1

Capitolo C16
grammatiche e linguaggi contestuali

Contenuti delle sezioni

a. linguaggi finitamente presentabili p. 2

5 pagine

C16 0.01 le prime delle pagine che seguono presentano brevemente le modalità che vengono seguite per definire svariati tipi di linguaggi formali, in genere infiniti numerabili, attraverso strumenti individuati da definizioni (necessariamente) finite e tendenzialmente costruttive.

Successivamente si concentra l'attenzione sopra quelli cha sono chiamati *context sensitive languages* e che qui chiamiamo linguaggi contestuali o in breve linguaggi-S.

Si tratta di una collezione di linguaggi che amplia sensibilmente quella dei linguaggi acontestuali, ma come questa si può individuare mediante un ben definito tipo di grammatica generativa e mediante un preciso tipo di riconoscitori, strumenti che estendono, risp., l'insieme delle grammatiche acontestuali e l'insieme dei riconoscitori a pila.

Anche per questi linguaggi si dimostra costruttivamente l'equivalenza dei due tipi di presentazioni mostrando come definire un riconoscitore equivalente-L a partire da una grammatica e viceversa come ottenere una grammatica equivalente-L a un assegnato riconoscitore.

C16 a. linguaggi finitamente presentabili

C16 a.01 Iniziamo con una definizione generica dei linguaggi finitamente presentabili o **linguaggi-FP**, definizione che non si occupa di dettagli formali o operativi ma ha il solo scopo di delineare una prospettiva generale.

Denotiamo con A un alfabeto i cui caratteri chiamiamo caratteri terminali, e consideriamo una coppia $\langle \boldsymbol{S}, \mathcal{M} \rangle$ costituita da uno strumento $\boldsymbol{S}$ in grado di elaborare stringhe su A e da un meccanismo di governo (e supervisione) $\mathcal{M}$ in grado di controllare le singole azioni di $\boldsymbol{S}$ e più complessivamente le elaborazioni che $\boldsymbol{S}$ porta avanti.

In alcuni casi lo strumento $\boldsymbol{S}$ è in grado di eseguire una unica elaborazione che ha la possibilità di procedere illimitatamente e che in taluni passi operativi propone una stringa w come elemento di un linguaggo associato L tipicamente registrandola su un nastro di emissione.

In altri casi ogni elaborazione di $\boldsymbol{S}$ viene innescata da dati iniziali provenienti dall'esterno e che lo stesso $\mathcal{M}$ ha giudicato ammissibili e procede fino a una conclusione auspicata e talora garantita dalla quale si ricava una stringa w da presentare come stringa su A e di conseguenza assegnata ad un linguaggio associato L.

Una tale coppia $\langle \boldsymbol{S}, \mathcal{M} \rangle$ la chiamiamo strumento governato.

Ad $\boldsymbol{S}$ e a $\mathcal{M}$ chiediamo solo genericamente di essere definiti con chiarezza, ossia in modo che le loro azioni siano ampiamente giudicate esaminabili in dettaglio, riproducibili e quindi attendibili.

Il linguaggio L che conveniamo di individuare con una scrittura della forma $L := \boldsymbol{S}^{\mathcal{M}}$, viene detto **linguaggio finitamente presentato** dallo strumento governato $\langle \boldsymbol{S}, \mathcal{M} \rangle$.
Oltre al termine "linguaggio finitamente presentato" qui usiamo anche la sua abbreviazione linguaggio-FP.

C16 a.02 Si possono riconoscere facilmente vari tipi di presentazioni finite di linguaggi trovati in precedenza.

Ogni linguaggio fornito da una espressione ben formata $\mathcal{E}$ nella quale compaiono alcuni operatori definiti sulle stringhe ed estesi nel rispetto dell'operazione unione per potersi applicare a linguaggi definiti in modo progressivo.
Questo in particolare è il caso dei linguaggi razionali individuati dalle espressioni razionali [C12e].

Analoga considerazione si può sviluppare per i linguaggi accettati da una macchina sequenziale $\boldsymbol{R}$ che può essere un riconoscitore-R o un riconoscitore-F e dal meccanismo della accettazione ^{Rcn}.

Altri linguaggi si ottengono sottoponendo i linguaggi di una collezione nota (in genere in quanto presentata finitamente) a una macchina sequenziale di un determinato genere.
In particolare abbiamo visto che sottoponendo i linguaggi razionali a trasduttori-R si ottengono ancora linguaggi razionali [C13b].

Un altro genere di presentazione finita riguarda le collezioni di linguaggi generati dalle grammatiche di un dato tipo; a questo genere appartengono i linguaggi acontestuali forniti dalle grammatiche-F e dal meccanismo della derivazione di stringhe che siano costituite solo da terminali.

Un altro genere di presentazione finita è costituito da sistemi di equazioni con il ruolo degli strumenti e dal meccanismo di risoluzione univoca, come abbiamo visto per i sistemi di equazioni lineari che forniscono i linguaggi razionali [C12f].

Possono considerarsi tali anche i linguaggi generati dalle macchine generatrici definite più permissivamente denotate con MSPG, linguaggi che chiamiamo ricorsivamente enumerabili [B18b]; più in particolare abbiamo i linguaggi ricorsivi [B18e] che devono soddisfare la richiesta di generazione opportunamente controllata.

Un altro tipo di linguaggi-FP si può individuare utilizzando come strumento un computer dotato di un linguaggio di programmazione e come meccanismo di controllo un programma in grado di governare elaborazioni del computer finalizzate a produrre stringhe su un alfabeto A che soddisfano determinati requisiti.

Questo ampliamento di prospettiva, forse eccessivamente olistico, porta a delineare una similitudine (che proponiamo solo come generica suggestione) tra lo strumento governato $\langle \boldsymbol{S}, \mathcal{M} \rangle$ e la coppia costituita dall'hardware e da un software procedurale di un comune computer.

C16 a.03 Dalla genericità della definizione delle coppie $\langle \boldsymbol{S}, \mathcal{M} \rangle$, unita alla varietà degli esempi accennati e immaginabili senza difficoltà, segue la difficoltà, forse la impossibilità, di definire a un buon livello di generalità i confini della collezione dei linguaggi finitamente presentati.

Resta quindi la necessità di definire con precisione ogni specifico tipo di strumenti e di meccanismi per il loro governo finalizzati alla definizione di linguaggi su un dato alfabeto. Queste definizioni sono indispensabili perché si possano dedurre con attendibilità le loro caratteristiche e le loro possibilità applicative.

La sola conclusione che si può trarre a partire dalla sola definizione da prospettiva riguarda il cardinale dei linguaggi finitamente presentati.

Per questo cominciamo con l'osservare che la condivisibilità della definitezza dei linguaggi-FP implica che tutte le coppie $\langle \boldsymbol{S}, \mathcal{M} \rangle$ che sono in grado di individuarli devono potersi esprimere con un complesso di richieste formulato con una chiarezza ampiamente comprensibile e condivisibile.

A ciascuno di questi complessi di richieste quindi, per quanto possa essere elaborato, è lecito chiedere che si possa esprimere con una stringa finita sopra un insieme di segni che denotiamo con $\overline{A}$ che può essere molto esteso, che in ogni momento degli studi deve essere finito e che nel tempo può essere ampliato.

C16 a.04 È opportuno almeno delineare quali contenuti vanno attribuiti a questo alfabeto $\overline{A}$.

Evidentemente questo insieme di segni deve contenere ciascuno degli alfabeti A sui quali sono costruibili le stringhe dei linguaggi da definire, e deve contenere tutti i segni che consentono di precisare gli strumenti per costruirle e i meccanismi per il governo di questi strumenti. tra questi ci devono essere i segni che servono a definire le strutture dei generi macchine sequenziali e sistemi di riscrittura e i segni in grado di documentare le evoluzioni dei riconoscitori, le istruzioni per le MSM e le MSPG, i processi di derivazione per le grammatiche e i processi per le interpretazioni delle collezioni di espressioni delle varie forme.

In $\overline{A}$ devono entrare, in particolare i segni di quelli che chiamiamo alfabeti ausiliari tra i quali segnaliamo un alfabeto delle variabili (o ausiliari) delle grammatiche-F e un alfabeto che consenta di individuare gli stati dei vari tipi di riconoscitori, di trasduttori e di strumenti generatori.

Queste premesse potrebbero indurre a dichiarare che $\overline{A}$ può contenere infiniti segni, ma qui vogliamo invece tenerci lontani da dichiarazioni con concessioni infinitarie.

Per questo assumiamo che $\overline{A}$ sia costituito solo dai segni che sono stati effettivamente utilizzati per definire linguaggi-FP (segni che supponiamo potrebbero essere storicamente accertati).

Questa richiesta a carattere limitativo implica che *Aol* sia un alfabeto finito in quanto tutti i tipi di segni che può comprendere possono provenire solo da un insieme finito di attività di formalizzazioni, ciascuna delle quali contribuisce con un insieme finito di segni.

La definizione di $\overline{A}$, naturalmente, deve essere tenuta aperta a nuove definizioni di coppie $\langle S, \mathcal{M} \rangle$ che ci si aspetta continuino ad essere proposte per presentare finitamente nuovi linguaggi, nonché nuovi tipi di strumenti e di meccanismi di controllo, in relazione a prevedibili nuove esigenze conoscitive e operative.
Quindi assumiamo che $\overline{A}$ sia finito, certamente molto esteso e ampliabile senza limiti, ma solo attraverso definizioni formalmente accurate o condivisibilmente formalizzabili con precione.

C16 a.05 A questo punto è interessante osservare che ogni alfabeto finito ordinato $\langle \alpha_1, \alpha_2, ..., \alpha_N \rangle$ (in questa forma può essere presentato $\overline{A}$) si può codificare con una sequenza di n stringhe su sole due o tre lettere attraverso sostituzioni biiettive come le seguenti:

$$\left\updownarrow \begin{matrix} \alpha_1 & \alpha_2 & \dots & \alpha_n \\ b\,a & b\,a^2 & \dots & b\,a^N \end{matrix} \right\updownarrow \qquad \text{e} \qquad \left\updownarrow \begin{matrix} \alpha_1 & \alpha_2 & \dots & \alpha_N \\ b\,a\,c & b\,a^2c & \dots & b\,a^nc \end{matrix} \right\updownarrow .$$

Questa osservazione consente di rendere quanto segue indipendente dalle estensioni degli alfabeti A sui quali si basano i linguaggi-FP che si possono prendere in considerazione.

Osserviamo anche che le conoscenze maturate fanno ragionevolmente prevedere che ad ogni presentazione finita di un linguaggio-FP si possano affiancare più presentazioni alternative dello stesso linguaggio.

Possiamo dunque sostenere che tutte le presentazioni finite di linguaggi formali, presenti e prevedibili, si possono collocare in un linguaggio su un alfabeto $\overline{A}$ finito (ed espandibile) e quindi possano costituire un insieme di stringhe al più numerabile.

Si possono inoltre prospettare senza difficoltà successioni di presentazioni finite di linguaggi diversi: per questo sono sufficienti e significativi i vari esempi incontrati di linguaggi razionali e di linguaggi acontestuali su precisi alfabeti.

Si può quindi concludere che la collezione di tutti i linguaggi finitamente presentabili sia potenzialmente numerabile.

C16 a.06 Accenniamo ad un ampliamento delle precedenti considerazioni riguardante il complesso delle attività per la individuazione e il controllo dei linguaggi formali.

È lecito pensare che anche l'insieme degli strumenti governati si può sequenzializzare in qualche modo e si può pensare di precisare la sequenza estendibile con il procedere del tempo delle attività computazionali affidabili.

Si può quindi pensare alla sequenzializzazione attraverso una composizione diagonale alla Cantor [B30a07] dei passi di generazione della totalità delle stringhe generabili da strumenti governati.

È dunque pensabile la sequenzializzazione di tutte le strighe generabili da strumenti governati e la proclamazione del fatto che tendono a costituire un insieme numerabile.

Si può pensare anche che queste stringhe possano essere generate munendole di indicazioni sul processo della loro generazione.

Inoltre, allargando la concezione sintattica dei linguaggi ad una loro concezione semantica, è pensabile che ciascuna stringa fornita da una presentazione finita possa essere accompagnata da indicazioni riguardanti una sua attendibile interpretazione.

Ciascuna di queste stringhe risulta quindi munita di un significato e si può considerare un elemento di conoscenza gestibile in modo attendibile.

In tal modo si può pensare che ciascuno di questi elementi di conoscenza, in linea di principio possa essere tenuto sotto controllo e che il loro complesso sia finito ma in continua crescita e inoltre sia potenzialmente controllabile affidabilmente.

Un tale scenario può servire a sostenere una estesa disponibilità delle conoscenze affidabili in relazione al loro utilizzo per procedere ad affrontare e auspicabilmente risolvere anche con sistematicità i problemi che le comunità e gli organismi devono senza tregua affrontare.

C16 a.07 Occorre aggiungere che il precedente scenario non deve essere visto in modo idilliaco per due ordini di motivi.

Le procedure che conducono a generare stringhe possono avere evoluzioni non controllabili e imprevedibili, come conseguenza della indecidibilità del problema dell'arresto della macchina di Turing o di simili risultati globalmente limitativi.

La effettiva interpretazione di una stringa generata in relazione al processo della sua generazione e il suo collegamento a un modo di risolvere una istanza di problema è reso enormemente oneroso dalle codifiche e dalle ridondanze che hanno condotto allo scenario unitario sopra delineato.

Quindi la utilizzazione dei risultati delle elaborazioni attendibili richiede esami specifici che possono essere molto onerosi e possono necessitare di continue verifiche, soprattutto in relazione alla loro adeguatezza nei confronti delle applicazioni che continuano a presentarsi, spesso con caratteristiche mutevoli.

Questi esami specifici dallo scenario unitario possono solo trarre il conforto di far parte di un complesso straordinariament ampio di attività portate avanti con modalità che favoriscono la riutilizzabilità dei risultati ottenuti, anche passando da un settore a un altro che nelle prime indagini è stato ingenuamente considerato lontano e scollegato da quello nel quale sono maturate esperienze positive.

C16 a.08 Per individuare collezioni di linguaggi ben controllabili che possono risultare utili, direttamente o indirettamente, alle applicazioni si devono individuare presentazioni finite precisamente formalizzabili.

Nel seguito di questo capitolo, come si è detto, ci proponiamo di introdurre due tipi di presentazioni di linguaggi equivalenti, l'uno costituito da grammatiche più generali delle acontestuali e uno costituito da riconoscitori con più possibilità operative dei riconoscitori a pila.

Procederemo a dimostrare che per ogni alfabeto A di terminali questi due tipi di presentazioni sono equivalenti-L, ovvero che individuano la stessa collezione di linguaggi, e che sono in grado di presentare, oltre naturalmente a tutti i linguaggi acontestuali, cioè all'intero $\mathbf{LngF}_A$, vari significativi linguaggi che non sono generabili da grammatiche acontestuali e non sono riconoscibili da riconoscitori a pila.

Testo fruibile in https://www.mi.imati.cnr.it/alberto/ e https://arm.mi.imati.cnr.it/Matexp/matexp_main.php