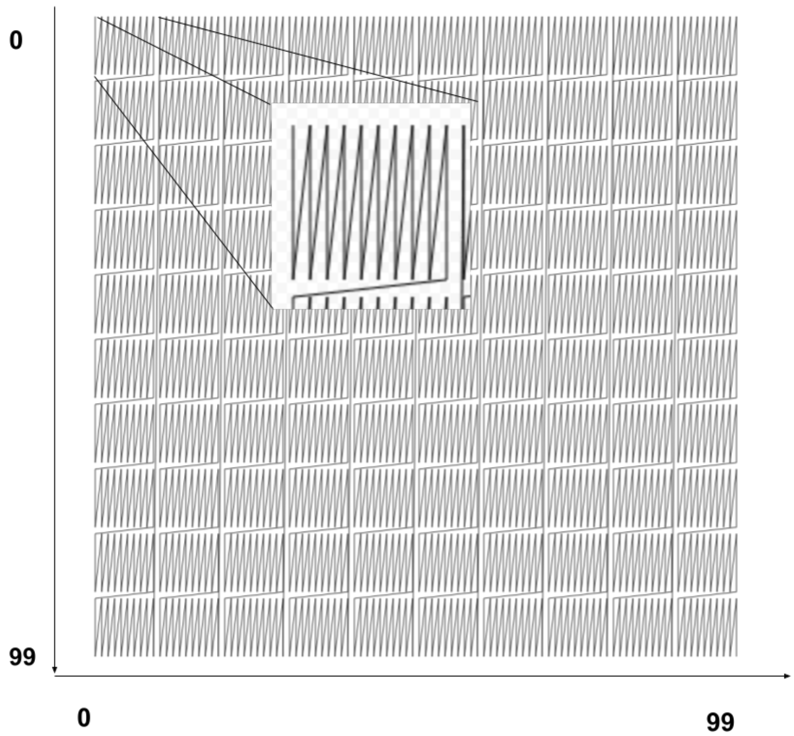
Capitolo Q2

Chiamiano line structure la parte della teoria che ci permette di rappresentare un quadrato (nel caso bidimensionale) come una linea. Utilizzando la formula (F1) passiamo da un quadrato ad una linea attraverso una funzione biettiva diversamente da quanto avviene per la filling curve di Hilbert per cui la funzione inversa, da linea a quadrato, associa 4 punti sul quadrato per ognuno della linea.



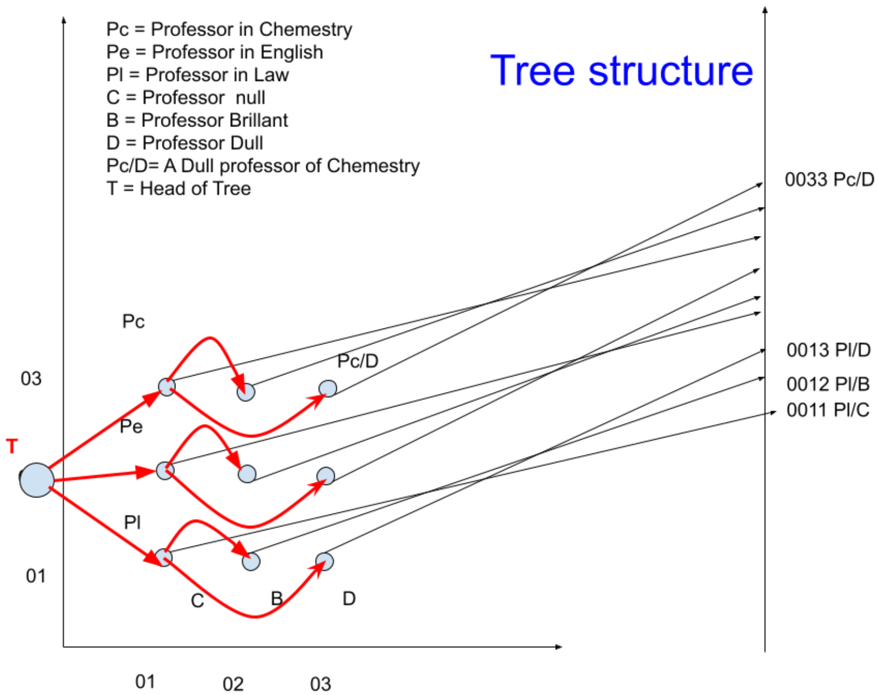
Le frecce rosse rappresentano come si muove la space filling curve all’interno del nostro quadrato. Le lunghezza delle frecce indica il valore assoluto della distanza tra i due punti successivi. Quindi se ipotizziamo che tra Pl/M e Pl/D ci sia distanza uno, tra Pl/D e Pe/B ci sarà distanza 2,5. Questo è in linea con la disposizione dei nostri concetti, Il professore di diritto mediocre e più vicino al professore di diritto noioso che il professore di diritto noioso al professore di inglese bravo.

La figura successiva rappresenta il modello precedente per due variabili con valori compresi tra 0-99



Capitolo Q3

Un’altra struttura che emerge dal nostro piano cartesiano è quella dell’albero. Gli alberi in informatica permettono la ricerca in tempo logaritmico e quindi sono molto utili quando si devono gestire molti dati. Se la struttura minima del dato che vogliamo rappresentare ha una testa che sorregge le altre parti, simile alla frase Soggetto - Predicato -Oggetto, dove il Soggetto regge il Predicato e il Predicato l’Oggetto, non c’è bisogno di inserire valori generici nelle dimensioni successive alla prima, come mostrato nel capitolo Q1. Negli altri casi si è costretti ad inserirli per una corretta costruzione semantica dell’albero.
Ipotizzando uno spazio a tre dimensioni [x1,x2,x3], l’albero che costruiremo avrà i valori di x1 che sorreggono come figli i valori di x2 i quali a loro volta sorreggeranno come figli i valori di x3. Ipotizzando x1 come Soggetto, x2 come Predicato ed x3 come Oggetto, si possono rappresentare con questa struttura i knowledge graph.
La figura successiva mostra la costruzione dell’albero in uno spazio a due dimensioni con l’utilizzo del valore generico C.



Capitolo Q4

Vedere lo stesso dato con relazioni differenti tra le parti e quindi con una figura finale differente, nel nostro caso un piano cartesiano, una linea e un albero, ci permette di introdurre il concetto di polimorfismo e di conoscenza. Le funzioni che permettono di passare da una struttura all'altra sono atti di conoscenza. Riordinare, eliminare, aggiungere relazioni a un dato, sono tutte funzioni di conoscenza. E tali funzioni di conoscenza permettono l'utilizzo dello stesso dato in ambiti e con finalità differenti. Nel nostro caso il piano cartesiano ci mostra tutte le possibili combinazioni del dato ed è il più ricco a livello informazionale. La linea ci permette la disposizione a scaffale. L'albero la ricerca informatica.
Quello che SSR ha avuto ai Grandi Magazzini vedendo il meccano ed immaginando le basi della Colon Classification è molto simile ad un atto di conoscenza, una ristrutturazione del dato che porta alla soluzione di un problema

Bibiografia

Bertani, Mauro. "Codici Dewey computazionalmente economici." ISKO Italia, 2019, http://www.iskoi.org/doc/firenze19/bertani.htm.

Miksa, Frank. "The influence of mathematics on the classificatory thought of S.R.Ranganathan." Knowledge Oganization for Information Retraival, 1997, pp. 167-179.

Peano, Giuseppe. “Sur une curve qui remplit toute une aire plane.” Mathematische Annalen, vol. 36, pp. 157-160. DOI:10.1007/BF01199438.

Ranganathan, R. S. Prolegomena to Library Classification. New York, Asia Publishing House, 1967. 1 vols.