Semantics of Advanced Data Types

Patricia Johann Appalachian State University

June 17, 2021

Course Outline

Lecture 1: Syntax and semantics of ADTs and nested types

Lecture 2: Syntax and semantics of GADTs ✓

Lecture 3: Parametricity for ADTs and nested types

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

- In a parametric model, each type T[A] with a free type variable A has a set interpretation $T_0: Set \rightarrow Set$ and a relational interpretation $T_1: Rel \rightarrow Rel$ such that if R: Rel(A, B) then $T_1R: Rel(T_0A, T_0B)$ (and IEL holds).
- We argued last time that there are parametric models for calculi supporting ADTs and nested types.
- Can we also construct parametric models for calculi supporting GADTs?
- For this we need set and relational interpretations for GADTs as described above.
- Such interpretations exist for discrete semantics of GADTs.
- These satisfy the IEL (and AT).
- For fully functorial semantics of GADTs the situation is more complicated.

Set and Rel Interpretations of GADTs

• If

$\begin{array}{l} \mathsf{data}\ \mathsf{G}\ :\mathsf{Set}\ \rightarrow\ \mathsf{Set}\ \mathsf{where}\\ \mathsf{c}:\forall\{\mathsf{A}:\mathsf{Set}\}\rightarrow\mathsf{F}\ \mathsf{A}\rightarrow\mathsf{G}\ (\mathsf{K}\ \mathsf{A}) \end{array}$

then

- The set interpretation G_0 of G is μH_0 , where $H_0 \, J = Lan_{K_0} \, F_0$, i.e.,

 $G_0 \cong \mu J. Lan_{K_0} F_0$

This left Kan extension is in Set, and K_0 interprets K and F_0 interprets F in Set.

- The relational interpretation G_1 of G is μH_1 , where $H_1 J = Lan_{K_1} F_1$, i.e.,

 $G_1 \cong \mu J. Lan_{K_1} F_1$

This left Kan extension is in Rel, and K_1 interprets K and F_1 interprets F in Rel.

Set and Rel Interpretations of GADTs

• If

data G
$$: Set \rightarrow Set$$
 where
 $c : \forall \{A : Set\} \rightarrow FA \rightarrow G(KA)$

then

- The set interpretation G_0 of G is μH_0 , where $H_0 J = Lan_{K_0} F_0$, i.e.,

 $G_0 \cong \mu J. Lan_{K_0} F_0$

This left Kan extension is in Set, and K_0 interprets K and F_0 interprets F in Set.

- The relational interpretation G_1 of G is μH_1 , where $H_1 J = Lan_{K_1} F_1$, i.e.,

 $G_1 \cong \mu J. Lan_{K_1} F_1$

This left Kan extension is in Rel, and K_1 interprets K and F_1 interprets F in Rel.

Set and Rel Interpretations of GADTs

• If

data G : Set
$$\rightarrow$$
 Set where
c : \forall {A : Set} \rightarrow F A \rightarrow G (K A)

then

- The set interpretation G_0 of G is μH_0 , where $H_0 J = Lan_{K_0} F_0$, i.e.,

$$G_0 \cong \mu J. Lan_{K_0} F_0$$

This left Kan extension is in Set, and K_0 interprets K and F_0 interprets F in Set.

- The relational interpretation G_1 of G is μH_1 , where $H_1 J = Lan_{K_1} F_1$, i.e.,

$$G_1 \cong \mu J. Lan_{K_1} F_1$$

This left Kan extension is in Rel, and K_1 interprets K and F_1 interprets F in Rel.

A Problematic GADT

- Consider the GADT G whose single constructor c provides only an element for the instance G \top

data $\mathsf{G}:\mathsf{Set}\to\mathsf{Set}$ where $\mathsf{c}:\mathsf{G}\top$

- For this GADT, FU = T and KU = T for all U : T.
- The picture we should have in mind is:



A Problematic GADT

- Consider the GADT G whose single constructor c provides only an element for the instance G \top

data $\mathsf{G}:\mathsf{Set}\to\mathsf{Set}$ where $\mathsf{c}:\mathsf{G}\top$

• For this GADT, $FU = \top$ and $KU = \top$ for all $U : \top$.

The picture we should have in mind is:



A Problematic GADT

- Consider the GADT G whose single constructor c provides only an element for the instance G \top

```
data \mathsf{G}:\mathsf{Set}\to\mathsf{Set} where \mathsf{c}:\mathsf{G}\top
```

- For this GADT, $FU = \top$ and $KU = \top$ for all $U : \top$.
- The picture we should have in mind is:



- The set interpretation G_0 of G is $G_0A = (Lan_{\lambda u.1} \lambda u.1) A$.
- This is computed as

$$\Big(\bigcup_{U:Set^0, f:(\lambda u.1)} (\lambda u.1) U \Big) / \sim = \Big(\bigcup_{U:Set^0, f:1 \to A} 1 \Big) / \sim$$

• Here, U is the unique object of Set^0 , * is the unique element of the singleton set 1, and \sim is the smallest equivalence relation such that (U, f, *) and (U, f', *) are related if



commutes, i.e., if f = f'.

• Since the relation generating \sim is already an equivalence relation,

$$(U,f,*)\sim (U,f',*)$$
 iff $f=f'$

• So, for the fully functorial semantics,

$$G_0A = (Lan_{\lambda u,1} \lambda u, 1) A = \{f : 1 \to A\} = A$$

- The set interpretation G_0 of G is $G_0A = (Lan_{\lambda u.1} \lambda u.1) A$.
- This is computed as

• Here, U is the unique object of Set^0 , * is the unique element of the singleton set 1, and \sim is the smallest equivalence relation such that (U, f, *) and (U, f', *) are related if



commutes, i.e., if f = f'.

• Since the relation generating \sim is already an equivalence relation,

$$(U, f, *) \sim (U, f', *)$$
 iff $f = f'$

• So, for the fully functorial semantics,

$$G_0A = (Lan_{\lambda u,1} \lambda u, 1) A = \{f : 1 \to A\} = A$$

- The set interpretation G_0 of G is $G_0A = (Lan_{\lambda u.1} \lambda u.1) A$.
- This is computed as

$$\left(\bigcup_{U:Set^{0}, f:(\lambda u.1) \ U \to A} (\lambda u.1) \ U \ \right) / \sim = \left(\bigcup_{U:Set^{0}, f:1 \to A} 1 \right) / \sim$$

• Here, U is the unique object of $Set^0,$ * is the unique element of the singleton set 1, and \sim is the smallest equivalence relation such that (U,f,*) and (U,f',*) are related if



commutes, i.e., if f = f'.

• Since the relation generating \sim is already an equivalence relation,

$$(U,f,*)\sim (U,f',*)$$
 iff $f=f'$

So, for the fully functorial semantics,

$$G_0A = (Lan_{\lambda u,1} \lambda u, 1) A = \{f : 1 \to A\} = A$$

- The set interpretation G_0 of G is $G_0A = (Lan_{\lambda u.1} \lambda u.1) A$.
- This is computed as

$$\left(\bigcup_{U:Set^{0}, f:(\lambda u.1) \ U \to A} (\lambda u.1) \ U \ \right) / \sim = \left(\bigcup_{U:Set^{0}, f:1 \to A} 1 \right) / \sim$$

• Here, U is the unique object of $Set^0,$ * is the unique element of the singleton set 1, and \sim is the smallest equivalence relation such that (U,f,*) and (U,f',*) are related if



commutes, i.e., if f = f'.

• Since the relation generating \sim is already an equivalence relation,

$$(U, f, *) \sim (U, f', *)$$
 iff $f = f'$

So, for the fully functorial semantics,

$$G_0A = (Lan_{\lambda u.1} \lambda u.1) A = \{f : 1 \to A\} = A$$

- The set interpretation G_0 of G is $G_0A = (Lan_{\lambda u.1} \lambda u.1) A$.
- This is computed as

$$\left(\bigcup_{U:Set^{0}, f:(\lambda u.1)} \underbrace{(\lambda u.1)U}_{U \to A}\right) / \sim = \left(\bigcup_{U:Set^{0}, f:1 \to A} 1\right) / \sim$$

• Here, U is the unique object of $Set^0,$ * is the unique element of the singleton set 1, and \sim is the smallest equivalence relation such that (U,f,*) and (U,f',*) are related if



commutes, i.e., if f = f'.

• Since the relation generating \sim is already an equivalence relation,

$$(U, f, *) \sim (U, f', *)$$
 iff $f = f'$

So, for the fully functorial semantics,

$$G_0A = (Lan_{\lambda u.1} \lambda u.1) A = \{f : 1 \to A\} = A$$

- The relational interpretation G_1 of GR is $G_1R = (Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$.
- We need that if R : Rel(A, B) then

- Consider the relation $R = (1, 2, 1 \times 2)$, where 1×2 relates the single element * of 1 to both elements of 2.
- For the IEL to hold for GADTs we expect the domain of $(Lan_{\lambda u, Eq_1}\,\lambda u. Eq_1)\,R$ to be 1, but it is not!

- The relational interpretation G_1 of GR is $G_1R = (Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$.
- We need that if R : Rel(A, B) then

- Consider the relation $R = (1, 2, 1 \times 2)$, where 1×2 relates the single element * of 1 to both elements of 2.
- For the IEL to hold for GADTs we expect the domain of $(Lan_{\lambda u, Eq_1} \lambda u. Eq_1) R$ to be 1, but it is not!

- The relational interpretation G_1 of GR is $G_1R = (Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$.
- We need that if R : Rel(A, B) then

- Consider the relation $R = (1, 2, 1 \times 2)$, where 1×2 relates the single element * of 1 to both elements of 2.
- For the IEL to hold for GADTs we expect the domain of $(Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$ to be 1, but it is not!

- The relational interpretation G_1 of GR is $G_1R = (Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$.
- We need that if R : Rel(A, B) then

- Consider the relation $R = (1, 2, 1 \times 2)$, where 1×2 relates the single element * of 1 to both elements of 2.
- For the IEL to hold for GADTs we expect the domain of $(Lan_{\lambda u. Eq_1} \lambda u. Eq_1) R$ to be 1, but it is not!

IEL Does Not Hold for GADTs

• We can compute the domain of $(Lan_{\lambda u. Eq_1} \lambda u. Eq_1) R$ as

$$\begin{aligned} \pi_1((Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R) &= & \pi_1(\varinjlim_{U:Rel^0,m:(\lambda u.Eq_1) U \to R}(\lambda u.Eq_1) R) \\ &= & \pi_1(\varinjlim_{U:Rel^0,m:Eq_1 \to R} Eq_1) \\ &= & \varinjlim_{U:Rel^0,m:Eq_1 \to R} (\pi_1 Eq_1) \\ &= & \varinjlim_{U:Rel^0,m:Eq_1 \to R} 1 \\ &= & \bigcup_{U:Rel^0,m:Eq_1 \to R} 1 / \approx \end{aligned}$$

• Here, U is the unique object of Rel^0 , * is the unique element of the singleton set 1, and \approx is the smallest equivalence relation such that (U, m, *) and (U, m', *) are related if



commutes, so $(U, m, *) \approx (U, m', *)$ iff m = m'.

So for the fully functorial semantics,

 $\pi_1(G_1R) = (Lan_{\lambda u.Eq_1} \lambda u.1) R = \{m : Eq_1 \to R\} = \{(!,k_0), (!,k_1)\}$

where k_0 and k_1 send * to the two different elements of 2.

IEL Does Not Hold for GADTs

• We can compute the domain of $(Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$ as

$$\begin{aligned} \pi_1((Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R) &= & \pi_1(\varinjlim_{U:Rel^0,m:(\lambda u.Eq_1) U \to R}(\lambda u.Eq_1) R) \\ &= & \pi_1(\varinjlim_{U:Rel^0,m:Eq_1 \to R} Eq_1) \\ &= & \varinjlim_{U:Rel^0,m:Eq_1 \to R} (\pi_1 Eq_1) \\ &= & \varinjlim_{U:Rel^0,m:Eq_1 \to R} 1 \\ &= & \bigcup_{U:Rel^0,m:Eq_1 \to R} 1 / \approx \end{aligned}$$

• Here, U is the unique object of $Rel^0,$ * is the unique element of the singleton set 1, and \approx is the smallest equivalence relation such that (U,m,*) and (U,m',*) are related if



commutes, so $(U, m, *) \approx (U, m', *)$ iff m = m'.

So for the fully functorial semantics,

 $\pi_1(G_1R) = (Lan_{\lambda u, Eq_1} \lambda u.1) R = \{m : Eq_1 \to R\} = \{(!, k_0), (!, k_1)\}$ here k_0 and k_1 send * to the two different elements of 2.

IEL Does Not Hold for GADTs

• We can compute the domain of $(Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R$ as

$$\pi_1((Lan_{\lambda u.Eq_1} \lambda u.Eq_1) R) = \pi_1(\underset{U:Rel^0,m:(\lambda u.Eq_1) U \to R}{\operatorname{Iim} U:Rel^0,m:(\lambda u.Eq_1) U \to R}(\lambda u.Eq_1) R)$$

$$= \pi_1(\underset{U:Rel^0,m:Eq_1 \to R}{\operatorname{Iim} U:Rel^0,m:Eq_1 \to R} I_q)$$

$$= \underset{U:Rel^0,m:Eq_1 \to R}{\operatorname{Iim} U:Rel^0,m:Eq_1 \to R} 1 \approx$$

• Here, U is the unique object of $Rel^0,$ * is the unique element of the singleton set 1, and \approx is the smallest equivalence relation such that (U,m,*) and (U,m',*) are related if



commutes, so $(U, m, *) \approx (U, m', *)$ iff m = m'.

So for the fully functorial semantics,

 $\pi_1(G_1R) \;=\; (Lan_{\lambda u.\, Eq_1}\,\lambda u.1)\,R \;=\; \{m: Eq_1 \to R\} \;=\; \{(!,k_0),(!,k_1)\}$

where k_0 and k_1 send * to the two different elements of 2.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

- Since the set and relational interpretations of GADTs are not appropriately fibred, the IEL cannot possibly hold.
- We conclude that fully functorial GADTs do not admit parametric models if GADTs are given a traditional relational semantics!
- We get the naturality consequences of parametricity functorial GADTs just from the functorial semantics.
- But if we can prove inhabitation results, or prove representation independence, or derive short cut fusion or deep induction rules for them, then it has to be from something other than parametricity.
- By contrast, discrete (syntax-only) GADTs have parametric models (but not non-trivial functoriality).
- So we can prove inhabitation results and prove representation independence and derive short cut fusion and deep induction rules for them. But we cannot derive (non-trivial) naturality consequences of parametricity for them.
- So there's a trade-off, and the choice we make has consequences.

Course Outline

Lecture 1: Syntax and semantics of ADTs and nested types

Lecture 2: Syntax and semantics of GADTs ✓

Lecture 3: Parametricity for ADTs and (truly) nested types