

Approfondimento di AALP

type inference sul linguaggio base + subtyping + NatBool

Simone Ballarin
matricola 1207245
anno accademico 2018/19

Subtyping algoritmico

Vigono le stesse regole del subtyping dichiarativo con eccezione di S-Trans e S-Refl, inoltre devono essere aggiunte le regole NatNat e BoolBool.

Le regole quindi presenti sono:

$\overline{S <: Top}$	S-Top
$\overline{Nat <: Nat}$	NatNat
$\overline{Bool <: Bool}$	BoolBool
$\overline{Nat <: Bool}$	NatBool
$\frac{T_1 <: S_1 \quad S_2 <: T_2}{S_1 \rightarrow S_2 <: T_1 \rightarrow T_2}$	S-Arrow

Lemma 1

ENUNCIATO: $S <: S$ senza S-Refl

DIMOSTRAZIONE:

induzione strutturale su S

(casi base)

$S = Top$, allora la tesi vale per l'assioma S-Top

$S = Nat$, allora la tesi vale per l'assioma $Nat <: Nat$

$S = Bool$, allora la tesi vale per l'assioma $Bool <: Bool$

(casi induttivi)

$S = S_1 \rightarrow S_2$, per ipotesi induttiva $S_1 <: S_1$ e $S_2 <: S_2$, quindi per per S-Arrow, $S <: S$

□

Lemma 2

ENUNCIATO: $S <: T$ derivabile implica $S <: T$ derivabile senza S-Trans

DIMOSTRAZIONE:

Se $S <: T$ derivabile allora $S <: T$ derivabile senza S-Refl.

procedo ora per induzione sull'albero di $S <: T$ senza S-Refl:

(h=1)

In questi casi la regola S-Trans non può essere usata in quanto non è un assioma, quindi alberi di altezza uno non potranno mai contenere S-Trans

(h->h+1)

procedo in base all'ultima regola usata, se questa è diversa da S-Trans allora la tesi segue per ipotesi induttiva sui sottoalberi premessa.

Ultima regola usata S-Trans ottenuta da $S <: U$ e $U <: T$ di altezza al più h, per qualche U.

Procedo per casi in base all'ultima regola dei due sottoalberi premesse di S-Trans (si ricorda che l'albero non contiene S-Refl):

(Any, S-Top)

Se l'albero destro finisce con S-Top allora $T = Top$, quindi $S <: Top$ viene per S-Top.

(S-Top, Any)

Se l'albero sinistro termina con S-Top allora $U = Top$. Per ip. ind. abbiamo che l'albero $U <: T$ è derivabile senza S-Trans, quindi siccome non è presente S-Refl, l'ultima regola usata per $U <: T$ è S-Top (unica regola utilizzabile). Quindi $T = Top$ e $S <: T$ ottenibile con S-Top.

(S-Arrow, S-Arrow)

Allora $S=S1 \rightarrow S2$, $T=T1 \rightarrow T2$, $U=U1 \rightarrow U2$ e $U1 <: S1$, $S2 <: U2$, $T1 <: U1$ e $U2 <: T2$.

Dalle precedenti posso ottenere $T1 <: S1$ e $S2 <: T2$ usando S-Refl per ip. ind. posso dire che $T1 <: S1$ e $S2 <: T2$ si possono ottenere senza T-Trans. Da queste due poi ottengo con S-Arrow $S <: T$ senza transitività.

□

Lemma 3:

ENUNCIATO: $S <: T$ derivabile sse $S <: T$ derivabile alg

DIMOSTRAZIONE:

$[S <: T \rightarrow S <: T \text{ alg}]$

procedo per induzione sull'albero $S <: T$.

($h=1$)

Allora l'albero è ottenuto da un assioma. Nel caso di S-Top allora la derivazione è anche algoritmica, nel caso S-Refl basta usare Lemma 1, l'ultimo caso che rimane è NatBool che rimane anche nel typing algoritmico.

($h \rightarrow h+1$)

Procedo in base all'ultima regola usata.

S-Trans

Se è stata usata S-Trans allora segue per Lemma 2.

S-Arrow

Se è stata usata S-Arrow allora per ogni premessa esiste una derivazione algoritmica per ip. induttiva. Unendo questi due alberi con S-Arrow (presente anche nel typing algoritmico) si ottiene $S <: T$ algebricamente.

$[S <: T \text{ alg} \rightarrow S <: T]$

Le regole algebriche sono un sottoinsieme delle regole dichiarative ad eccezione delle regole NatNat e BoolBool. Questi due assiomi però possono essere sempre sostituiti da S-Refl per questo segue la tesi.

□

Correttezza e completezza rispetto subtyping algoritmico

Lemma 4:

ENUNCIATO: $S <: T$ derivabile algebricamente sse $(\Rightarrow) S \text{ T ha valore } True$.

DIMOSTRAZIONE:

$[\Rightarrow]$

Procedo per induzione sull'altezza:

($h=1$)

Allora l'albero è ottenuto da STop, NatNat, NatBool, BoolBool. Per ognuno di questi casi c'è un pattern matching che ritorna *True*.

($h \rightarrow h+1$)

Allora è ottenuto da S-Arrow. Le due premesse sono alberi algebrici di altezza al più h , quindi per ipotesi induttiva le due chiamate ricorsive fatte nel codice avranno come risultato *True*. Il calcolo evolve a *True* && *True* che quindi evolve a *True*.

$[\Leftarrow]$

Procedo per induzione sulla struttura del giudizio

(casi base)

allora si ricade in uno dei primi quattro casi del pattern matching, ognuno dei quali corrisponde ad un assioma algoritmico (essendo che ora ogni regola è guidata dalla forma del giudizio).

(caso induttivo)

Il quinto (il caso della funzione) si ottiene per ipotesi induttiva, se quest'ultimo caso è *True* vuol dire che le due chiamate ricorsive ritornano *True*. Quindi per ipotesi induttiva esiste un albero algoritmico per le premesse di S-Arrow, quindi la tesi.

□

Lemma 5:

ENUNCIATO: Se $S \leq T$ derivabile algebricamente allora $(\leq) S T = True$ altrimenti

$(\leq) S T = False$

DIMOSTRAZIONE:

La funzione termina sempre in quanto le chiamate vengono fatte sempre su tipi più piccoli. Per il lemma 4 la computazione termina a *True* sse esiste una derivazione algebrica. Se non esiste una derivazione algebrica allora siccome il codice deve terminare allora essendo che il risultato non è *True* è per forza di cose *False*.

L'algoritmo è corretto rispetto alla derivazione algebrica la quale è corretta rispetto alla derivazione dichiarativa.

□

Typing Algoritmico

Vigono le stesse regole del typing dichiarativo con eccezione di T-Sub che deve essere rimossa. Inoltre devono essere modificate le regole TIf e TApp con una loro controparte algoritmica TIFAlg, TAppAlg.

Le regole quindi presenti sono:

1. TSumAlg, uguale a TSum;
2. TMinAlg, uguale a TMin;
3. TNatAlg, uguale a TNat;
4. TVarAlg, uguale a TVar;
5. TTrueAlg, uguale a TTrue;
6. TFalseAlg, uguale a TFalse;
7. TArrowAlg, uguale a TArrow
8. TIfAlg

$$\frac{\Gamma \vdash b : B \quad B \leq Bool \quad \Gamma \vdash m : T \quad \Gamma \vdash n : S}{\Gamma \vdash \text{if } b \text{ then } m \text{ else } n : \text{lub}(T, S)}$$

9. TAppAlg

$$\frac{\Gamma \vdash t_1 : S_{11} \rightarrow S_{12} \quad \Gamma \vdash t_2 : S_2 \quad S_2 \leq S_{11}}{\Gamma \vdash t_2 : S_{12}}$$

Per ragioni di brevità nelle seguenti dimostrazione si darà per scontato che la conversione tra $S \leq T$ e $S < T$ algoritmico è sempre possibile per ogni S e T (in entrambi i versi). Quindi non si useranno sintassi diverse.

Correttezza Typing algoritmico

ENUNCIATO: Se $\Gamma \vdash t : T$ allora $\Gamma \vdash t : T$

DIMOSTRAZIONE:

procedo per induzione sull'altezza di $\Gamma \vdash t : T$, sia questa h .

($h=1$)

Allora l'ultima regola usata è TVarAlg, TNatAlg, TTrueAlg o TFalseAlg. In tutti questi casi la derivazione è già in forma dichiarativa siccome le regole sono uguali alla loro variante non algoritmica.

($h \rightarrow h+1$)

Procedo per casi sull'ultima regola usata.

TArrowAlg

Se l'ultima regola è TArrowAlg allora $\Gamma \vdash \text{fn } x : S.t : T$ ottenuta da $\Gamma, x : S \vdash t : T$ con altezza al più h .

Per ipotesi induttiva abbiamo $\Gamma, x : S \vdash t : T$ quindi usando TArrow si ottiene $\Gamma \vdash \text{fn } x : S.t : T$.

TAppAlg

Se l'ultima regola usata è TAppAlg allora $\Gamma \vdash t_1 t_2 : T_{12}$ ottenuta da $\Gamma \vdash t_1 : T_{11} \rightarrow T_{12}$, $\Gamma \vdash t_2 : T_2$ e $T_2 \leq T_{11}$, tutte ottenute con altezza al più h .

Per ipotesi induttiva sui primi due abbiamo $\Gamma \vdash t_1 : T_{11} \rightarrow T_{12}$, $\Gamma \vdash t_2 : T_2$, ora con TApp e TSub

otteniamo la seguente derivazione che dimostra la tesi:
$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \frac{\Gamma \vdash t_2 : T_2 \quad T_2 \leq T_{11}}{\Gamma \vdash t_2 : T_{11}}}{\Gamma \vdash t_1 t_2 : T_{12}} .$$

TSumAlg

Se l'ultima regola usata è TSumAlg allora $\Gamma \vdash t_1 + t_2 : Nat$, ottenuto da $\Gamma \vdash t_1 : Nat$ e $\Gamma \vdash t_2 : Nat$ con altezza al più h. Per ipotesi induttiva allora abbiamo $\Gamma \vdash t_1 : Nat$ e $\Gamma \vdash t_2 : Nat$, da queste due per TSum ottengo la tesi.

TMinAlg

Se l'ultima regola usata è TMinAlg allora la dimostrazione è analoga al caso TSumAlg.

TIFAlg

Se l'ultima regola usata è TIFAlg allora $\Gamma \vdash \text{if } b \text{ then } m \text{ else } n : L$, ottenuto da $\Gamma \vdash b : B$, $B < Bool$, $\Gamma \vdash m : T$, $\Gamma \vdash n : S$ con altezza al più h, dove $L = lub(T, S)$.

Per ipotesi induttiva abbiamo $\Gamma \vdash b : B$, $\Gamma \vdash m : T$, $\Gamma \vdash n : S$ e per definizione di lub abbiamo $T < L$ e $S < L$. Ora usando T-If e due volte T-Subs si può dimostrare la tesi nel seguente

$$\text{modo } \frac{\frac{B < Bool}{\Gamma \vdash b : B} \quad \frac{\Gamma \vdash m : T \quad T < L}{\Gamma \vdash m : L} \quad \frac{\Gamma \vdash n : S \quad S < L}{\Gamma \vdash n : L}}{\Gamma \vdash \text{if } b \text{ then } m \text{ else } n : L}.$$

Conclusione: il caso base e i casi induttivi sono dimostrati quindi per ipotesi induttiva si ha la tesi.

□

Completezza typing algoritmico (tipo minimo S)

ENUNCIATO: Se $\Gamma \vdash t : T$ allora $\Gamma \vdash t : S$ per qualche $S < T$

DIMOSTRAZIONE:

Si procede per induzione sull'altezza di $\Gamma \vdash t : T$.

(h=1)

Se l'ultima regola usata è TVar, TTrue, TFalse o TNat allora siccome queste regole sono identiche alla loro variante algoritmica l'albero è già in forma algoritmica.

(h->h+1)

Procedo per casi in base all'ultima regola usata.

TArrow

Se l'ultima regola usata è TArrow allora $\Gamma \vdash \lambda x : T_1. t : T_1 \rightarrow T_2$ ottenuto da $\Gamma, x : T_1 \vdash t : T_2$ con altezza al più h. Per ipotesi induttiva $\Gamma, x : T_1 \vdash t : S_2$ per qualche $S_2 < T_2$, ora da questa segue per TArrowAlg $\frac{\Gamma, x : T_1 \vdash t : S_2}{\Gamma \vdash \lambda x : T_1. t : S_2}$, siccome $S_2 < T_2$ per SFun abbiamo $T_1 \rightarrow S_2 < T_1 \rightarrow T_2$ che dimostra la tesi.

TApp

Se l'ultima regola usata è TApp allora $\Gamma \vdash t_1 t_2 : T_{12}$ ottenuto da $\Gamma \vdash t_1 : T_{11} \rightarrow T_{12}$ e $\Gamma \vdash t_2 : T_{11}$ con altezza h o inferiore. Per ipotesi induttiva allora $\Gamma \vdash t_1 : S_1$ e $\Gamma \vdash t_2 : S_2$ con $S_1 < T_{11} \rightarrow T_{12}$ e $S_2 < T_{11}$. Per inversione abbiamo $S_1 = S_{11} \rightarrow S_{12}$ con $T_{11} < S_{11}$ e $S_{12} < T_{12}$. Per transitività abbiamo $S_2 < S_{11}$.

Ora per TAppAlg abbiamo $\frac{\Gamma \vdash t_1 : S_{11} \rightarrow S_{12} \quad \Gamma \vdash t_2 : S_2 \quad S_2 < S_{11}}{\Gamma \vdash t_1 t_2 : S_{12}}$ che dimostra la tesi siccome $S_{12} < T_{12}$.

TSum

Se l'ultima regola usata è TSum allora $\Gamma \vdash t_1 + t_2 : Nat$ ottenuto da $\Gamma \vdash t_1 : Nat$ e $\Gamma \vdash t_2 : Nat$ entrambi con altezza al più h. Per ipotesi induttiva $\Gamma \vdash t_1 : L$ e $\Gamma \vdash t_2 : S$ con $L < Nat$ e $S < Nat$, per inversione abbiamo $L = Nat$ e $T = Nat$. Dalle due derivazioni algoritmiche si ottiene per TSumAlg $\Gamma \vdash t_1 + t_2 : Nat$, siccome $Nat < Nat$ per la regola algoritmica NatNat la tesi è dimostrata.

TMin

Se l'ultima regola usata è TMin allora la dimostrazione è analoga a TSum.

TIF

Se l'ultima regola usata è TIF allora $\Gamma \vdash \text{if } b \text{ then } m \text{ else } n : T$ ottenuto da $\Gamma \vdash b : Bool$, $\Gamma \vdash m : Nat$ e $\Gamma \vdash n : Nat$, tutte di altezza al più h . Per ipotesi induttiva abbiamo $\Gamma \Rightarrow b : A$, $\Gamma \Rightarrow m : B$ e $\Gamma \Rightarrow n : C$ con $A <: Bool$, $B <: T$ e $C <: T$, per inversione su questi abbiamo $A = Bool$ o $A = Nat$.

Procediamo per casi in base al valore di A :

1. se $A = Nat$ per NatBool e TIfAlg abbiamo $\Gamma \Rightarrow \text{if } b \text{ then } m \text{ else } n : lub(B, C)$;
2. se $A = Bool$ per BoolBool e TIfAlg abbiamo $\Gamma \Rightarrow \text{if } b \text{ then } m \text{ else } n : lub(B, C)$.

Siccome T è maggiorante di B e C allora il lub esiste. Per definizione di lub possiamo dire $lub(B, C) <: T$. Questo dimostra la tesi.

TSub

Se l'ultima regola usata è TSub allora $\Gamma \vdash t : T$ ottenuta da $\Gamma \vdash t : U$ e $U <: T$ entrambe con altezza al più h . Per ipotesi induttiva $\Gamma \Rightarrow t : U'$ con $U' <: U$, per transitività abbiamo $U' <: T$ queste due cose dimostrano la tesi.

Conclusione: il caso base e i casi induttivi sono dimostrati quindi per ipotesi induttiva si ha la tesi.

□

Correttezza algoritmo di typing rispetto a typing algoritmico

ENUNCIATO: se $\text{typeOf } \Gamma t = Just T$ allora $\Gamma \Rightarrow t : T$

DIMOSTRAZIONE:

Procedo per induzione strutturale sul termine t .

(casi base)

$\text{typeOf } \Gamma (EVar x) = Just T$

In questo caso si deve avere $\text{lookup } x \Gamma = Just T$, questo vuol dire che $(x : T) \in \Gamma$, sapendo questo tramite la regola TVarAlg possiamo ottenere la tesi.

$\text{typeOf } \Gamma (ENum n) = Just Nat$

il costruttore *Enum* rappresenta un valore naturale del linguaggio, da questo per TNatAlg deriva la tesi.

$\text{typeOf } \Gamma (EBool b) = Just Bool$

il costruttore *EBool* rappresenta un valore booleano del linguaggio, quindi $b = True$ o $b = False$ da questo per TTrueAlg o per TFalseAlg deriva la tesi.

(casi induttivi)

$\text{typeOf } \Gamma (Efn x S term) = Just S \rightarrow T$

nel codice viene fatta la seguente chiamata $\text{typeOf } ((x, S) : \Gamma) term = Just T$, per ipotesi induttiva allora $\Gamma, (x : S) \Rightarrow term : T$, da questa per TArrowAlg deriva la tesi.

$\text{typeOf } \Gamma (EAp m1 m2) = Just T_{12}$

nel codice vengono fatte le seguenti chiamate :

1. $\text{typeOf } \Gamma m1 = Just T_1$
2. $\text{typeOf } \Gamma m2 = Just T_2$

per ipotesi induttiva quindi si ha $\Gamma \vdash m1 : T_1$ e $\Gamma \vdash m2 : T_2$. Siccome sappiamo che il risultato finale è $Just\ T_{12}$, possiamo dire che la guardia booleana dell'if è vera quindi $T_2 \leq T_{11}$ e, il pattern matching sulla forma di T_1 è ricaduto nel secondo caso $T_1 = T_{11} \rightarrow T_{12}$, da quando appena detto applicando TAppAlg deriva la tesi.

$typeOf\ \Gamma\ (ECond\ b\ m1\ m2) = Just\ lub(T_1, T_2)$

le tre chiamate ricorsive presenti nel codice mi permettono di avere per ipotesi induttiva i seguenti giudizi algoritmici $\Gamma \vdash m1 : T_1$, $\Gamma \vdash m2 : T_2$ e $\Gamma \vdash b : B$. Inoltre siccome il risultato finale sappiamo essere $Just\ lub(T_1, T_2)$ possiamo dire che la guardia dell'if è vera quindi $B \leq Bool$, il codice prosegue nel *case of* intermedio, e quindi $lub(T_1, T_2)$ esiste.

Ora usando TIfAlg possiamo deriva la tesi.

$typeOf\ \Gamma\ (ESum\ m1\ m2) = Just\ Nat$

le due chiamate ricorsive presenti nel codice mi permettono di avere per ipotesi induttiva i seguenti giudizi algoritmici $\Gamma \vdash m1 : T_1$, $\Gamma \vdash m2 : T_2$, essendo che il risultato finale è $Just\ Nat$, posso affermare che il controllo $T_1 = Nat \ \&\&\ T_2 = Nat$ da esito positivo. Applicando TSumAlg ai due giudizi ottengo la tesi.

$typeOf\ \Gamma\ (EMin\ m1\ m2) = Just\ Nat$

analogo a $typeOf\ \Gamma\ (ESum\ m1\ m2) = Just\ Nat$.

□

Completezza algoritmo di typing rispetto a typing algoritmico

ENUNCIATO: se $\Gamma \vdash t : T$ allora $typeOf\ \Gamma\ t = Just\ T$

DIMOSTRAZIONE:

Procedo per induzione sull'altezza della derivazione $\Gamma \vdash t : T$, sia questa h.

(h=1)

Procedo per casi in base alla regola usata.

TVarAlg

allora $\Gamma \vdash t : T$ e $(t : T) \in \Gamma$, il codice esegue il lookup di t dalla struttura che rappresenta il contesto. Siccome la variabile è nel contesto allora il lookup avrà successo e l'algoritmo ritorna il tipo letto dal contesto.

TNatAlg

$\Gamma \vdash n : Nat$, l'algoritmo grazie al pattern matching riconosce il costruttore *ENum* e ritorna *Just Nat*

TBoolAlg

$\Gamma \vdash b : Bool$, l'algoritmo grazie al pattern matching riconosce il costruttore *EBool* e ritorna *Just Bool*

(h->h+1)

TFunAlg

allora $\Gamma \vdash fn\ x : S.t : T$ ottenuta da $\Gamma, x : S \vdash fn.t : T$ con altezza al più h. Per ipotesi induttiva la chiamata $typeOf\ ((x : S) : \Gamma)\ t = Just\ T$, siccome la chiamata da un risultato allora siamo nel secondo caso del *case of* e infatti viene ritornato *Just (TArrow S T)*.

TAppAlg

allora $\Gamma \vdash t_1 t_2 : T_{12}$ ottenuta da $\Gamma \vdash t_1 : T_{11} \rightarrow T_{12}$, $\Gamma \vdash t_2 : T_2$ e $T_2 \leq T_{11}$ con altezza al più h , per ipotesi induttiva abbiamo $\text{typeOf } \Gamma t_1 = \text{Just } (T_{\text{Arrow}} T_{11} T_{12})$ e $\text{typeOf } \Gamma t_2 = \text{Just } T_2$ inoltre per la correttezza dell'algoritmo di subtyping abbiamo $T_2 \leq T_{11} = \text{True}$. Dato quanto appena detto il codice esegue il secondo caso del primo *case of*, e poi il secondo di quello più annidato. Nel *case of* più annidato la guardia avrà esito positivo, quindi verrà ritornato $\text{Just } T_{12}$. Questo dimostra la tesi.

TIFAlg

allora $\Gamma \vdash \text{if } b \text{ then } m \text{ else } n : L$, ottenuto da $\Gamma \vdash b : \text{Bool}$, $\Gamma \vdash m : T$, $\Gamma \vdash n : S$ con altezza al più h , dove $L = \text{lub}(T, S)$. Per ipotesi induttiva $\text{typeOf } \Gamma b = \text{Just Bool}$, $\text{typeOf } \Gamma m = \text{Just } T$, $\text{typeOf } \Gamma n = \text{Just } S$, per quanto appena detto la computazione procede fino al secondo caso del *case of*, dentro questo viene calcolato il lub siccome questo esiste ed è L allora viene trovato e ritornato $\text{Just } L$.

TSumAlg

allora $\Gamma \vdash t_1 + t_2 : \text{Nat}$, ottenuto da $\Gamma \vdash t_1 : \text{Nat}$ e $\Gamma \vdash t_2 : \text{Nat}$. Per ipotesi induttiva $\text{typeOf } \Gamma t_1 = \text{Just Nat}$ e $\text{typeOf } \Gamma t_2 = \text{Just Nat}$. Per quanto appena detto siamo nel secondo caso del *case of* e quindi ritorniamo Just Nat .

TMinAlg

analogo.

□

ENUNCIATO : $\Gamma \vdash t : T$ derivabile allora $\text{typeOf } \Gamma t = \text{Just } T$ altrimenti $\text{typeOf } \Gamma t = \text{Nothing}$

DIMOSTRAZIONE:

Siccome l'algoritmo effettua sempre chiamate ricorsive con termini più piccoli allora termina certamente.

Da completezza e correttezza abbiamo $\Gamma \vdash t : T$ sse $\text{typeOf } \Gamma t = \text{Just } T$. Siccome $\forall T. \Gamma \vdash t : T$ non derivabile, allora $\forall T \text{ typeOf } \Gamma t \neq \text{Just } T$. Dato che l'algoritmo termina certamente allora non potendo essere un valore Just deve essere $\text{typeOf } \Gamma t = \text{Nothing}$.

□