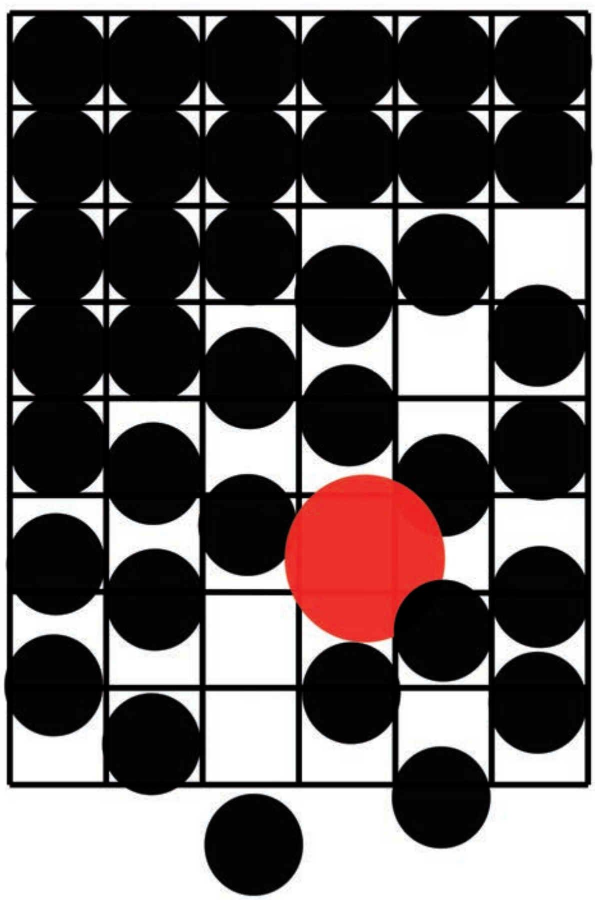
Roberto Spagnuolo

EVOLUZIONE DEI SISTEMI COMPLESSI E REGOLAMENTAZIONE



FrancoAngeli

ad Elena

Non est quod timeas ne operam perdideris, si tibi didicisti.

Lucio Anneo Seneca

Roberto Spagnuolo

EVOLUZIONE DEI SISTEMI COMPLESSI E REGOLAMENTAZIONE

FrancoAngeli

In copertina: elaborazione grafica dell'Autore.

Indice

Introduzione

La domanda che mi vado ponendo ormai da anni, e che è divenuto una sorta di fil rouge di pressoché ogni mio interesse culturale, è questa: quale è il ruolo della regolazione imposta ad un sistema, nell'evoluzione del sistema stesso? Un sistema sociale prevede delle regole nell'intento di stabilizzare il sistema. Ora la mia domanda è quindi: vi è un valore ottimale della regolazione? La incertezza, la follia (Erasmo, 2012) devono avere un loro spazio per stabilizzare il sistema o devono essere bandite in modo che il sistema sia stabilizzato su comportamenti predefiniti?

Questa domanda la ritengo fondamentale. Ognuno infatti dovrebbe interrogarsi su quale è il suo comportamento rispetto al sistema in cui agisce, e far scaturire le proprie convinzioni non da luoghi comuni, ma da precise e maturate convinzioni. La verità nessuno la ha in tasca, ma la domanda è comunque un dovere perché è il dubbio (Cartesio, 2012) il seme della conoscenza.

Il problema se lo sono posto non pochi filosofi. Il filosofo più noto relativamente a questo argomento è Rousseau (Rousseau, 2009). Egli vedeva un contrasto tra libertà individuale necessaria perché l'individuo raggiunga la sua finalità, e le regole del sistema, le leggi, che necessariamente limitano tale libertà. Egli giunse ad una conclusione piuttosto drastica sulla quale sono state costruite teorie sociali spesso catastrofiche. Egli sosteneva che la finalità dell'individuo coincide con quella dell'insieme degli individui per cui anche chi non conosce tale finalità deve perseguirla essendo a ciò forzato dalle regole collettive. Il problema, ammesso e non concesso che in un sistema che tende soprattutto alla stabilità – prerequisito per ogni vantaggio individuale – si renda opportuna una uniformità di obiettivi individuali, il problema è: chi formalizza questa finalità collettiva-individuale e chi la traduce in regole di comportamento?

A questo problema Rousseau dà una risposta poco convincente: l'assemblea. Cioè un insieme di individui rappresenta statisticamente il fine della collettività. Questa idea, la cui validità è tutta da dimostrare, è tutt'ora alla base di molti sistemi di regolazione sociale. Una obiezione a tale metodo è semplice: è evidente la disomogeneità degli individui e pertanto il loro "peso" in una decisione collettiva. Inoltre, in genere, è tramite la parola che si influenzano le opinioni altrui e non è affatto dimostrato che l'opinione migliore sia quella meglio espressa.

Mi sono orientato, in tanti anni di riflessione, a considerare la soluzione di questo problema nella logica della dinamica dei sistemi complessi secondo la quale è possibile individuare uno spazio limitato intorno al quale il sistema può oscillare (quello che viene chiamato attrattore). Lo spazio deve essere limitato e il sistema trovare il suo stato stazionario oscillando, ma lo stato stazionario non deve essere uno stato di equilibrio, altrimenti il sistema "muore", cessa di esistere.

A livello colloquiale, si faccia caso a quei gruppi di amici dove vi è qualcuno che si comporta fuori delle righe. Spesso viene criticato per i suoi eccessi, le sue battute non sempre eleganti, ma egli è spesso la molla del sistema. Se non ci fosse, il gruppo si arenerebbe in discorsi ricorrenti, magari equilibrati, ma per questo per nulla evolutivi. Un gruppo sociale deve muoversi, muoversi al meglio su quell'orlo del caos del quale Paul Valery, poeta e matematico, diceva: "tra ordine e disordine vi è un momento meraviglioso".

Nel trattare questo argomento vi sono in genere due vie, vie imposte dalla complessità dell'argomento e dalla relativa novità che ancora porta in sé. Una è quella delle opinioni, opinioni che sono discutibili ma non confutabili, l'altra via è quella del formalismo scientifico, formalismo che però offusca le opinioni, diviene un discorso tra specialisti e soprattutto ha il limite del dimostrato e non del dimostrabile.

Non sono uno specialista di dinamica dei sistemi complessi ed oltretutto su questo tema vi sono ancora molte cose che devono ancora essere studiate. Si consideri che la "teoria del caos" risale, non con questa denominazione, a Poincaré, e si sviluppa come teoria autonoma con Lorenz negli anni '60 del secolo scorso.

Ora il problema stilistico che mi pongo è quanto una opinione debba e possa essere confortata da una formulazione formale e quanto una formulazione formale limiti una opinione. Questo, ripeto, in un campo ancora inesplorato. Pertanto sono arrivato alla conclusione che un'esposizione formalizzata possa essere un formalismo dell'opinione. Ovvero possa non avere un suo spazio autonomo, ma essere impiegato per generare delle immagini con un formalismo consolidato, di opinioni altrimenti difficili da illustrare.

Vi sono alcuni sviluppi recenti della conoscenza che possono essere di stimolo e possono costituire un sottofondo culturale di conoscenza che ci aiuta a inquadrare il problema dei sistemi dinamici in un ambito più libero. Infatti il primo passo da fare è quello di aprirsi a informazioni che possono anche collidere con le nostre convinzioni o addirittura con il buon senso. Potremo sempre rigettarle, ma prima è opportuno valutarle senza posizioni preconcette.

I temi che trovo rilevanti sono moltissimi e nel seguito solo pochi di essi troveranno spazio per una esposizione anche sommaria.

Uno dei problemi di base è il contrasto tra la realtà come la percepiamo secondo i canoni più immediati e come in effetti essa potrebbe essere sotto la superficie dell'apparenza. L'apparenza della realtà, un po' per pigrizia un po' per rassicurarsi, pare governata da leggi semplici e da principi di causa effetto. Che ciò non sia vero è inquietante e quindi è un pensiero che si tende istintivamente a rifiutare. Chi non è "addetto ai lavori" può farlo tranquillamente, ma ciò determina una carenza culturale diffusa. Almeno interrogativi, se non soluzioni, il mondo attuale in evoluzione verso la complessità più esplicita dovrebbe suggerirne ad ognuno.

Del resto la fisica newtoniana è sufficiente a descrivere il mondo in cui ci muoviamo, ma si sa che ha dei limiti che si palesano quando si superano i limiti del quotidiano.

I concetti di complessità, di sistemi dinamici, di caos, di biforcazioni, di entropia e di informazione non fanno parte del bagaglio culturale di ognuno e tanto meno sono temi della istruzione scolastica, solidamente ancorata a una cultura ottocentesca, determinista se non illuminista.

Il problema però è che la complessità è entrata nella nostra vita quotidiana e pertanto, anche se la cultura istituzionale, la cultura della didattica di Stato, la ignora, è indispensabile fare i conti con essa.

La stessa "democrazia" è un concetto vuoto se non se ne specificano i termini. Molte persone hanno parole chiave come libertà e democrazia che ritengono concetti semplici e senza sfumature, contraddizioni, mezzi di attuazione incerti. La statistica applicata a sistemi politici mostra che essa non è affatto una risposta significativa alle esigenze della collettività.

Nel passato a questi problemi, anche politici, vi sono state risposte in genere estreme che hanno oscillato tra la dittatura di pochi alla anarchia. In effetti è molto più ragionevole ritenere che tra un minimo ed un massimo assoluti vi possa essere uno stato stazionario ma non di estremo, non di equilibrio.

Come si vede già in questa affermazione banale, la terminologia matematica, presa con i suoi formalismi anche più blandi, può aiutare a costruire immagini che trovino una solidità nel formalismo stesso.

Ciò può portare all'obiezione che il solo formalismo non può condurre alla dimostrabilità di ciò che rappresenta. Ciò è vero, ma è anche vero che esso consente di dare forza ad un'immagine senza quella gratuità che altrimenti avrebbe.

Uno degli "oggetti" più complessi che oggi maneggiamo pur senza renderci conto che maneggiamo un oggetto costruito sull'orlo del caos, è il software, e da riflessioni sul software inizieremo il nostro viaggio.

Infatti questo libro, pur essendo un saggio in ambito tecnico, è, almeno nelle mie intenzioni, una panoramica su argomenti spesso non ancora pienamente risolti e sui quali delle riflessioni sono più produttive che non delle incerte conclusioni.

Era mio desiderio dare a questo libro un titolo che richiamasse una delle più singolari raccolte che io conosca: *Ab hic et ab hoc* di Americo Scarlatti. Il lettore infatti troverà informazioni su argomenti che un tecnico oggi deve conoscere, se non dettagliatamente nei metodi, come "cultura" scientifica. Si tratterà di perceptroni, del dilemma del prigioniero, di reti bayesiane. Ma insieme a questo tenterò di dimostrare, o almeno dare indicazioni per poter valutare, come le moderne teorie del caos e della complessità rimettono in discussione le convinzioni maturate negli ultimi secoli. Oltre a questo vi sono delle considerazioni che intendono suggerire degli ampliamenti del tema trattato e anche delle riflessioni che intendono inquadrare i temi tecnici trattati in un ambito culturale e politico più vasto.

Tornando al titolo che avrei voluto dare al libro, che è anche una chiave di lettura, un famoso proverbio ironizza sui discorsi delle donne: "Quando conveniunt Domitilla, Sibylla, Drusilla sermones faciunt et ab hic et ab hoc et ab illa". Et ab hic et ab hoc è un latinetto usato per evidenziare la caoticità di un discorso. Ma, soprattutto, *Et ab hic et ab hoc* è il titolo di una collana di 12 libri di *Stranezze, bizzarrie, scherzi e bisticci letterari*, come recita il sottotitolo, pubblicata dal 1915 e al 1934 da Carlo Mascaretti (1855-1928) meglio noto con lo pseudonimo di Americo Scarlatti. Gli ultimi volumi, dopo la morte di Mascaretti, vennero pubblicati dalla nipote sulla scorta degli appunti dell'autore.

Il primo e più sorprendete dei volumi, ripubblicato da Salani nel 1988 (Scarlatti, 1988), è dedicato alle Amenità letterarie. Per darne un'idea, non si sa che brano scegliere perché sono tutti imperdibili. Proviamo.

De Amicis, incontrando l'amico Ferdinando Fontana, lo prese all'improvviso per un braccio e declamò:

> Vedi?... là su quella via,
> Sotto l'ampio arco del ciel,
> Havvi un uomo che si avvia

Dietro il pigro somarel!

Eppur nulla, nulla esiste.
E illusion la nostra fu!...
Tutto è nulla!... Ahi cosa triste,
Non ci siam né io né tu!

Il giorno dopo, incontrando di nuovo l'amico, gli disse di averci ripensato e di aver modificato la seconda strofa:

No, la strada non esiste,
Non v'è l'uomo, non v'è il ciel!...
Pure un dubbio in me persiste:
Che rimanga l'asinel!...

E con il dito indicò se stesso.

La "stranezza", la irregolarità ovvero lo spazio fuori delle regole, del consueto, del solito e rassicurante sentiero, sono le uniche possibilità di "scoperta" cioè di una "variazione". La variazione è fondamentale, è la ricerca di uno stato stazionario più produttivo, è l'abbandono della certezza per il guadagno di spazi ancora non sfruttati da altri. È il viaggio, è il mito di Ulisse.

Alcune volte, quando qualcuno inventa una nuova favola, ci sono molti altri che la ripetono, la riscrivono, la chiosano, la spiegano, tentano cioè di essere regolari con la irregolarità, l'invenzione di altri, ma ciò lo fanno solo dopo che la favola è stata accettata, è divenuta cioè "regolare".

Qui abbiamo tentato di inventare nuove favole sul tema della complessità organizzata, favole, non verità, racconti più di fantasia e di voglia di vedere nuovi paesaggi, che di inchiodare su un foglio delle "verità". La particolarità di questo testo è il tentativo di raccontare delle favole appoggiandoci alla stampella della matematica.

Perché questo? Perché l'uso della matematica nella scoperta è una favola che racconta se stessa e, come tale, si mette in dubbio da se stessa. Cioè, in fondo, c'è uno strumento per stabilire una qualsiasi verità? No di certo e pertanto un uso critico, certe volte provocatorio, della matematica per raccontare favole, è uno strumento per esplorare senza mai dare certezze che fermerebbero il gioco eterno del mutamento, della invenzione, da "invenio", trovo, cioè, e si trova dove non si è mai cercato, non sui sentieri già battuti.

Non è molto noto che Gödel scrisse un teorema per dimostrare l'esistenza di Dio sul principio ontologico di Anselmo d'Aosta. Fu pubblicato postumo (Gödel, 2006), Gödel era incerto se farlo o meno. Sembra una favola con quei simboli misteriosi che all'inesperto (e per essere esperti ce ne vuole!) paiono più formule alchemiche che simbolismi di logica. Forse quel teorema era ed è un pezzo d'arte. Alfred Jarry, drammaturgo francese di fine '800, inventò la "patafisica", ovvero la "scienza delle soluzioni immaginarie" e scrisse, a proposito di Dio, nell'ultimo capitolo del suo *Gesta ed opinioni del dottor Faustroll patafisico* (Jarry, 1984), intitolato *Della superficie di Dio*: Dio è il punto tangente di zero e dell'infinito. E ciò dimostrandolo con formalismi matematici che potremmo definire della "matematica dell'assurdo". Ora: esiste il teatro dell'assurdo, della vita non si capisce molto e la si può considerare spesso assurda, quindi perché non dovrebbe esistere una matematica dell'assurdo?

1. *Il software*

Per dare una definizione di software, in italiano più opportunamente “programma di calcolo”, preferiamo partire dalla definizione di algoritmo. Un algoritmo è un procedimento formale che risolve un determinato problema attraverso un numero finito di passi. L’algoritmo è alla base della nozione teorica di calcolabilità: un problema è calcolabile quando è risolvibile mediante un algoritmo. La programmazione, e quindi il software, costituisce essenzialmente la traduzione di un algoritmo in un programma, a prescindere dal linguaggio in cui è scritto, che può essere effettivamente eseguito da un calcolatore.

Il software, prescindendo da ogni valutazione di complessità tecnologica, viene visto come una sorta di “corredo” per l’hardware (in italiano più propriamente elaboratore o calcolatore elettronico) che consente di ottenere la soluzione di un predefinito problema che possa trovare una rappresentazione logico-numerica.

Questa finalità consente di vedere il software come una scatola chiusa, la “black-box” di una certa letteratura. Ora se provassimo ad aprire questa scatola, entreremmo nell’ambito di un’analisi specialistica delle modalità di costruire questa scatola. Non è questo il nostro scopo. Il nostro intento invece è quello di descriverne la natura restandone “fuori”. Un po’ come chi guida una macchina può farlo semplicemente conoscendo la risposta della vettura a certi comandi e non necessariamente come tale risposta venga attuata. Col software però, poiché si ritiene sia ambito appunto di uno specialista, si tende ad evitare una conoscenza anche della “scatola nera” vista da fuori, cioè non solo di poter ignorare la meccanica di autovettura, ma anche le modalità di una guida consapevole.

Teniamo quindi la scatola chiusa ed esaminiamo la relazione tra ingresso ed uscita dei dati nella e dalla scatola. Vorremmo vedere se è possibile trovare dei criteri per comprenderla senza aprirla. Si tenga presente questa affermazione per l’approccio che seguiremo.

Se vediamo il software come una "scatola nera", potremo solo esaminare i dati in uscita in rapporto con quelli in ingresso. Questo approccio ci conduce a poter usare gli strumenti di una disciplina ben consolidata: la teoria dell'informazione. Infatti il software può essere visto come un canale di trasmissione opportunamente codificato nel quale le informazioni viaggiano con una probabilità che vi sia una relazione tra segnale in entrata e segnale in uscita perfettamente formalizzabile.

Nella teoria dell'informazione, se X è lo spazio dei segnali che possono essere trasmessi e Y è lo spazio dei segnali che possono essere ricevuti, la probabilità condizionata, indicata con p(Y|X), è la probabilità di Y, dato X ed è una proprietà del canale, nella teoria dell'informazione, legata al "rumore" del canale che nella nostra analogia con il software indica la probabilità di ottenere la soluzione Y dato X.

Per fare un po' di chiarezza impiegando formulazioni con le quali non si ha sempre familiarità, la probabilità condizionata p(y|x) è legata alla probabilità congiunta p(y,x) dalla relazione:

$$p(y \mid x) = p(y,x) / p(x) \text{ [1.1]}$$

Il software, la "scatola nera", stabilisce quindi una relazione probabilistica tra lo spazio dei dati e quello dei risultati. Non esaminiamo quindi le modalità di una eventuale elaborazione dei dati, ma la capacità di mettere in relazione stati di ingresso e di uscita. Cioè di valutare che si ottenga y_i, se è stato trasmesso x_i cioè la probabilità condizionata $p(y_i \mid x_i)$ che si legge appunto: probabilità di y_i dato x_i.

Useremo un esempio, rappresentato nella tabella 1.1. La probabilità condizionata di x_1 dato y_1, è data dalla 1.1 ed è pari a 3/20 / 1/5 = 3/4.

La tabella 1.1 si può vedere come un canale di trasmissione leggermente disturbato. Se infatti si "riporta" il valore di 1/20 fuori diagonale sommandolo al termine diagonale si ottiene la tabella 1.2.

Nella tabella 1.2 è immediato rilevare che la probabilità condizionata $p(y_1|x_1)$ sia l'unità, cioè la certezza.

La probabilità condizionata delle probabilità y_i,x_i, che cioé mettano in relazione un evento con un solo esito, estesa a tutti i possibili eventi, è data dalla relazione:

$$P(Y \mid X) = \sum_i \frac{p(y_i,x_i)}{p(x_i)} p(y_i) \text{[1.2]}$$

dove si impiega la probabilità marginale p(y) di ogni possibile uscita.

Nel caso della tabella 1.1, P(Y | X) = 0.36. Se immaginiamo una distribuzione uniforme delle probabilità congiunte p(Y | X)= 0.2. Si deve notare che in questo modello assumiamo che tutte le variabili siano dipendenti ovvero che per nessuna p(y)=0.

P(Y,X	x_1	x_2	x_3	x_4	x_5	P(y)
y_1	3/20	0	0	0	0	3/20
y_2	1/20	1/5	0	0	0	1/4
y_3	0	0	1/5	0	0	1/5
y_4	0	0	0	1/5	0	1/5
y_5	0	0	0		0	1/5
P(x)	1/5	1/5	1/5	1/4	3/20	

Tab 1.1. Tabella delle probabilità congiunte tra gli eventi x e gli esiti y che rappresenta la incertezza che a un simbolo in entrata corrisponda uno ed un solo simbolo in uscita.

p(y,x)	x_1	x_2	x_3	x_4	x_5	p(y)
y_1	1/5	0	0	0	0	1/5
y_2	0	1/5	0	0	0	1/5
y_3	0	0	1/5	0	0	1/5
y_4	0	0	0	1/5	0	1/5
y_5	0	0	0	0	1/5	1/5
p(x)	1/5	1/5	1/5	1/5	1/5	

Tab. 1.2. Come la tabella 1.1, ma qui le probabilità congiunte indicano che effettivamente ad ogni simbolo in entrata corrisponde un solo simboli in uscita.

Riferendosi alla teoria dell'informazione, si può parlare di informazione mutua I(Y|X). Date, cioè, due variabili casuali X ed Y con probabilità congiunta p(x,y), la mutua informazione è definita come entropia relativa tra p(x,y) e p(x), p(y). La mutua informazione rappresenta l'informazione che una certa variabile contiene circa un'altra variabile. L'informazione mutua è legata all'entropia dell'informazione dalla relazione:

$$I(Y \mid X) = H(Y) - H(Y \mid X)[1.3]$$

La entropia condizionata H(Y|X) di due variabili casuali X ed Y con distribuzione di probabilità congiunta p(x,y) è data da:

$$H(Y \mid X) = \sum\nolimits^{y,x} p(y,x) \log_2 \frac{1}{p(y|x)} =$$
$$- \sum\nolimits^{y,x} p(y,x) \log_2 \frac{p(y,x)}{p(x)}[1.4]$$

Precisiamo che nella teoria dell'informazione si usa il logaritmo in base 2 con riferimento all'elemento di informazione minima, il bit, ma il cambio di base avviene impiegando una costante e pertanto nel nostro percorso, necessariamente qualitativo, la base del logaritmo è ininfluente.

Non stupisca il fatto che quando p(y|x) = 1, l'entropia sia nulla mentre ci si aspetta che la informazione aumenti. Ma in effetti va notato che se il risultato che si ottiene è atteso, non si ha alcuna informazione appunto perché il risultato era previsto, l'informazione è il risultato non atteso. In questo senso l'entropia H può essere vista come "incertezza" e con ciò evitando una certa perplessità sul concetto di "ordine" e "disordine" associati spesso al concetto di entropia. Qui ci limitiamo a notare che una sequenza di simboli tutti eguali ha entropia, e quindi informazione, nulla, mentre una sequenza di simboli molto differenziati ha maggiore entropia e maggiore informazione.

Va subito chiarito che il concetto di informazione qui è riferito al rapporto ingresso-uscita della black-box non è relativo all'uso che si può fare del risultato ottenuto dalla elaborazione dei dati. Quindi se si ritiene comunemente che il software produca informazione o ordini l'informazione, questo concetto si riferisce al rapporto che esso ha con il sistema esterno al funzionamento della scatola nera. Qui invece analizziamo solo l'affidabilità della scatola nera in termini probabilistici e cioè quale probabilità vi sia, dato uno stato in ingresso, di ottenere l'evento atteso in uscita. In questo senso il significato che assume l'informazione è molto diverso.

L'entropia, nella teoria dell'informazione, si può vedere anche come la possibilità di comprimere un messaggio: più è uniforme, meno complessa è la codifica compressa del messaggio. Un messaggio costituito da 100 simboli eguali, ad esempio A, si può codificare banalmente come 100A, cosa che non è possibile se i simboli sono tutti diversi. Quindi 100A ha meno informazione e minore entropia della successione di simboli diversi.

Il concetto di entropia associato al modello che abbiamo proposto per descrivere la "scatola nera" ci consente di fare altre osservazioni sul software che ereditiamo tutte dalla teoria dell'informazione e quindi da una teoria consolidata. È chiaro che non abbiamo l'intenzione di definire un'analoga metrica per il software. Di metriche del software ve ne sono molte e tutt'ora per nulla esaustive. Qui, lo abbiamo detto, ci preme condurre un'analisi qualitativa, ma basata su argomentazioni dimostrabili.

La teoria dell'informazione definisce "capacità del canale" C il valore della mutua informazione in funzione dell'errore di trasmissione p_e.

$$C(p_e) = 1 - H(p_e) \quad [1.5]$$

dove $H(p_e)$ è l'entropia associata alla probabilità di errore. Nel nostro caso:

$$p_e = \sum^i \sum_{i \neq j}^j p(x_i, y_j) \quad [1.6]$$

$$H(p_e) = -p_e \log_2 p_e - (1 - p_e) \log_2 (1 - p_e) \quad [1.7]$$

Per il calcolo, si tenga presente che

$$\lim_{x \to 0} x \log x = 0 \quad [1.8]$$

È interessante notare che nella formula della probabilità di errore, come del resto avevamo già intuito, la probabilità di errore è legata ai termini fuori diagonale che non relazionano i valori in ingresso con quelli attesi in uscita.

Se si assume che la funzione che esprime le variabili in entrata sia ordinabile e lo sia anche quella che esprime le variabili in uscita, è allora possibile mettere in rapporto biunivoco entrate ed uscite ordinando righe e colonne della matrice in modo che la larghezza di banda sia minima. Se la matrice è diagonale, la probabilità di errore è certamente nulla. In merito si potrebbe estendere il discorso alla diagonalizzazione della matrice legata agli autovalori di questa, ma non pare che questa tecnica possa portarci strumenti utili per l'analisi che stiamo conducendo.

La capacità di canale in funzione della probabilità di errore è data dalla 1.5. L'entropia della probabilità di errore che impiegheremo, a fini illustrativi, è relativa ad un evento binario il cui andamento è rappresentato in figura 1.1.

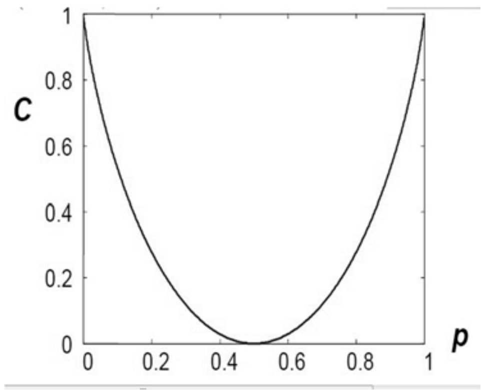


Fig. 1.1. Capacità di canale binario in funzione della probabilità di errore.

Il concetto di "capacità" di canale introduce il problema del costo relazionato all'errore, argomento che vedremo nel prossimo capitolo, e come il rapporto costi-benefici esista anche per il software e come la sua natura immateriale non voglia dire affatto che si possa ottenere un software "error free", in assoluto, e che poi lo si possa ottenere a costi linearmente dipendenti dalla probabilità di errore.

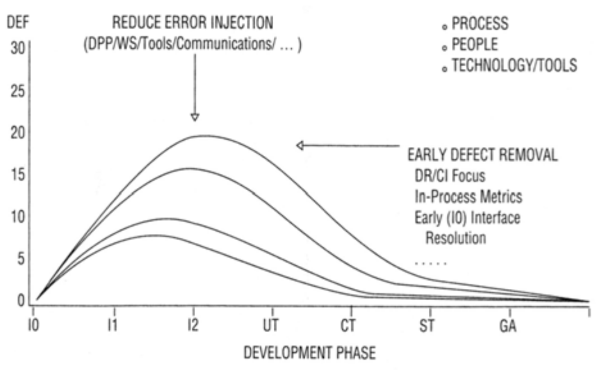


Fig. 1.2. Modello di Reylegh per la qualità del software. La figura è tratta da Kan, 2000.

La figura 1.2 mostra l'andamento, in quattro diversi casi, del numero di difetti nel software durante lo sviluppo e le fasi di testing prima del rilascio.

La figura 1.2 è tratta da Stephen H. Kan (Kan, 2000) e mostra il modello probabilistico di Rayleigh dell'andamento della difettosità nella vita del software. Le quattro curve rappresentano casi diversi e non sviluppi dello stesso caso. Questa funzione non ha relazione diretta con il nostro modello ma è ottimamente indicativa dell'andamento del processo di "diagonalizzazione" nel tempo. Dalla fase iniziale in cui vi è una immissione di errori a quella successiva in cui gli errori vengono eliminati. Dalla fase I2 inizia il processo che abbiamo chiamato "diagonalizzazione" e la cui evoluzione vedremo meglio nel prossimo capitolo. L'eliminazione avviene sempre in seguito ad una verifica tra rapporto ingresso-uscita e quindi è chiaro che il problema di verifica di una soluzione, come vedremo parlando dei problemi NP, è un problema risolvibile mentre può non esserlo il problema di individuazione iniziale della soluzione.

La complessità condizionale

In questo nostro percorso non ci interessiamo della rappresentazione della conoscenza ma del "costo" e dell'errore connessi ai sistemi che possiamo mettere in relazione uno stato in ingresso con un evento in uscita. Lo stato quindi è una collezione di informazioni che non entrano singolarmente nel sistema ma rappresentano collettivamente le variabili di ingresso. Cioè, con un esempio banale, non sono ingressi del sistema "albero", "fiori vermigli", "frutto balausta" per ottenere una relazione molti ad uno con "melograno", il nostro "stato" è proprio "melograno", il rapporto è sempre uno ad uno. La rappresentazione della conoscenza non è il nostro argomento.

Possiamo però impiegare una interessante rappresentazione della conoscenza per il nostro fine. Tale strumento è dato dalle reti bayesiane caratterizzate da grafi connessi privi di cicli i cui nodi sono caratterizzati da una tabella di probabilità connesse relative alle variabili "genitori". In genere in queste reti si usa l'inferenza statistica che è sostanzialmente l'analisi degli stati vero o falso che può assumere una variabile. Tratteremo nel seguito le reti bayesiane. Qui citiamo il teorema di Bayes in quanto può essere utile per il seguito della trattazione.

$$P(P|N) = P(N|P)\ P(P) / P(N) \quad [1.9]$$

L'interesse di questo teorema è nel fatto che mette in relazione probabilità a priori con probabilità a posteriori. Con un semplice esempio che può risultare interessante:

- sappiamo che per lo 0.8 dei giorni è nuvolo, $P(N)=0.8$,
- sappiamo che piove lo 0.5 dei giorni, $P(P) = 0.5$