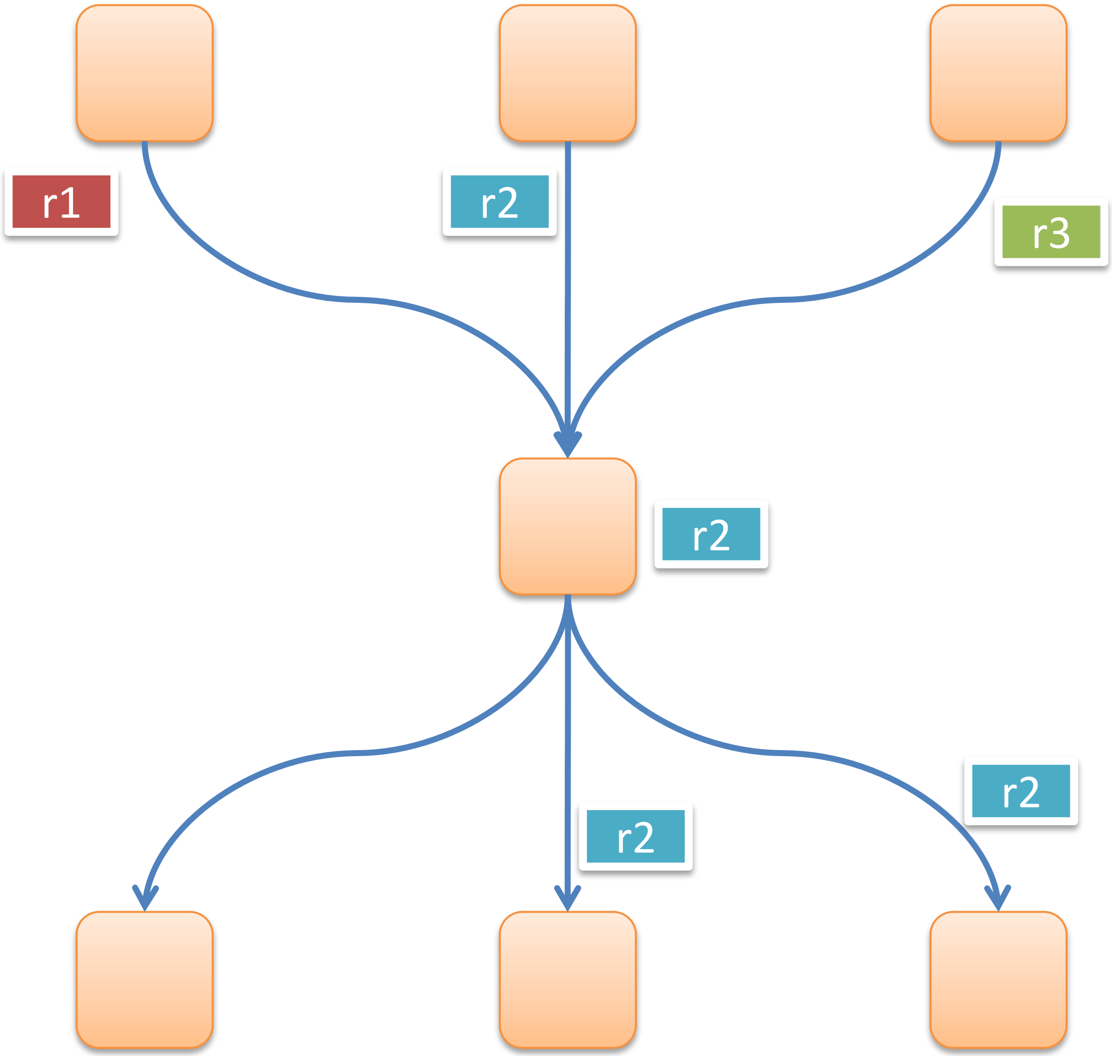
Neighbor-specific BGP: An algebraic exploration

Alexander J. T. Gurney
Timothy G. Griffin

University of Cambridge

ICNP 2010 – Kyoto

Interdomain routing with BGP

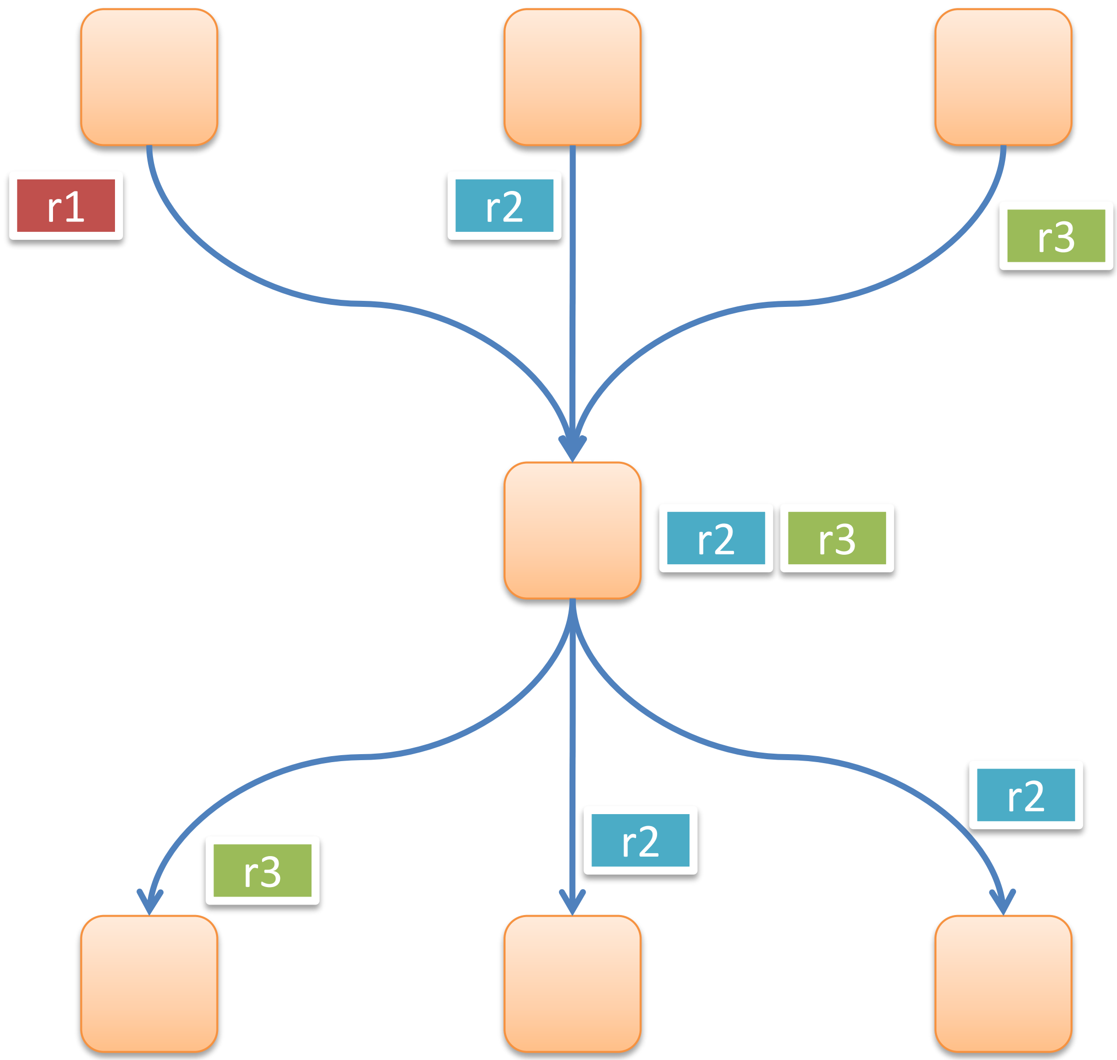


Routes come in

Choose the best one

Tell some neighbors

Neighbor-specific BGP



Routes come in

Choose several good routes

Tell different routes to neighbors

A good idea?

Yes

- Exploit route diversity
- Customers might like tailored services
- Can improve stability

No

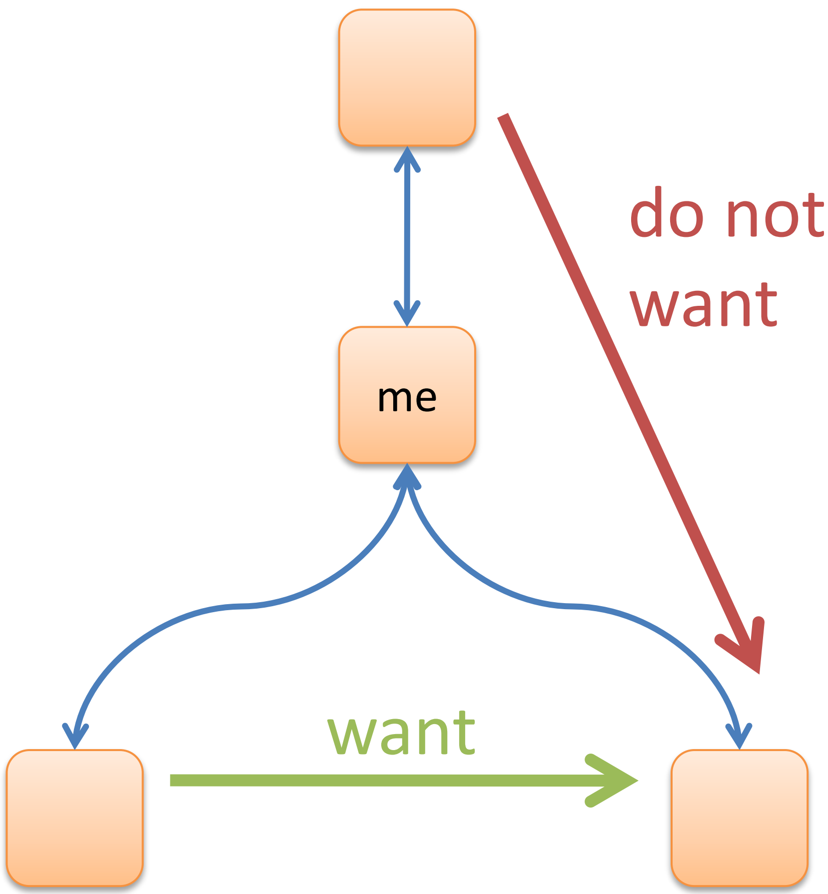
- More complex
- Need new forwarding infrastructure
- More demands on routers and routing tables

Our question: What can we say about *correctness* of neighbor-specific routing protocols?

Correctness

- Correctness arguments for BGP are founded on *local properties* that imply *global stability*.
  - Because all actions taken are with local knowledge
- The main local property is whether *route choice* conforms to some criterion
- This will guarantee convergence to a locally optimal routing solution

Example: Partial customer



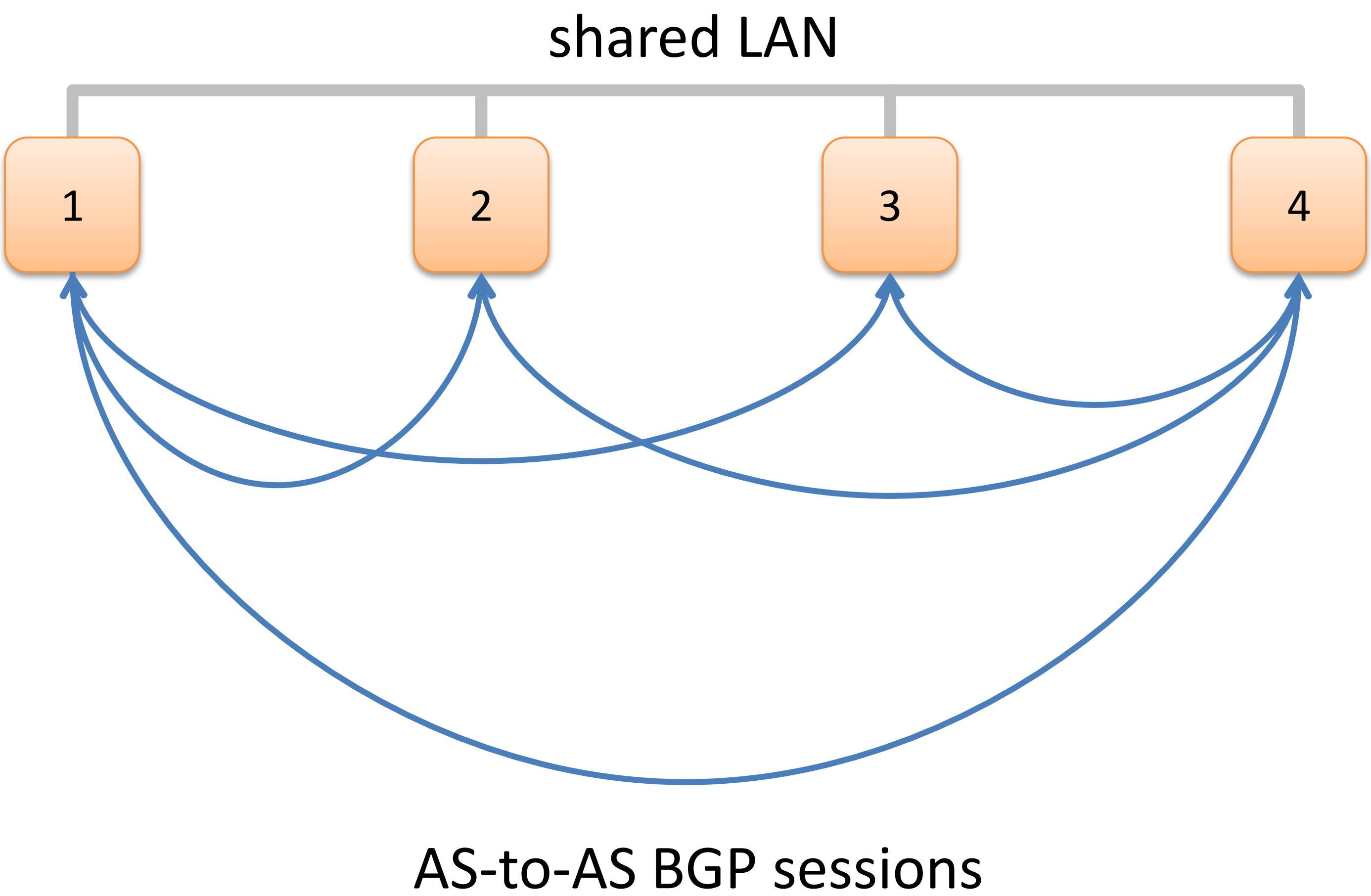
Partial customer

Ordinary customers want to learn my best routes to every destination.

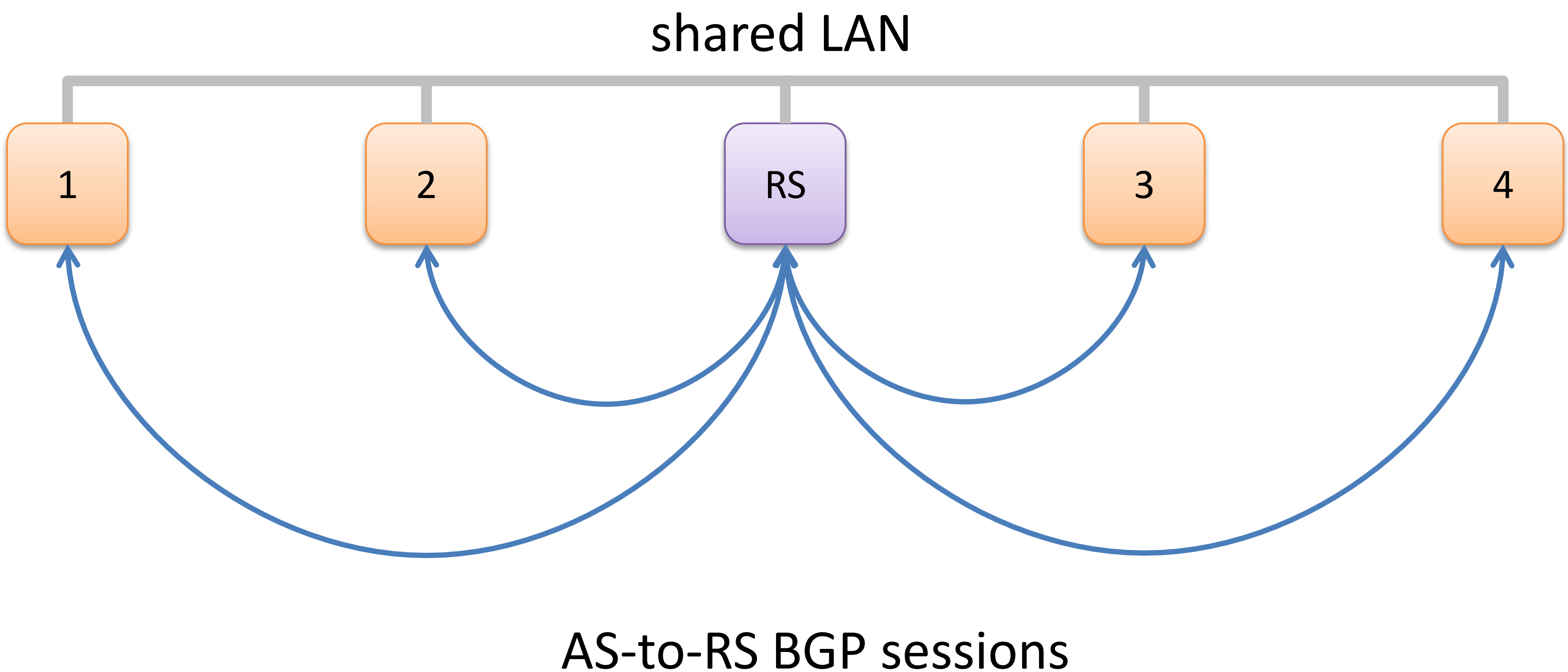
Partial customers only care about *some* of my connections (perhaps those in a particular city or region).

They would like to know the best routes I have that go via the neighbors they care about.

Example: Internet exchange points

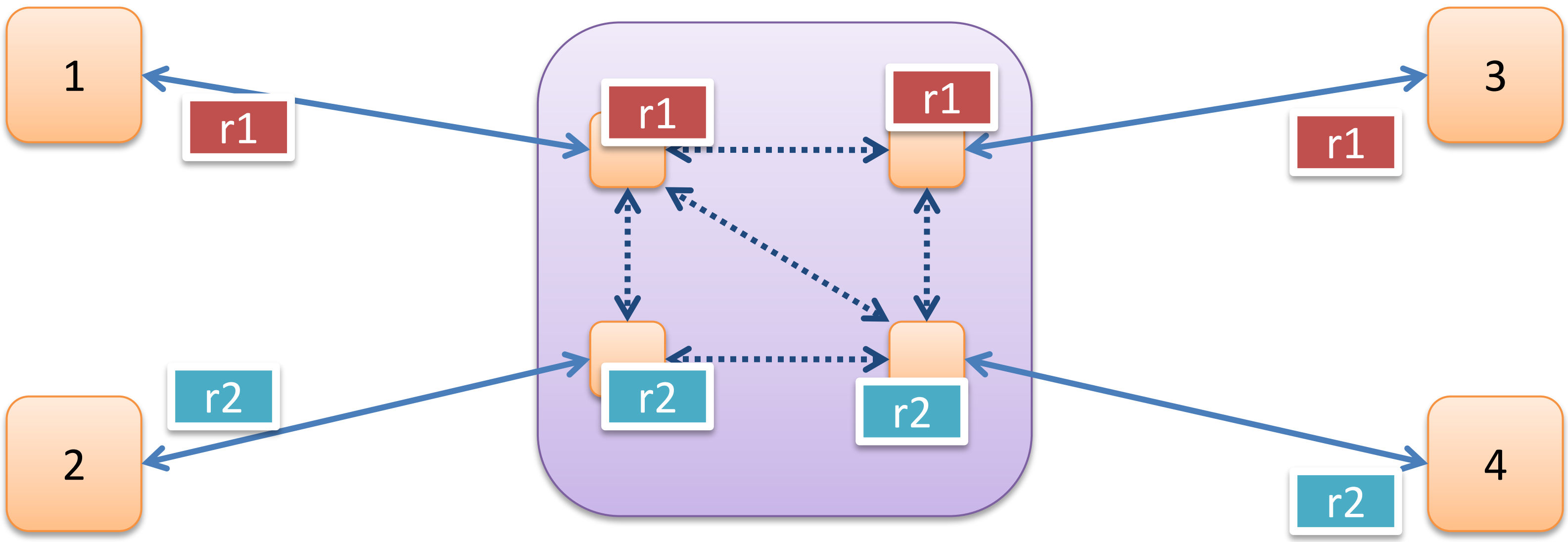


Route server model



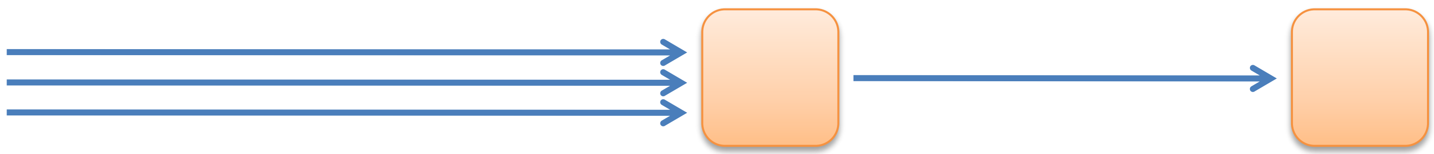
- RS does no forwarding
- Probably two independent RSes
- O(n) vs O(n^2) sessions
- But what about policy?

Multi-RIB architecture



- Multiple internal RIBs, with import and export policy
- Emulates the "original" connectivity
- From the outside, this is neighbor-specific BGP

Algebraic model



Routes come in

Choose the best

Pass it on

S = the set of all possible routes

$\leq$ = preference among routes

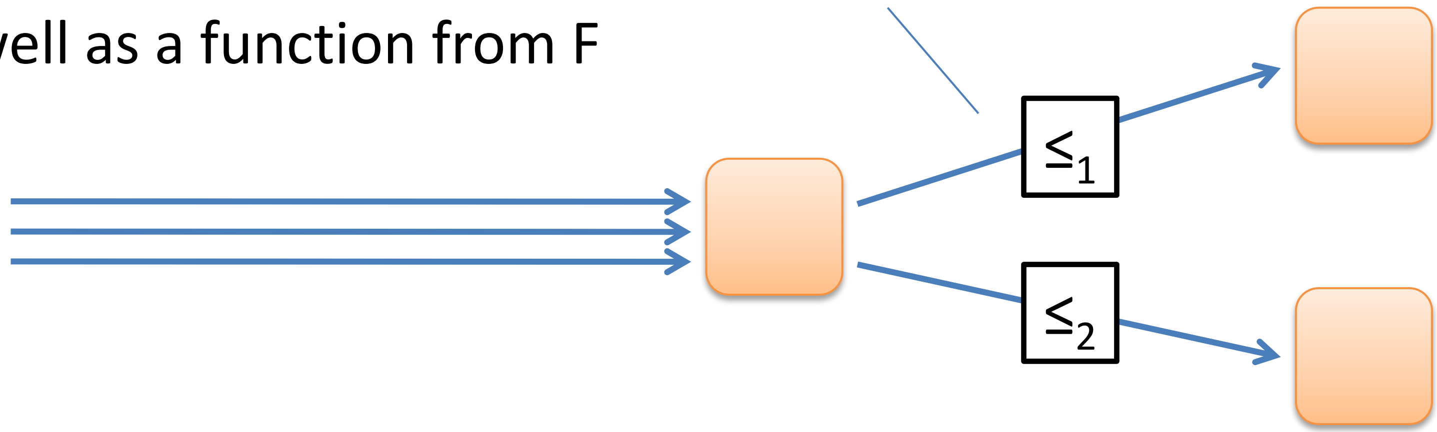
“$p < q$” means p is preferred to q

F = set of functions from S to S

f(p) is the extension of p along the arc labelled f

NS algebraic model

Now each arc has an associated order relation as well as a function from F



Routes come in

Choose the best

Pass them on

S = the set of all possible routes

≤ = preference among routes

F = set of functions from S to S

Route selection

- Incoming *sets* of routes: R_1, R_2, R_3, R_4, ...
- Choose the best by $\min(\leq, R_1 \cup R_2 \cup \ldots) = B$
- Extend routes B to $f(B) = \{ f(b) \mid b \text{ in } B \}$
- Export $\min(\leq_i, f(B))$ along arc *i*

- Here, $\min(\leq, R) = \{r \text{ in } R \mid \forall s \text{ in } R: \neg(s < r)\}$

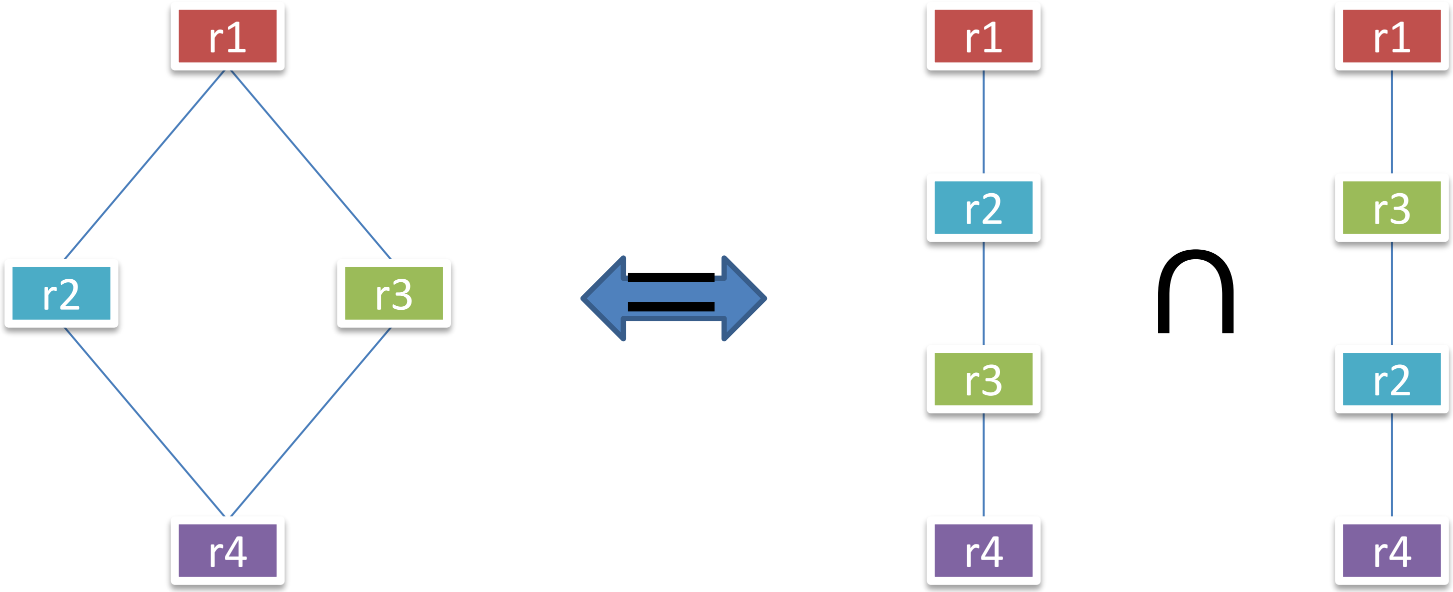
Previous correctness results

- Ordinary “single-path” BGP:
  - Gao-Rexford valley-freedom
  - “Strictly inflationary” path algebra
  - Restrict to shortest-paths or similar
- Equal-cost multipath variant: same conditions carry across (*mutatis mutandis*)
- Neighbor-specific multipath: Gao-Rexford OK, anything else unknown

Path algebra condition for NS

- Normal BGP: It's enough to have $p < f(p)$ whenever $f(p)$ is a permitted path
  - This is the "strictly inflationary" condition
- Multipath: Same condition (!)
- NS-BGP: Same condition suffices, *provided* that each per-arc order $\leq_i$ is an *extension* of the "consensus order" $\leq$.

Orders and extensions



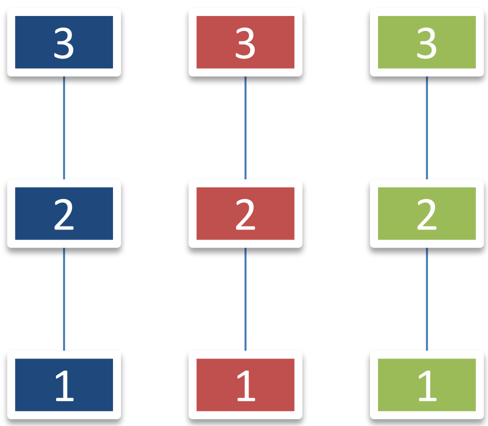
Orders and extensions

- An extension is a relational superset
  - If $a \leq b$ (original order) then $a \leq_E b$ (extended order)
  - Doesn't have to be linear
- Intersect all the extensions to get the original order back
- If the consensus order (intersection of all per-arc orders) satisfies strict inflation, then convergence is ensured.

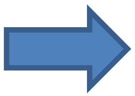
We are using preorders

- Reflexive and transitive
- Routes can be equivalent in preference
  - Example: ASPATHs of the same length
- Or they can be incomparable
  - If neither $a \leq b$ not $b \leq a$ is true, then we can choose either way for the extension
  - Incomparability in the consensus order means: decisions about which route is better are made *locally*

Example: Route classes

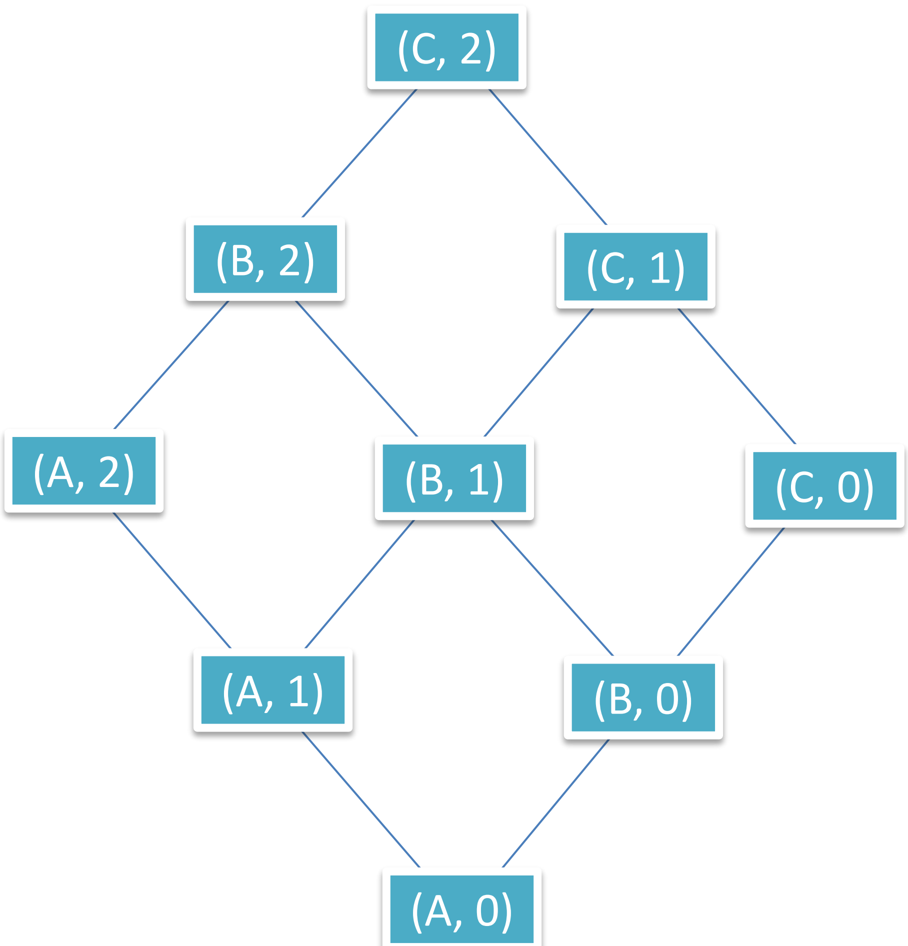


Three different classes of route: routes from different classes are incomparable.



In an extension, we can decide how the classes are ordered (here, blue is best and red is worst).

Example: Lexicographic choice

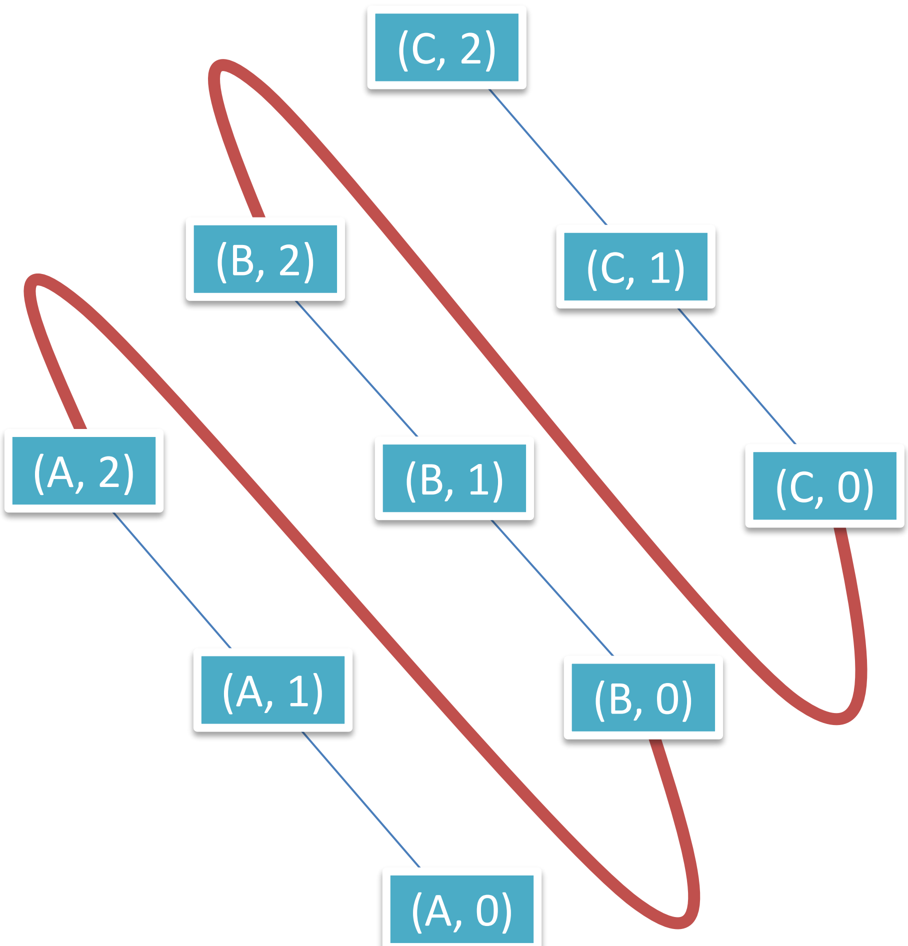


Here, each route has two attributes. One has A < B < C, and the other has 0 < 1 < 2.

We have not decided which attribute will be the dominant one.

So for example, (A, 2) is incomparable to (B, 0).

Example: Lexicographic choice



In an extension, we can make a determination either way.

This results in the formation of a lexicographic product.

In this case, A-B-C dominates: it is considered first, and then 0-1-2 is the tiebreaker.

Summary so far

- In neighbor-specific routing, we have multiple routes, and an order (for selection) appears on each arc
- Convergence is guaranteed if the consensus order (intersection of all individual orders) has the "strict inflationary" property
- What about NS-*BGP*?

NS-“BGP”, approximately

- Routes have these attributes, considered in order:
  1. Local preference *bigger is better*
  2. AS path *shorter paths are better*
  3. (Extra stuff) *no preferences*
- Extensions can place any order on part 3, but they cannot do anything else
  – The intersection (for part 3) is the discrete order

What makes it strictly inflationary?

"Local preference then AS path then extra stuff"

Strictly inflationary
(extended paths are always longer)

Restrict to inflationary policies
(not necessarily strict)

Anything goes!

Special case: Gao-Rexford

Consequences

- A great deal of freedom is possible: the additional attributes can be anything
- But neighbor-specific choices *cannot* be allowed to override the attributes responsible for global safety
  - In particular, local preference.

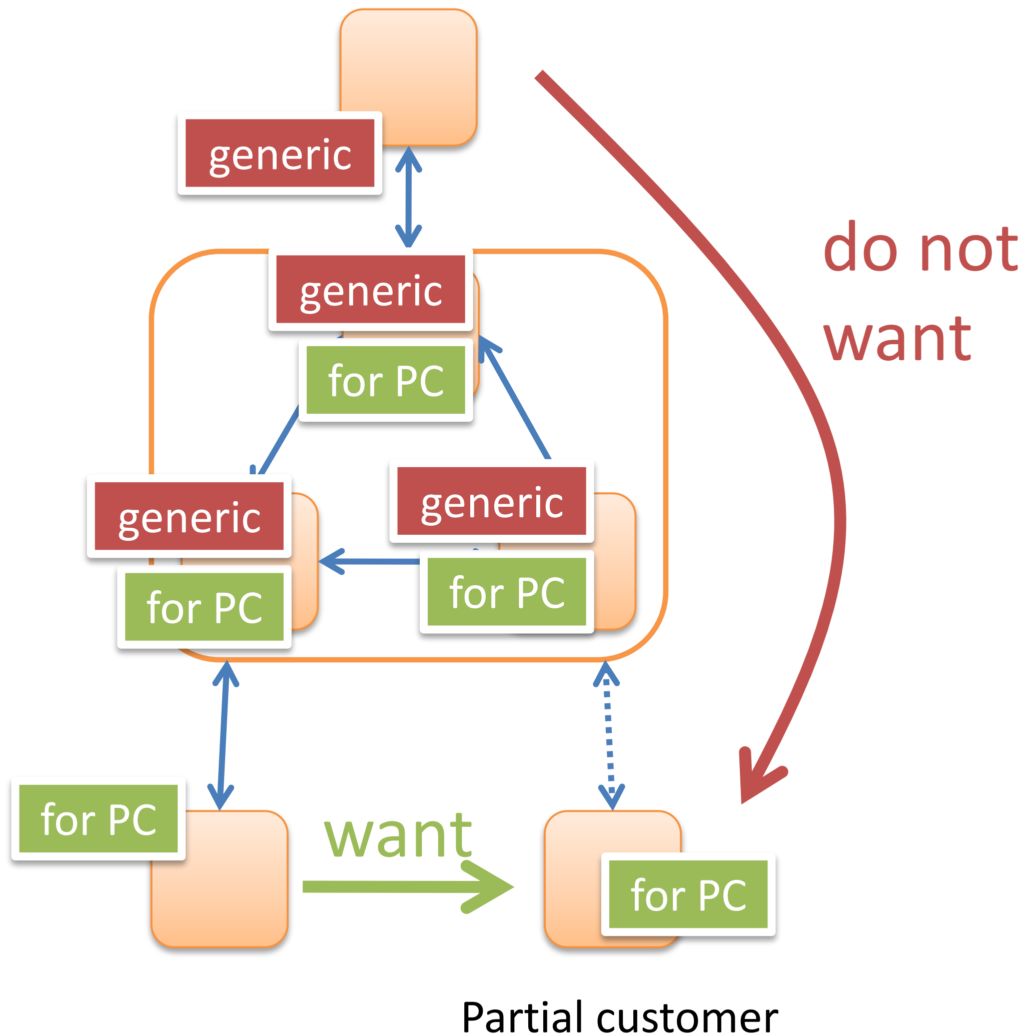
Partial customer, revisited

We need two route classes: those suitable for anyone, and those which are for the partial customer.

On most arcs, these are considered equivalent.

But a different order is used on the PC-facing arc, to choose the desired routes.

(In fact, the generics should be filtered out as well.)



Partial customer, revisited

- We can show that this scheme converges, using the "extension" idea
- *But* it can only be applied if local preference and AS path conditions permit!
  - We have to set these attributes when receiving a route, so that the neighbor-specific preferences are not overruled.
- Similar remarks apply to the IXP example

Proof generality

- Any arc can carry multiple routes (not just the iBGP sessions) if desired
- We assume a routing model based on local choices, converging to Nash equilibrium
- Not just BGP-like policies are covered, but any satisfying the conditions
- Once the consensus order has been proved safe, any of its extensions are then OK to use – even those not originally considered.

Conclusions

- There are a lot of possibilities for adding neighbor-specificity to policy-based routing protocols.
  - Some of them even work!
- Neighbor-specificity must be kept under control if we value convergence
- There are many unanswered questions about deployment of this technology