An algebraic framework for a unified view of route-vector protocols

João Luís Sobrinho

Instituto Superior Técnico, Universidade de Lisboa

Instituto de Telecomunicações

Portugal

Outline

1. Context for route-vector protocols (BGP, etc.)
2. Basics of the algebraic theory of routing
3. Optimality of paths (IGRP) ■
4. Usable connectivity and visibility (BGP) ■
5. Termination in loop-free states (BGP) ■
6. Survey of applications
7. Conclusions

Outline

1. Context for route-vector protocols (BGP, etc.)

Route-vector protocols

- Routing
  - Selection of paths in a network
- Routing protocols
  - Distributed algorithms to select paths in a network
- Route-vector protocols
  - Separate computation per destination
  - Routes learned from neighbors; “best” route announced to neighbors

Route-vector protocols in the Internet - I

- Border Gateway Protocol (BGP)
  - The inter-domain routing protocol of the Internet
  - Routing policies: LOCAL-PREF, AS-PATH, COMMUNITY, MULTI-EXIT-DISC, etc.
  - Used as well in the enterprise and in data-centers
- Routing Information Protocol (RIP)
  - Shortest paths

Route-vector protocols in the Internet - II

- Interior Gateway Routing Protocol (IGRP) and Enhanced IGRP (EIGRP)
  - Quality-of-service paths
- Interconnection of routing instances
  - Administrative Distance and Route Redistribution
- Wireless networks
  - Many metrics: hop-count, capacity, loss rate, interference level, energy consumption, etc.

Issues with route-vector protocols

- Non-termination (oscillations)
- Forwarding loops
- Sub-optimal paths
- Constraints on the usability of paths
- Hidden destinations

Limitations of case-by-case analysis

- Easy to overlook undesirable behaviors
- Repetition of arguments and of errors
- No insight across applications
- Little margin for automated management of routing configurations

Algebraic theory of routing

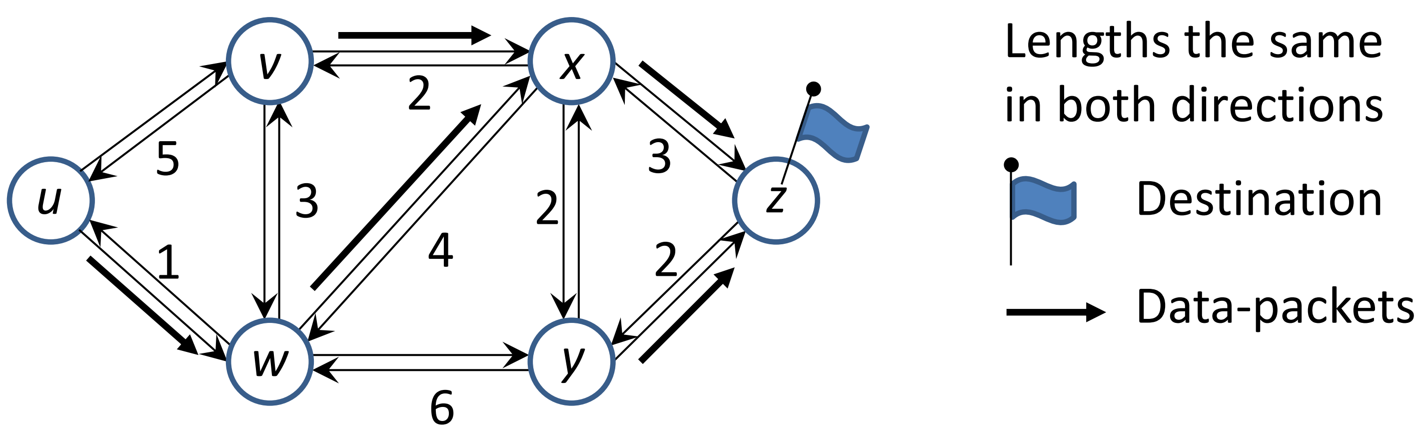
- Provides unified view of route-vector protocols
- Relates local routing decisions to global routing behaviors
- Facilitates specification, design, configuration, and analysis
- Gives lots of insight!

Outline

2. Basics of the algebraic theory of routing

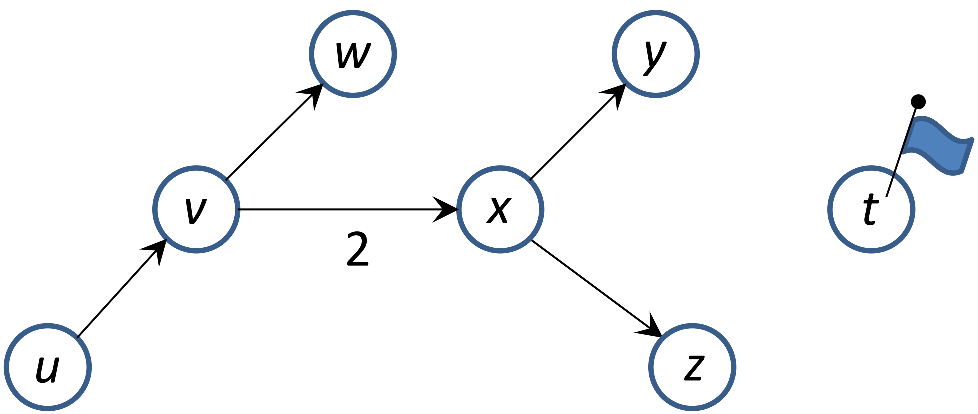
Shortest-path routing

- Each link has a **length**
- Length of a path is the **sum** of the lengths of its links
- Select paths of **minimum** length (shortest paths)



Distance-vector protocol - I

- Separate computation per destination

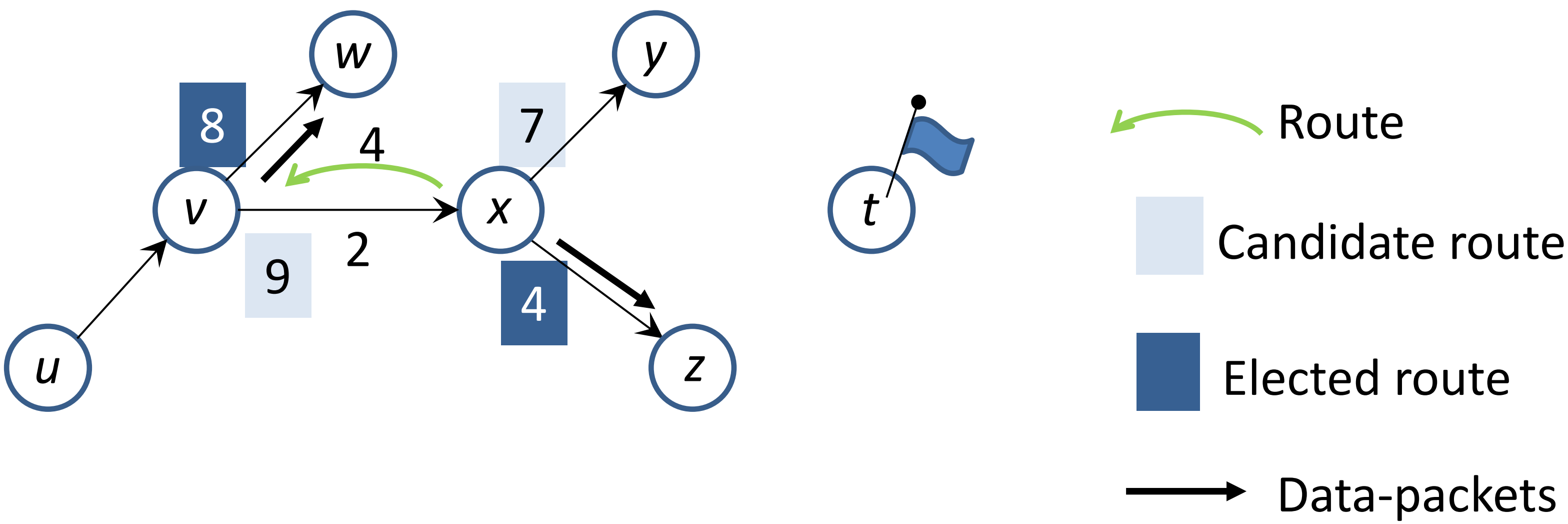


Only shown:

- Destination *t*
- Link *vx* of **length** 2
- Neighbors of *v* and of *x*

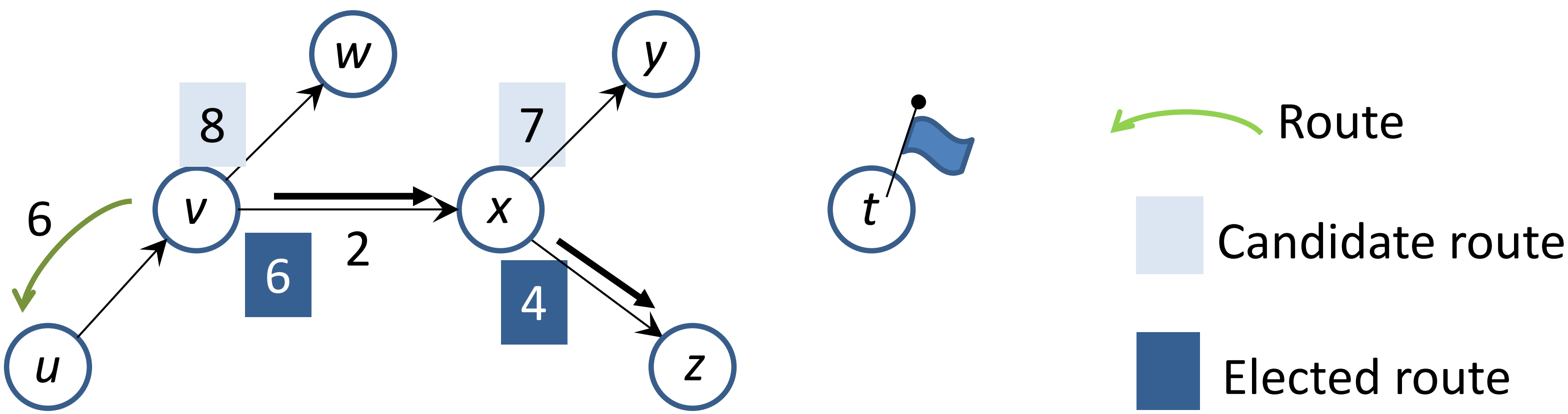
Distance-vector protocol - II

- Routes associate a **length** to a destination
- Local state: candidate routes and elected route



Distance-vector protocol - III

- Reception of a route
  - **extension** into a candidate route (+)
  - **election** of a route (min)
  - elected route sent to neighbors



Distances are computed
Data-packets travel along shortest paths

Question about the algorithm

- Can the simple algorithm underlying distance-vector protocols be used to compute other types of paths, related, for instance, to quality-of-service?

Question about the algorithm

- Can the simple algorithm underlying distance-vector protocols be used to compute other types of paths, related, for instance, to quality-of-service?

Idea: create framework for generic path attributes and how they are combined by the operations of election and extension

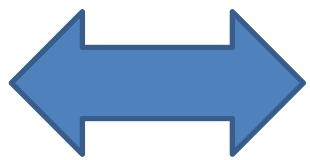
Routing algebra $(\Sigma, \bullet, \sqcap, \otimes)$

- Attributes, Σ; unreachability, $\bullet \in \Sigma$
- Election operation, $\sqcap$
  - Selectivity: $\alpha \sqcap \beta$ is either α or β, for $\alpha, \beta \in \Sigma$
  - Commutativity: $\alpha \sqcap \beta = \beta \sqcap \alpha$, for $\alpha, \beta \in \Sigma$
  - Associativity: $(\alpha \sqcap \beta) \sqcap \gamma = \alpha \sqcap (\beta \sqcap \gamma)$, for $\alpha, \gamma, \beta \in \Sigma$
  - Identity: $\alpha \sqcap \bullet = \alpha$, for $\alpha \in \Sigma$
- Extension operation, $\otimes$
  - Associativity: $(\alpha \otimes \beta) \otimes \gamma = \alpha \otimes (\beta \otimes \gamma)$, for $\alpha, \gamma, \beta \in \Sigma$
  - Annihilation: $\alpha \otimes \bullet = \bullet$, for $\alpha \in \Sigma$

[Sobrinho, 2002]

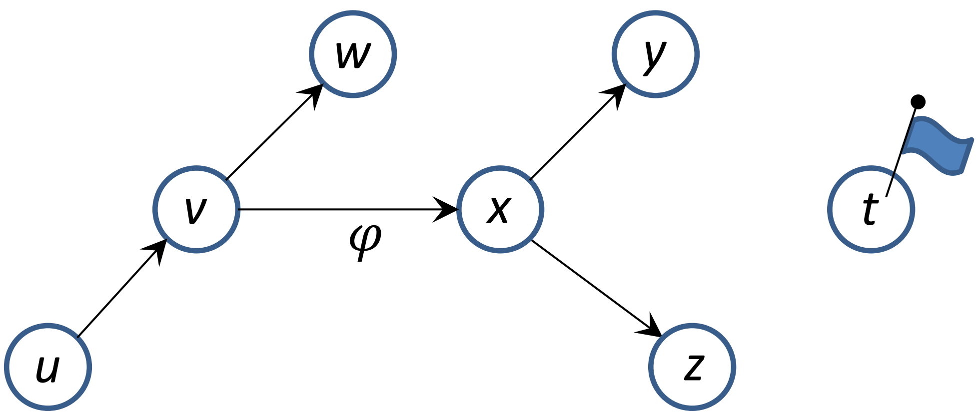
Equivalence between election and order

$\alpha \preccurlyeq \beta$ if $\alpha \sqcap \beta = \alpha$ for $\alpha, \beta \in \Sigma$

Election operation $\sqcap$  Total order $\preccurlyeq$

Route-vector protocol - I

- Separate computation per destination

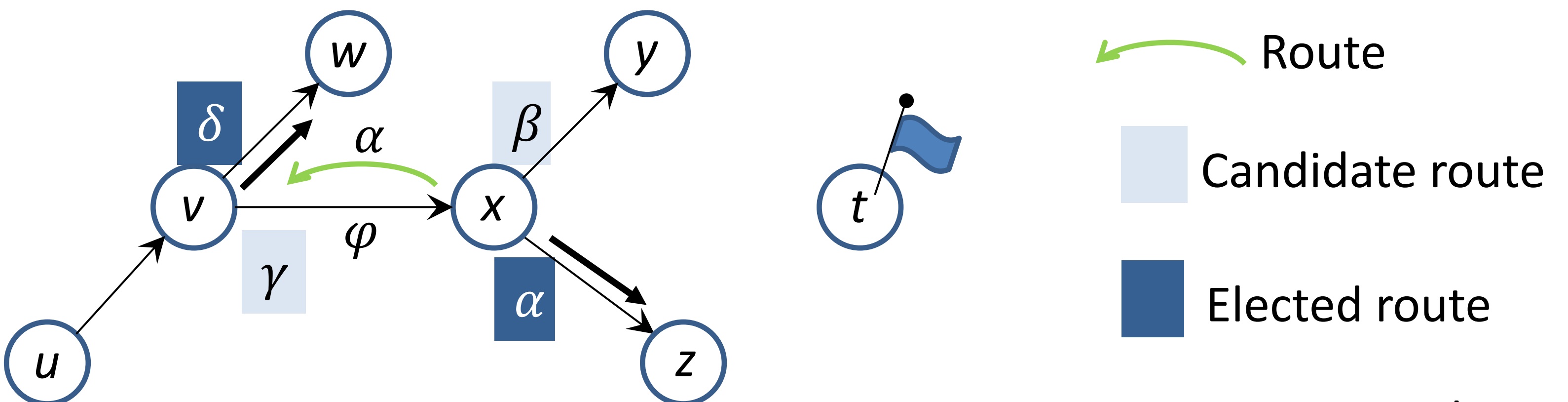


Only shown:

- Destination *t*
- Link *vx* with **attribute** φ
- Neighbors of *v* and of *x*

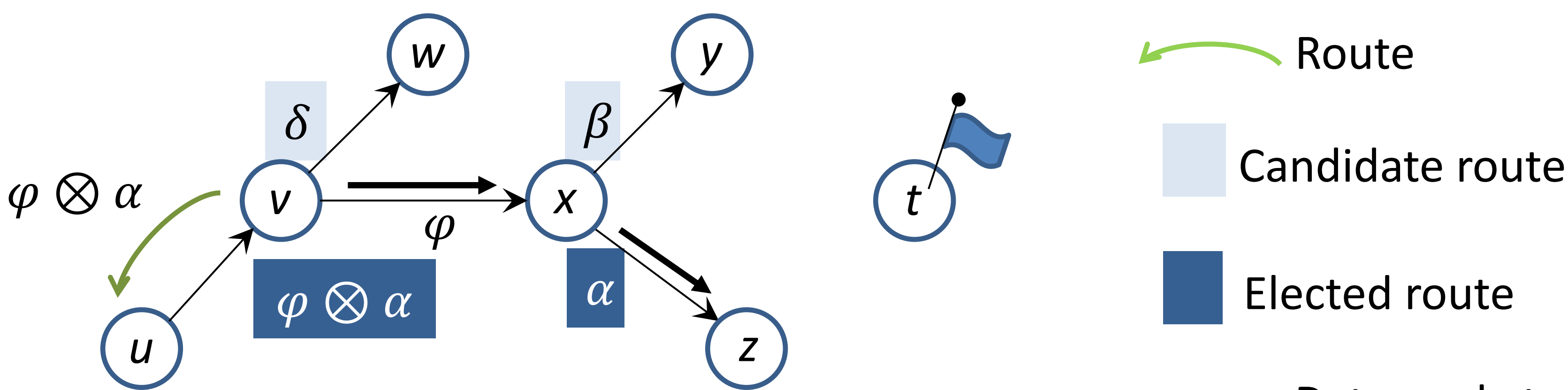
Route-vector protocol - II

- Routes associate an **attribute** to a destination
- Local state: candidate routes and elected route



Route-vector protocol – III

- Reception of a route
  - **extension** into a candidate route ($\otimes$)
  - **election** of a route ($\sqcap$)
  - elected route sent to neighbors



Assuming $\varphi \otimes \alpha \prec \delta$

Routing algebras and shortest paths

$(\Sigma, \bullet, \sqcap, \otimes)$

Route-vector protocol

$(\mathbb{Z} \cup \{+\infty\}, +\infty, \min, +)$

Distance-vector protocol

In practice, lengths are finite and addition is truncated

Outline

3. Optimality of paths (IGRP)

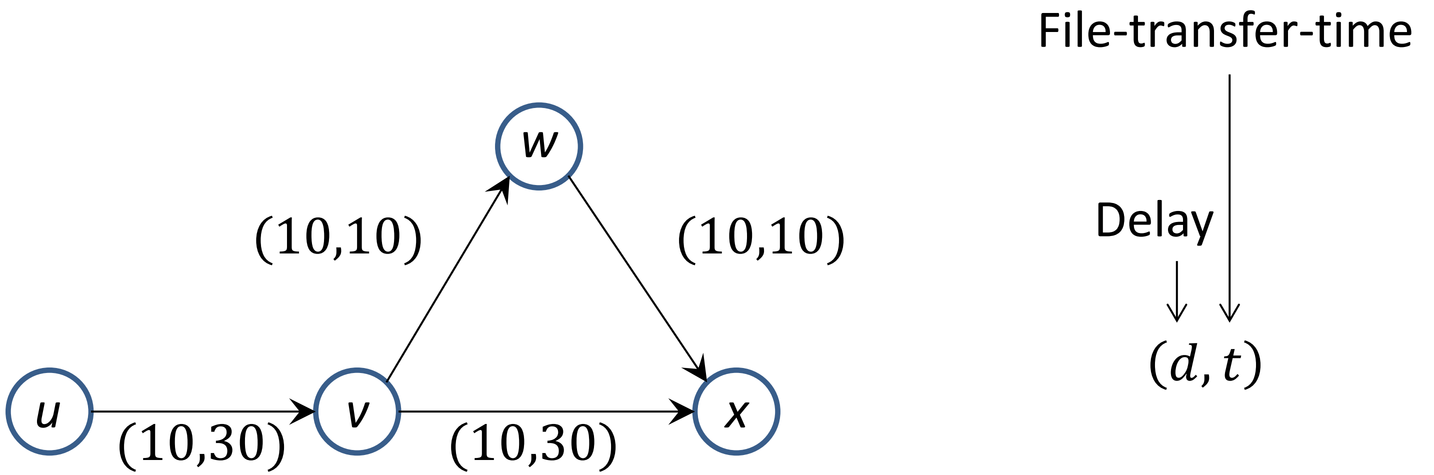
Quickest-path routing

- Each link has a **delay** and a **file-transfer-time**
- Delay of a path: **sum** of the delays of its links
- File-transfer-time of a path: **maximum** file-transfer-time among those of its links
- Select paths of **minimum** latency (quickest paths)
  - latency of a path: delay plus file-transfer-time

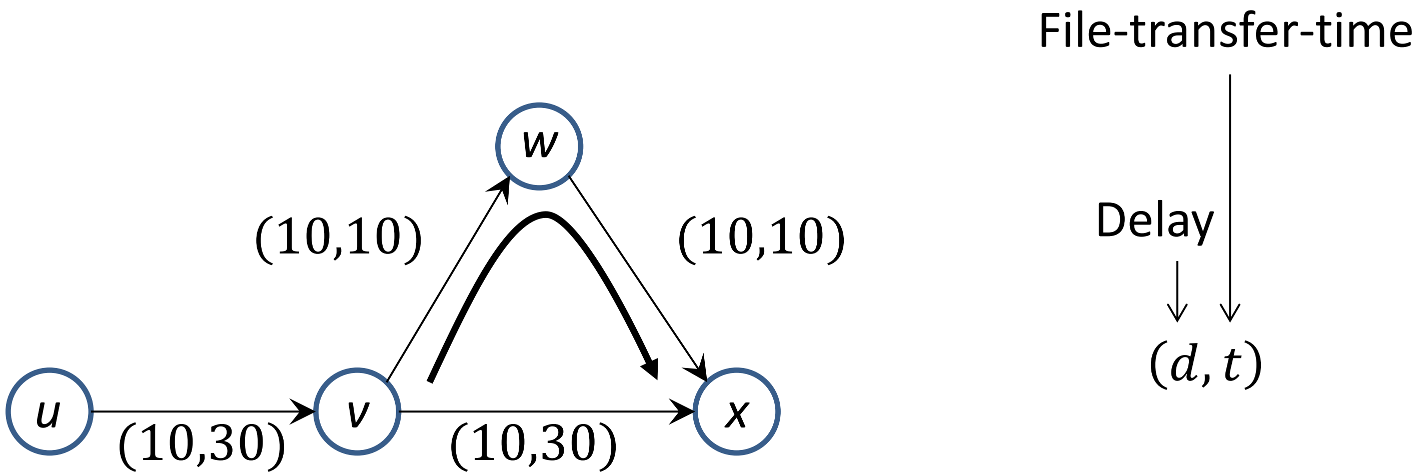
Quickest-path routing algebra

- Attributes
  - Pairs (d, t), with delay d and file-transfer-time t
- Total order
  - $(d_1, t_1) \prec (d_2, t_2)$ if $d_1 + t_1 < d_2 + t_2$
- Extension
  - $(d_1, t_1) \otimes (d_2, t_2) = (d_1 + d_2, \max\{t_1, t_2\})$

Quickest-path network - I



Quickest-path network - II



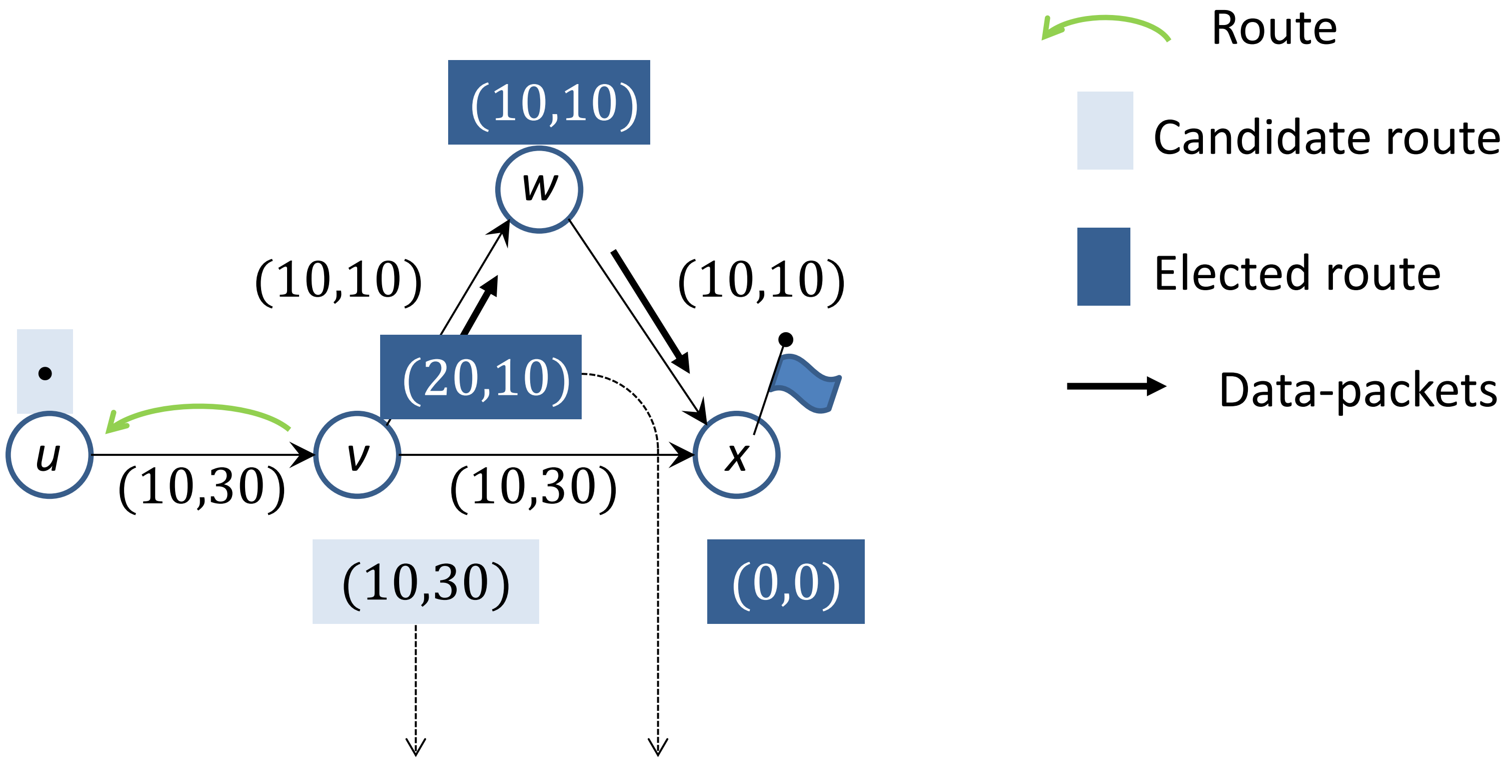
Pair of path *vwx*

$$(10, 10) \otimes (10, 10) = (10 + 10, \max\{10,10\}) = (20, 10)$$

Latency of path *vwx*

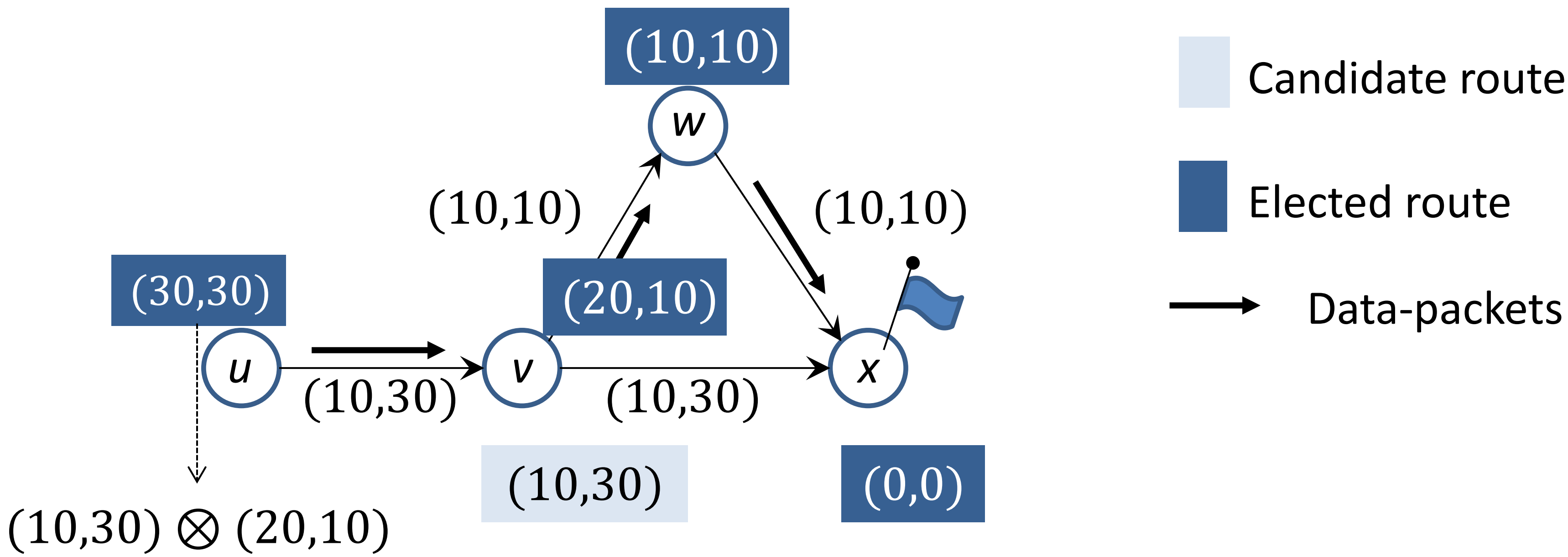
$$20 + 10 = 30$$

Internal Gateway Routing Protocol (IGRP)



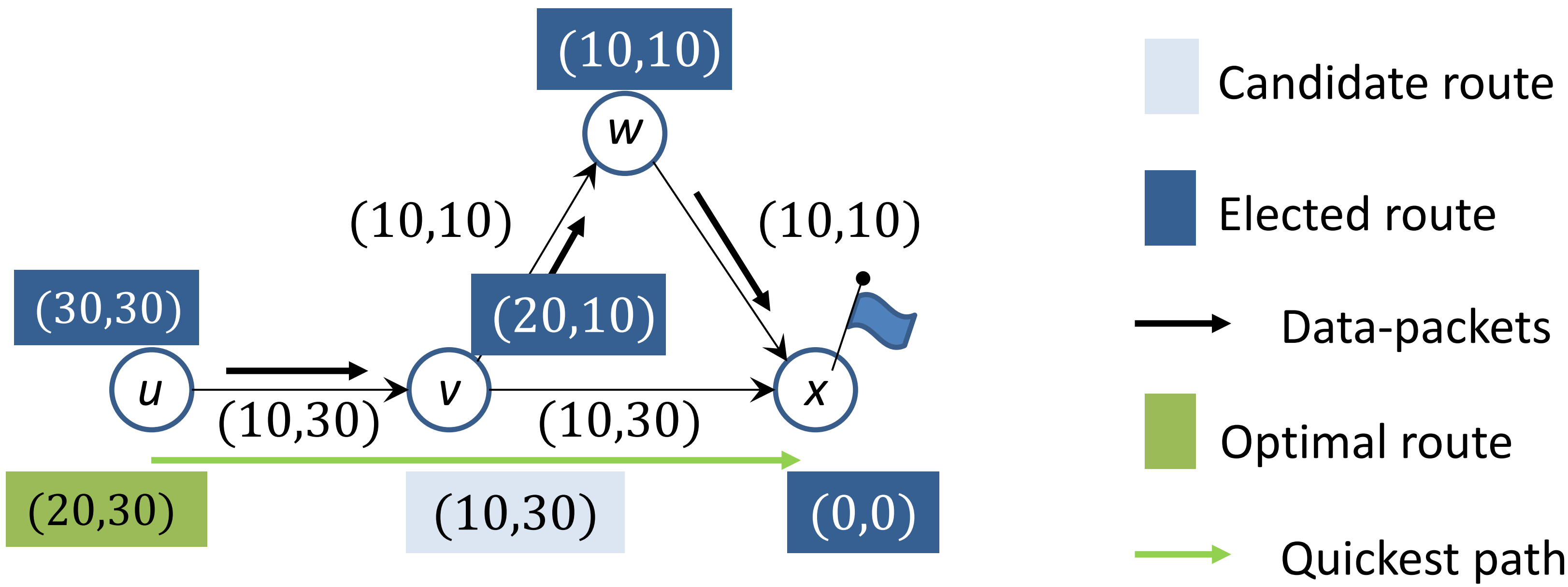
Latency $10 + 30 = 40 \quad > \quad 30 = 20 + 10$

IGRP



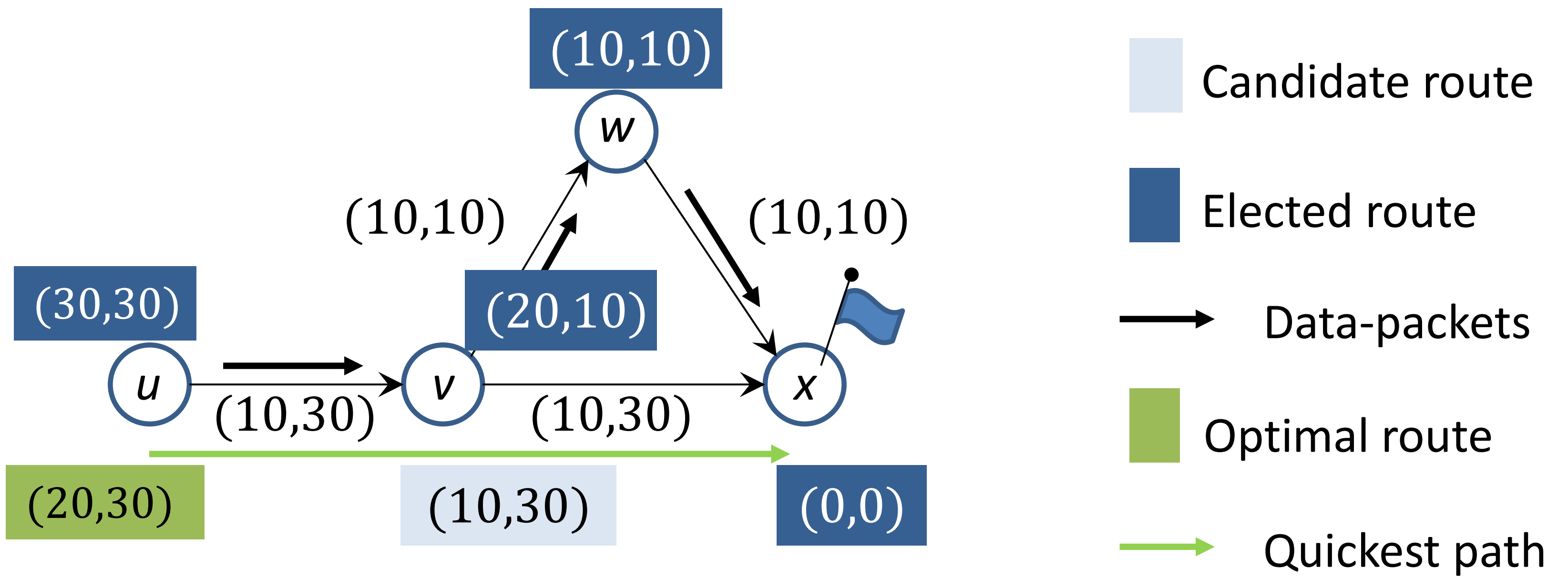
u* elects route $(30, 30)$, latency 60, corresponding to path *uvwx

IGRP: routes are not optimal



Optimal route at *u* is $(20,30)$, latency 50, corresponding to path *uvx*!

IGRP: no quickest paths



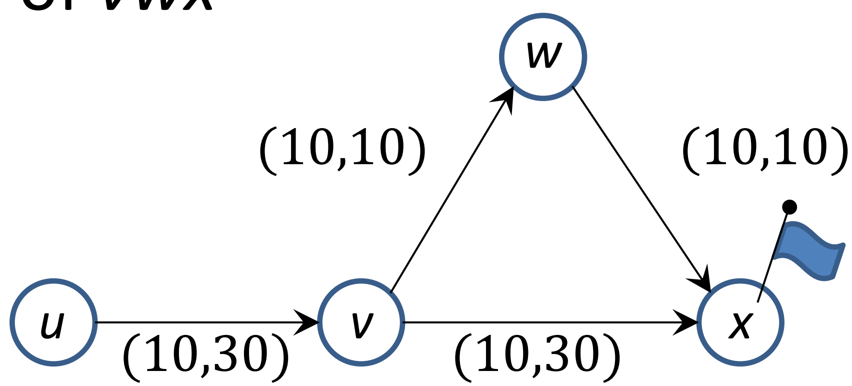
Data-packets do not travel along quickest paths!

[Sobrinho, 2002]

[Gouda and Schneider, 2003]

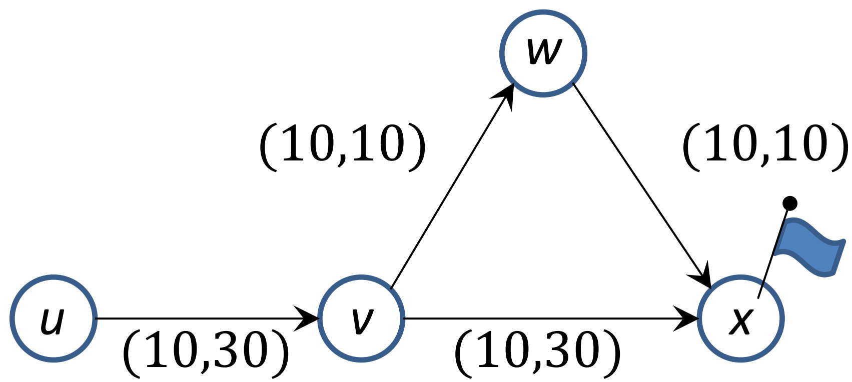
Semantic explanation

- *vwx* has smaller file-transfer-time, but larger delay, than *vx;* latency of *vwx* is smaller
- *uv* has a large transfer-time, meaning a low arrival rate of data-packets at *v*
- Once at *v*, data-packets do not benefit from the smaller transfer-time of *vwx*



Algebraic explanation

- The pair of link *uv* inverts the order between pairs



$(\mathbf{10}, \mathbf{30}) \otimes (10,30) = (20,30)$ $\qquad (20,10)$

$\prec \qquad \prec$

$(\mathbf{10}, \mathbf{30}) \otimes (20,10) = (30,30)$ $\qquad (10,30)$

Question about optimality

- When does a route-vector protocol compute optimal routes?

Isotonicity

- Attribute γ is isotone if extension does not invert preferences

$$\forall_{\alpha,\beta} \quad \alpha \preccurlyeq \beta \;\Rightarrow\; \gamma \otimes \alpha \preccurlyeq \gamma \otimes \beta$$

  – Extension distributes over election

$$\forall_{\alpha,\beta} \quad \gamma \otimes (\alpha \sqcap \beta) = (\gamma \otimes \alpha) \sqcap (\gamma \otimes \beta)$$

- Routing algebra is isotone if every attribute is isotone

Optimality of routes

Isotone
routing algebra

→

Optimality,
every network,
every destination

Distributed computation of a global optimum

[Sobrinho, 2002]

Isotonicity: quickest and shortest paths

Quickest-paths routing algebra

Not isotone

$$[\ (20,10) \prec (10,30)$$
$$\wedge$$
$$(10,30) \otimes (20,10) \succ (10,30) \otimes (10,30)\ ]$$

Sub-optimal paths

Shortest-paths routing algebra

Isotone

$$[\ l \leq m \ \Rightarrow n + l \leq n + m\ ]$$

Optimal paths

Outline

4. Usable connectivity and visibility (BGP)

Inter-domain routing

- Internet: the network of networks
  - Tens of thousands of Autonomous Systems (ASs)
  - Hundreds of thousands of destination IP prefixes
- Border Gateway Routing Protocol (BGP)
  - Route-vector protocol running among the ASs
- Routing policies
  - ASs configure BGP to satisfy their economic interests

Economic relationships between ASs

- Provider-customer relationship
  – Customer pays provider to transit its traffic

- Peer-peer relationship
  – Peers exchange traffic between them and their customers often without monetary compensations

Gao-Rexford (GR) policies: routes

- BGP messages carry **reachability** information
  - Autonomy and privacy
- GR routes
  - Customer route: reachability learned from a customer
  - Peer route: reachability learned from a peer
  - Provider route: reachability learned from a provider
  - Unreachability

[Gao and Rexford, 2001]

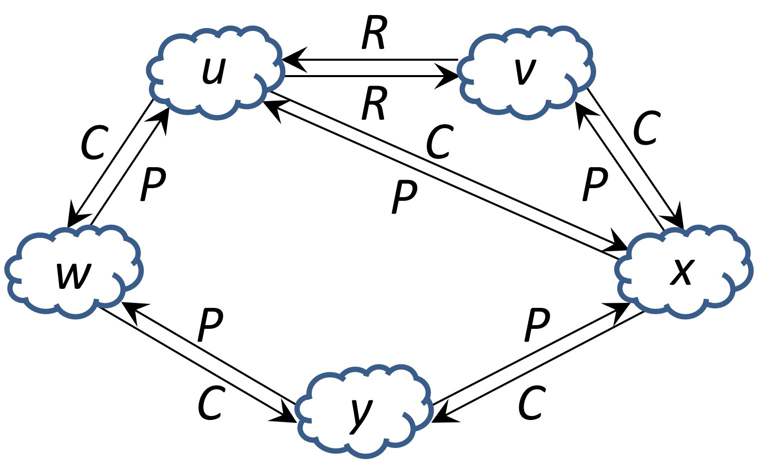
GR policies: preferences and exports

- GR preferences
  - First customer routes
  - Then peer routes
  - Then provider routes
- GR exports
  - All routes exported to customers
  - Customer routes exported to all neighbors

[Gao and Rexford, 2001]

GR network

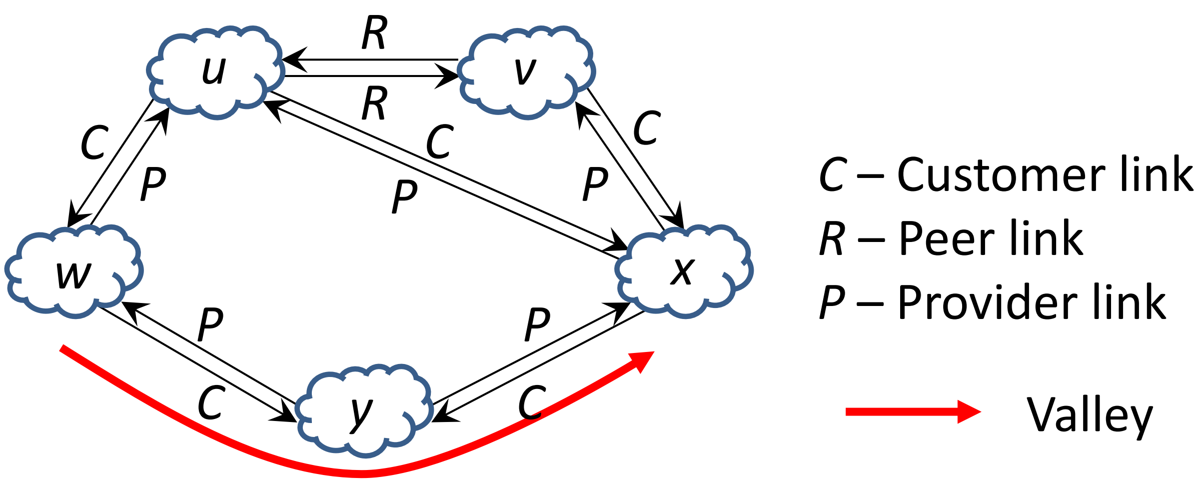
- *u* and *v* are peers
- *u* is a provider of *w* and *x*
- *v* is a provider of *x*
- *w* and *x* are providers of *y*



C – Customer link, provider to customer
R – Peer link, peer to peer
P – Provider link, customer to provider

GR network: unusable paths

- Valley
  - Customer or peer link then peer or provider link
- Unusable paths
  - Any path containing a valley



Questions about usability

A. Are unusable paths inherent to routing based exclusively on reachability information?

B. Can we quantify the usable connectivity of a network?

Questions about algebraic modeling

- Can **arbitrary** routing policies set with BGP be modeled algebraically?

Questions about algebraic modeling

- Can **arbitrary** routing policies set with BGP be modeled algebraically?

Idea: generalize extension from a binary operation to a set of maps on attributes

Routing algebra $(\Sigma, \bullet, \sqcap, \mathcal{T})$

- Attributes, Σ; unreachability, $\bullet \in \Sigma$
- Election operation, $\sqcap$
  - Selectivity: $\alpha \sqcap \beta$ is either α or β, for $\alpha, \beta \in \Sigma$
  - Commutativity: $\alpha \sqcap \beta = \beta \sqcap \alpha$, for $\alpha, \beta \in \Sigma$
  - Associativity: $(\alpha \sqcap \beta) \sqcap \gamma = \alpha \sqcap (\beta \sqcap \gamma)$, for $\alpha, \gamma, \beta \in \Sigma$
  - Identity: $\alpha \sqcap \bullet = \alpha$, for $\alpha \in \Sigma$
- Maps on Σ, called extenders, $\mathcal{T}$
  - Closure: $ST \in \mathcal{T}$, for $S, T \in \mathcal{T}$
  - Annihilation: $T(\bullet) = \bullet$, for $T \in \mathcal{T}$

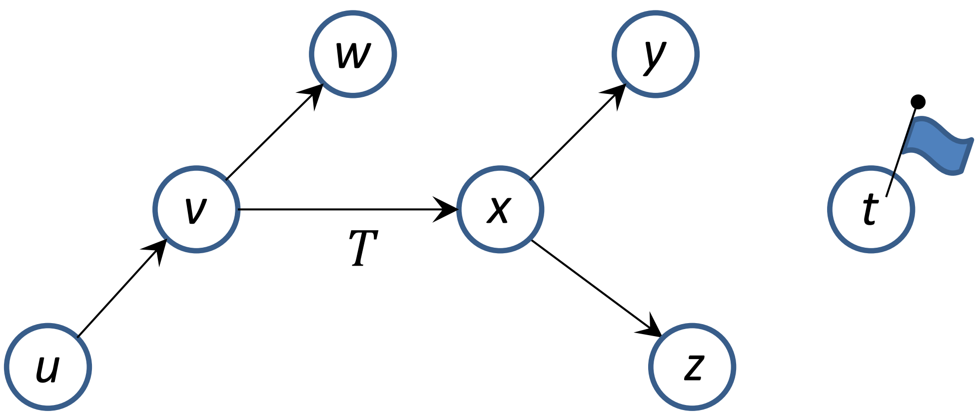
[Sobrinho, 2005]

Route-vector protocol - I

- Separate computation per destination

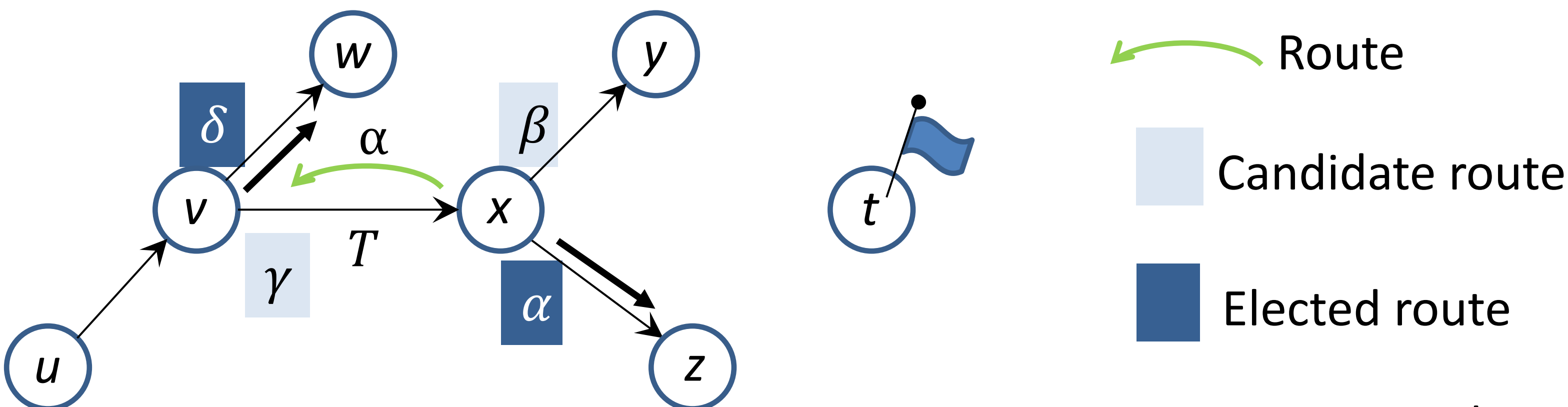
Only shown:

- Destination *t*
- Link *vx* with **extender** *T*
- Neighbors of *v* and of *x*



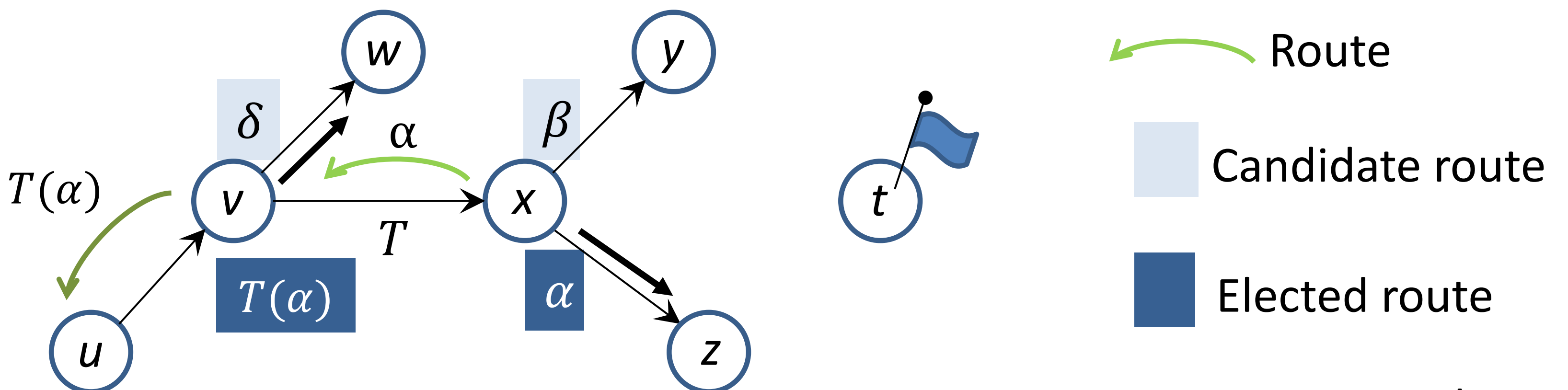
Route-vector protocol - II

- Routes associate an **attribute** to a destination
- Local state: candidate routes and elected route



Route-vector protocol - III

- Reception of a route
  - **extension** into a candidate route ($\mathcal{T}$)
  - **election** of a route ($\sqcap$)
  - elected route sent to neighbors



Assuming $T(\alpha) \prec \delta$

Isotonicity

- Extender T is isotone if it is an increasing map

$$\forall_{\alpha,\beta} \quad \alpha \preccurlyeq \beta \ \Rightarrow T(\alpha) \preccurlyeq T(\beta)$$

  – Extender is an endomorphism

$$\forall_{\alpha,\beta} \quad T(\alpha \sqcap \beta) = T(\alpha) \sqcap T(\beta)$$

- Routing algebra is isotone if all extenders are isotone

GR routing algebra: attributes and order

- Attributes
  - $\{c, r, p, \bullet\}$

c – Customer route
r – Peer route
p – Provider route

- Total order
  - $c \prec r \prec p \prec \bullet$

Customer routes,
then peer routes,
then provider routes

GR routing algebra: extenders

- Extenders
  - closure of $\{C, R, P\}$
  -

C – Customer link
R – Peer link
P – Provider link

Attributes

Extenders	c	r	p	$\bullet$
C	c	$\bullet$	$\bullet$	$\bullet$
R	r	$\bullet$	$\bullet$	$\bullet$
P	p	p	p	$\bullet$

$C(c) = c$ – Customer route exported to provider becoming a customer route

$C(r) = C(p) = \bullet$ – Peer and provider routes not exported to provider

GR routing algebra: isotonicity

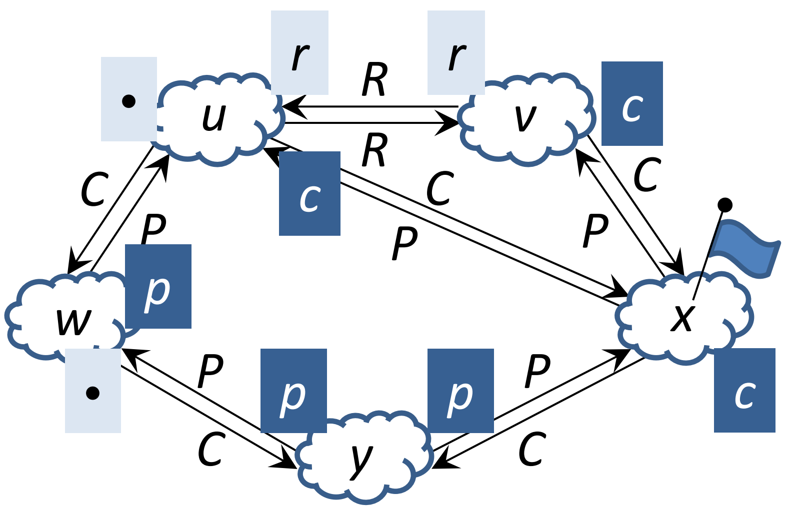
Attributes

Extenders	*c*	*r*	*p*	•	
	c	*r*	*p*	•	Increasing
C	*c*	•	•	•	Increasing
R	*r*	•	•	•	Increasing
P	*p*	*p*	*p*	•	Increasing

The Gao-Rexford routing algebra is isotone

GR network: stable state of BGP

- *u* and *v* are peers
- *u* is a provider of *w* and *x*
- *v* is a provider of *x*
- *w* and *x* are providers of *y*



C – Customer link
R – Peer link
P – Provider link

Candidate route

Elected route

Questions about usability

A. Are unusable paths inherent to routing based exclusively on reachability information?

Modeling reachability: next-hop

- Extender T is next-hop if its image has a single attribute different from unreachability

$$\forall_{\alpha,\beta} \quad T(\alpha) \prec \bullet \wedge T(\beta) \prec \bullet \;\Rightarrow\; T(\alpha) = T(\beta)$$

- Routing algebra is next-hop if all extenders are next-hop

Next-hop: use cases

- Inter-domain routing
  - Autonomy and privacy

- Interconnection of routing instances
  - Circumvention of comparison of attributes from different routing instances

GR routing algebra: next-hop

Attributes

Extenders	*c*	*r*	*p*	•
C	*c*	•	•	•
R	*r*	•	•	•
P	*p*	*p*	*p*	•

Same attribute

The Gao-Rexford routing algebra is next-hop

Usability of paths

Next-hop routing algebra ⇒ Some paths are unusable, every network with cycles (at least three nodes)

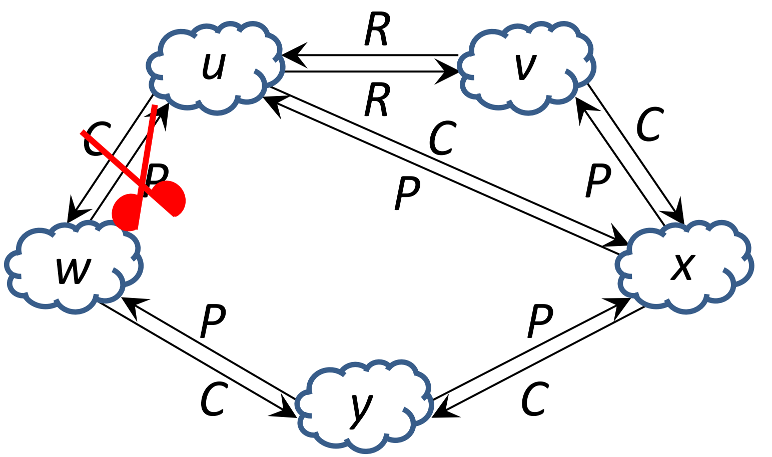
In order to avoid "bad behaviors," reachability information cannot be propagated all the way around a cycle

[Sobrinho, 2016]

Questions about usability

B. Can we quantify the usable connectivity of a network?

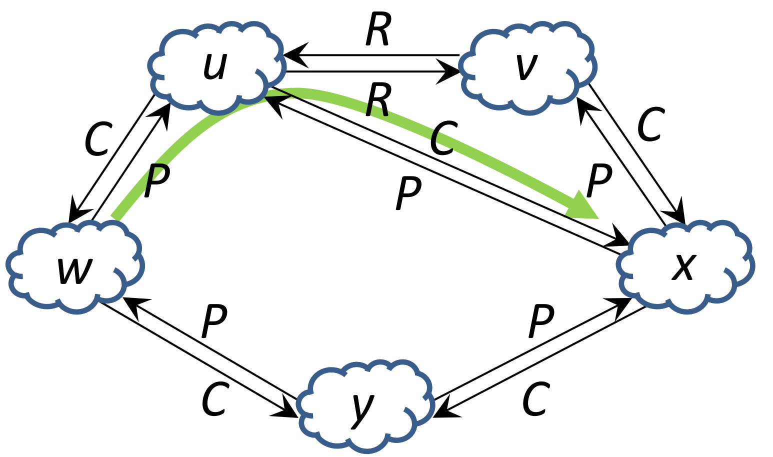
GR connectivity: usable separation



C – Customer link
R – Peer link
P – Provider link

One link **usably** separates *w* from *x*

GR connectivity: usable disjointness



C – Customer link
R – Peer link
P – Provider link

One link **usably** separates *w* from *x*

One usable link-disjoint path from *w* to *x*

Usable connectivity: duality and computation

Next-hop and isotone
routing algebra

Minimum number of links that **usably** separates source from target	**=**	Maximum number of **usable** link-disjoint paths from source to target

Common quantity computable in polynomial-time

[Sobrinho and Quelhas, 2012]

Question about visibility

- Given a usable path to a destination, will every node along the path be able to reach it?

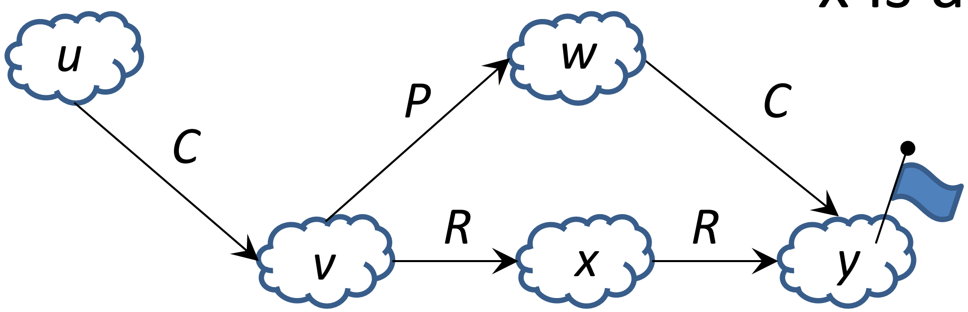
GR with Backups (GRBack)

- All exports are allowed **except** those from one provider to another
  - Violation of GR export rules
- Backup routes are usable routes other than customer, peer, or provider routes
- Backup routes increase **avoidance level** for every violation of GR export rules

[Gao et al., 2001]

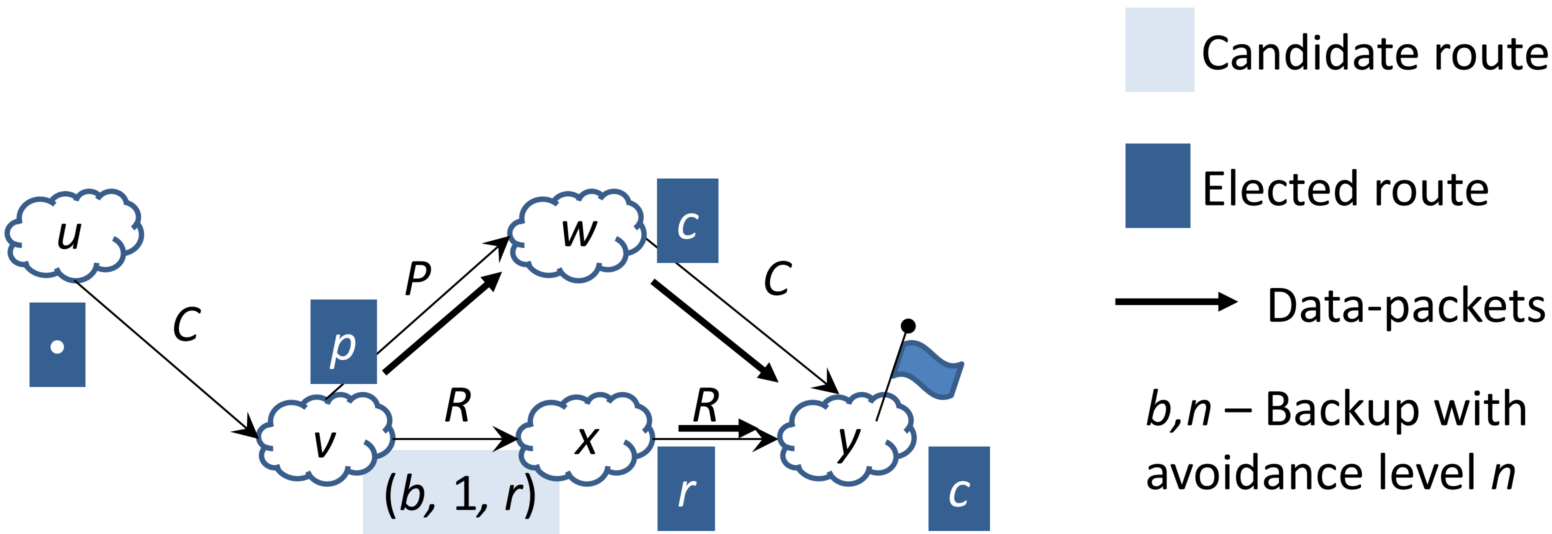
GRBack: visibility - I

- *u* and *w* are providers of *v*
- *w* is a provider of *y*
- x is a peer of *v* and *y*



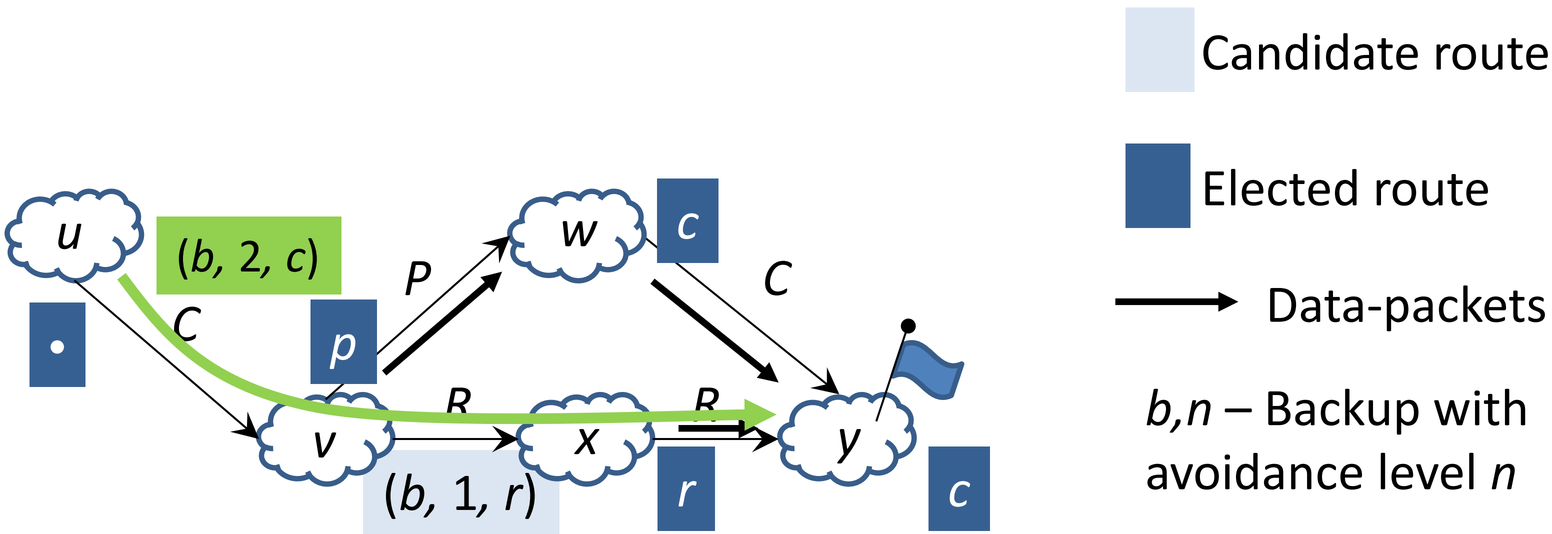
Links shown only in one direction

GRBack: visibility - II



u* does not reach *y

GRBack: visibility - III



There is a usable path from *u* to *y*

u* does not reach *y

***y* is not visible from *u*!**

Visibility of destinations

Isotone
routing algebra

Visibility,
every network,
every destination

[Sobrinho and Quelhas, 2012]

Outline

5. Termination in loop-free states (BGP)

GR with Peer+s (GRPeer+)

- Routes learned from a peer+ (peer+ routes) preferred to customer routes
  - Violation of GR preference rules

GRPeer+ routing algebra: attributes; order

- Attributes
  - $\{\boldsymbol{r}^{+}, c, r, p, \bullet\}$

r^{+} – Peer+ route
c – Customer route
r – Peer route
p – Provider route

- Total order
  - $\boldsymbol{r}^{+} \prec c \prec r \prec p \prec \bullet$

Peer+ routes,
then customer routes,
then peer routes,
then provider routes

GRPeer+ routing algebra: extenders

- Extenders
  - closure of $\{\boldsymbol{R}^{+}, C, R, P\}$
  -

R^{+} – Peer+ link
C – Customer link
R – Peer link
P – Provider link

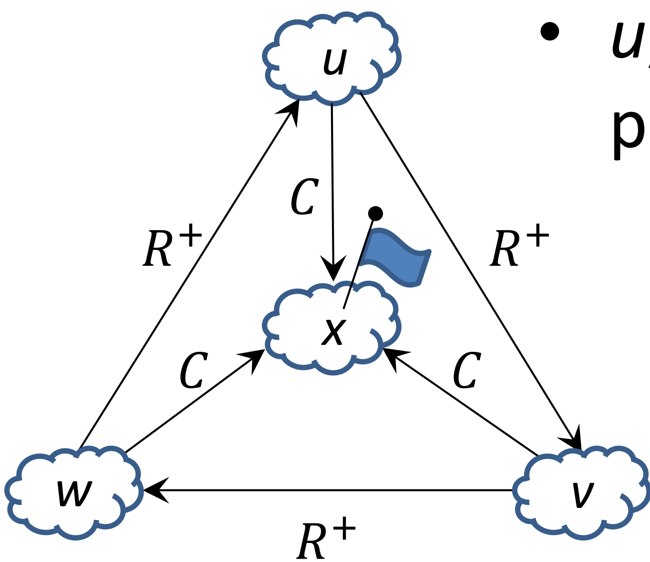
Attributes (columns) / Extenders (rows)

	$\boldsymbol{r}^{+}$	c	r	p	$\bullet$
$\boldsymbol{R}^{+}$	$\bullet$	$\boldsymbol{r}^{+}$	$\bullet$	$\bullet$	$\bullet$
C	$\bullet$	c	$\bullet$	$\bullet$	$\bullet$
R	$\bullet$	r	$\bullet$	$\bullet$	$\bullet$
P	p	p	p	p	$\bullet$

Only customer routes are exported to peer+s

Peer+ routes are exported only to customers

GRPeer+ network



- *u*, *v*, and *w* are providers of *x*
- *u*, *v*, and *w* are mutual peers
- *u*, *v*, and *w* prefer their clockwise peer (peer+) to *x*

R^+ – Peer+ link
C – Customer link

Links shown only in one direction

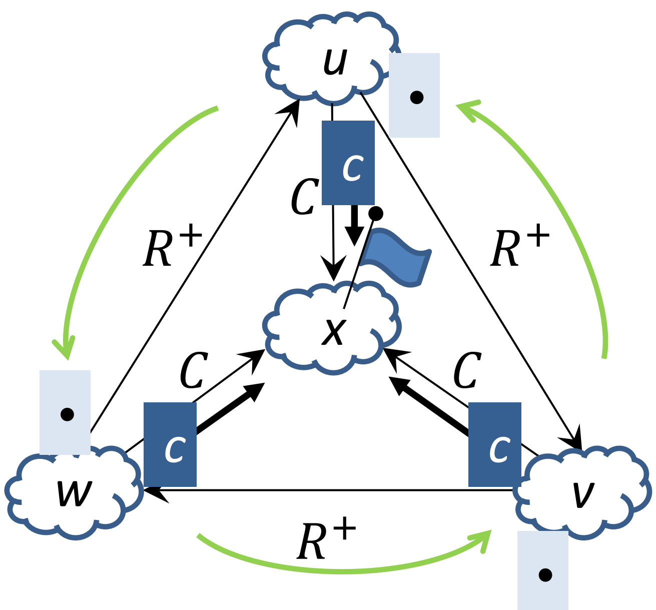
Non-termination - I

Candidate route

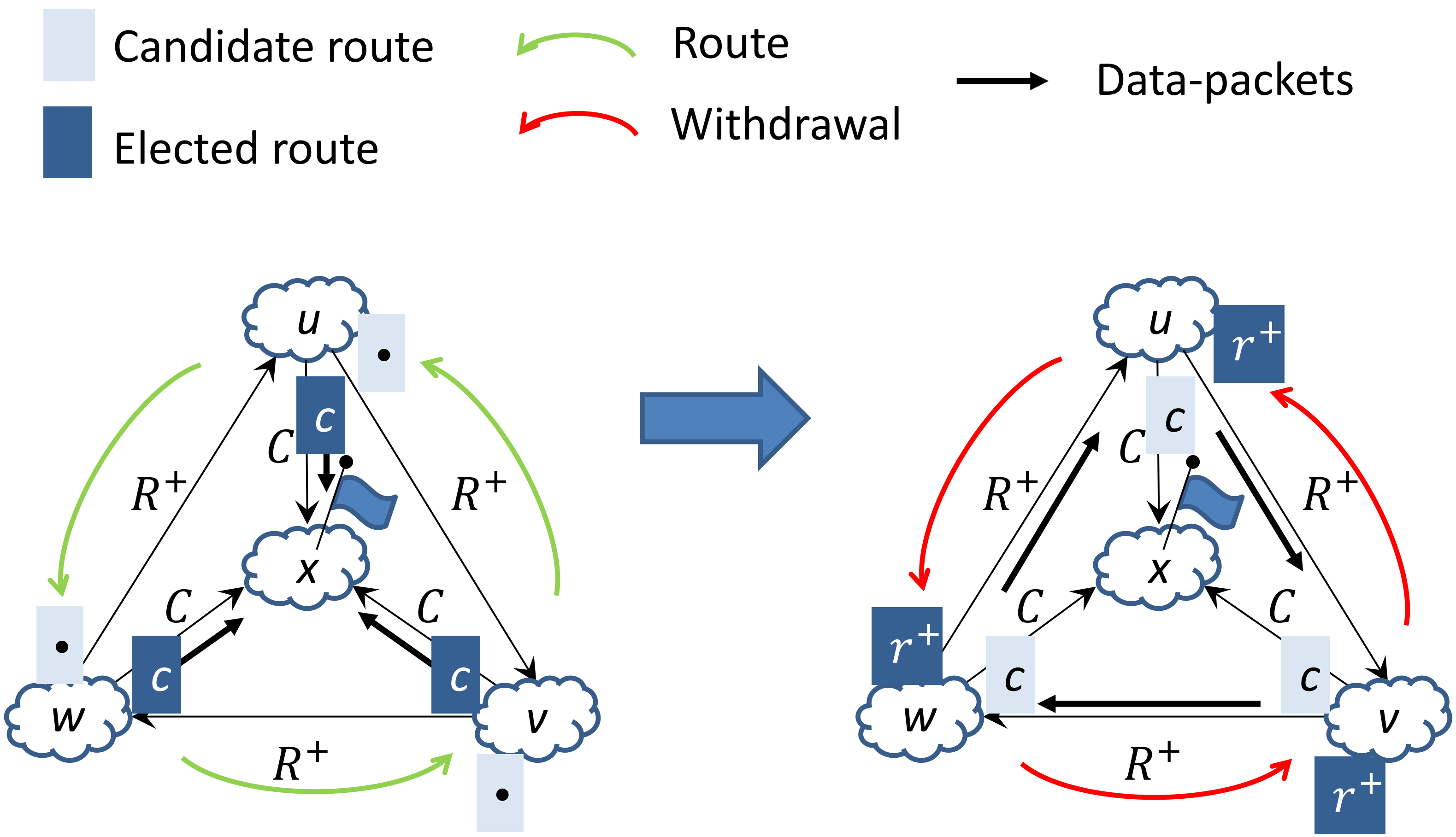
Elected route

Route

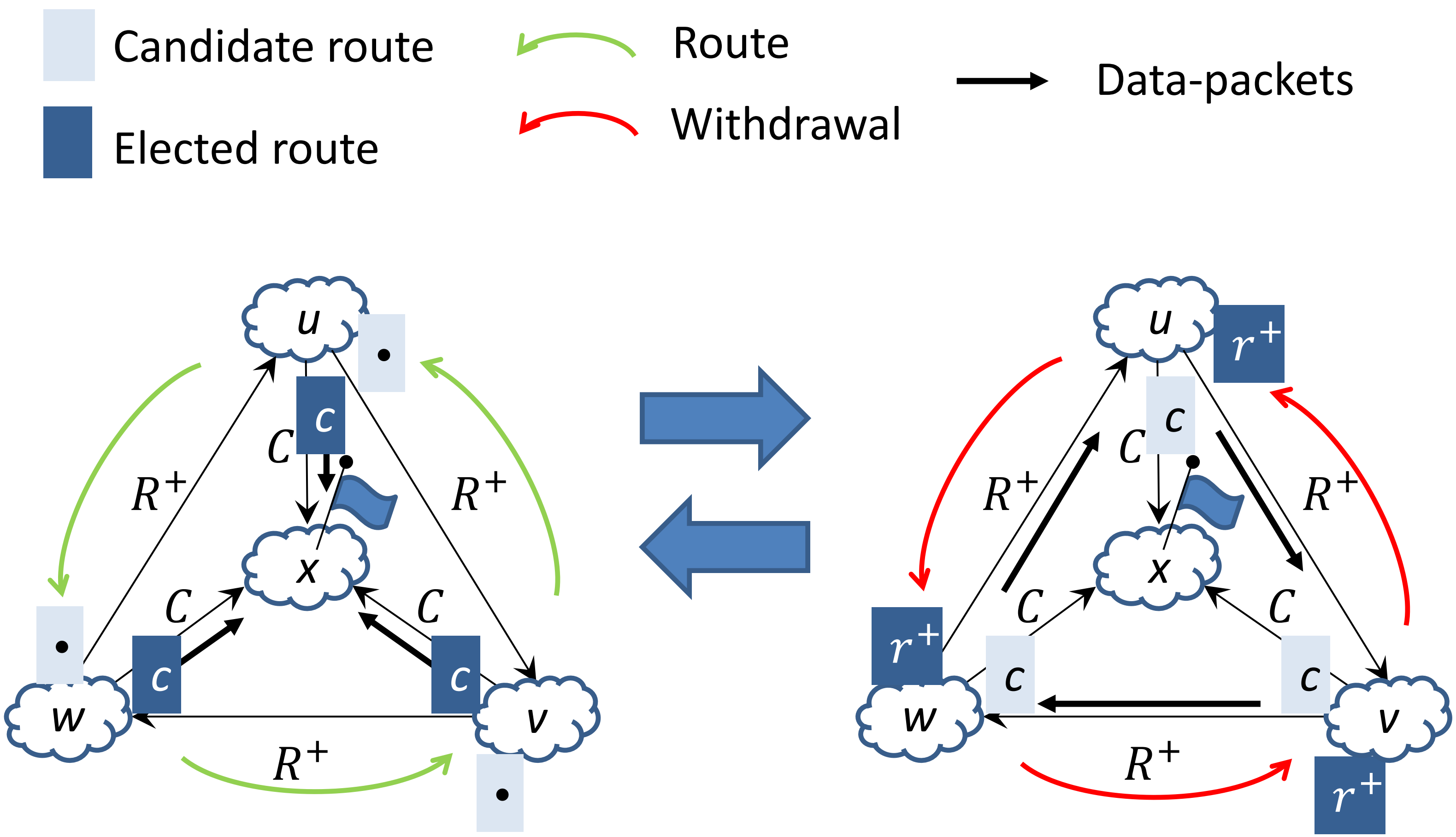
Data-packets



Non-termination - II



Non-termination - III



Correctness

- Termination
  - Stable state is reached, eventually
- No forwarding loops in stable state
  - Elected routes not learned around a cycle

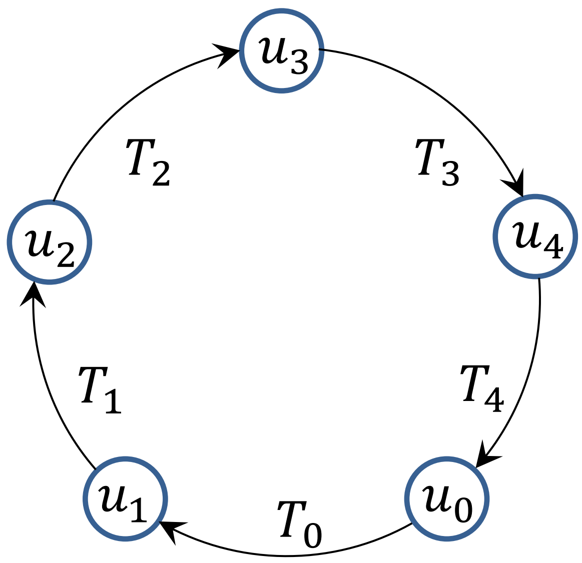
Question about correctness

- Can we characterize correctness in terms of routing configurations around the cycles of a network?

Strictly absorbent cycle - I

- Cycle $u_0 u_1 \cdots u_{n-1} u_0$, with T_i the extender of $u_i u_{i+1}$, is strictly absorbent if

$$\forall_{\alpha_0 \prec \bullet, \alpha_1 \prec \bullet, \ldots, \alpha_{n-1} \prec \bullet} \ \exists_i \ \alpha_i \prec T_i(\alpha_{i+1})$$

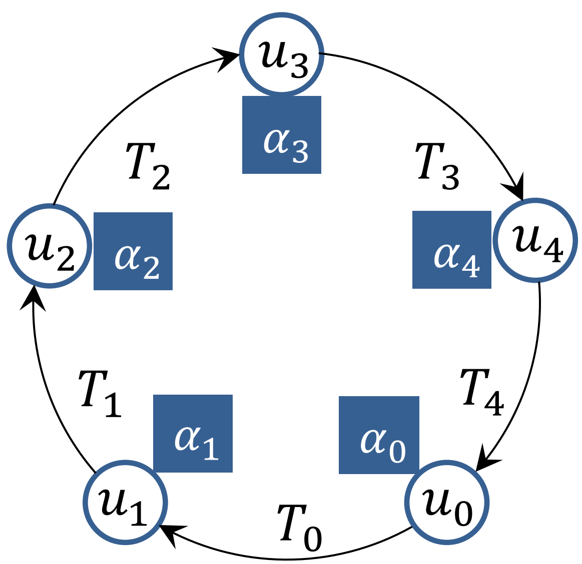


Strictly absorbent cycle - II

- Cycle $u_0 u_1 \cdots u_{n-1} u_0$, with T_i the extender of $u_i u_{i+1}$, is strictly absorbent if

$$\forall_{\alpha_0 \prec \bullet, \alpha_1 \prec \bullet, \ldots, \alpha_{n-1} \prec \bullet} \ \exists_i \ \alpha_i \prec T_i(\alpha_{i+1})$$

$\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$
external to the cycle

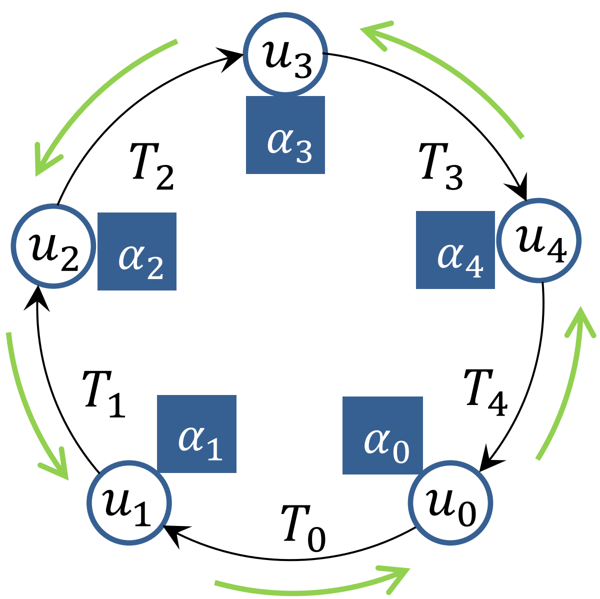


Strictly absorbent cycle - III

- Cycle $u_0 u_1 \cdots u_{n-1} u_0$, with T_i the extender of $u_i u_{i+1}$, is strictly absorbent if

$$\forall_{\alpha_0 \prec \bullet, \alpha_1 \prec \bullet, \ldots, \alpha_{n-1} \prec \bullet} \; \exists_i \; \alpha_i \prec T_i(\alpha_{i+1})$$

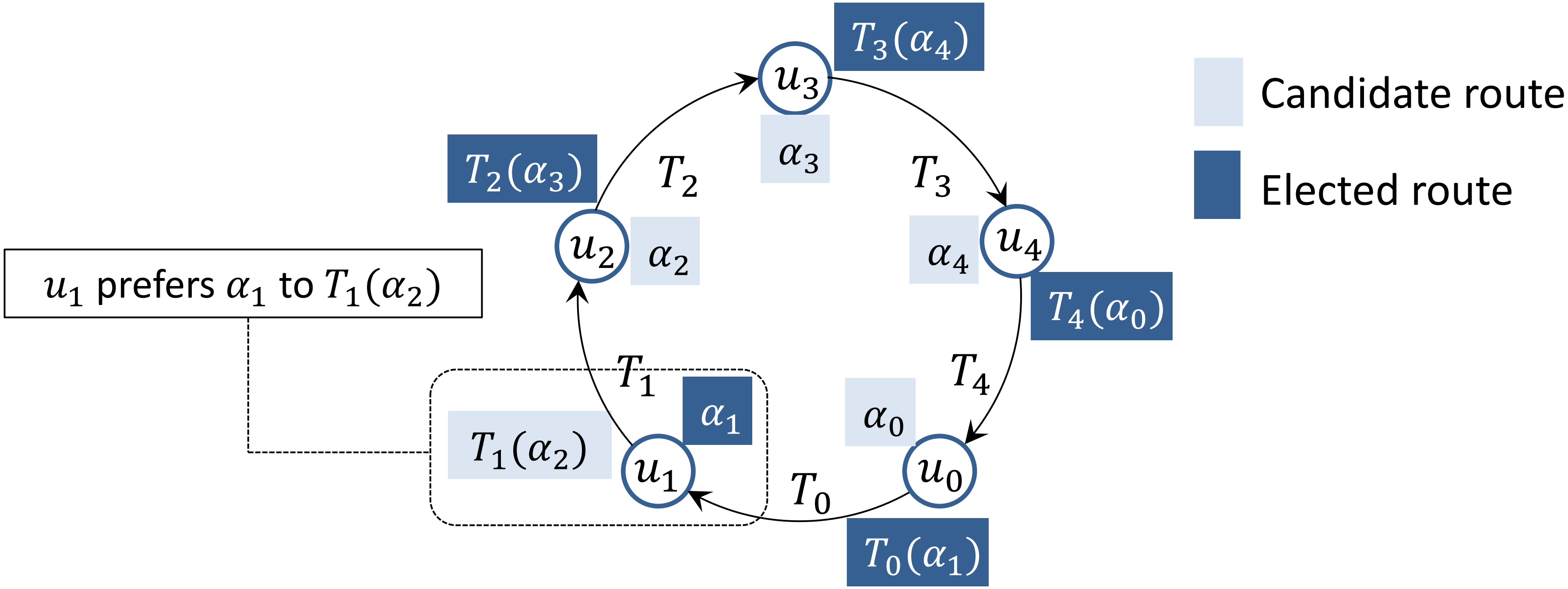
$\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$
sent around the cycle



Strictly absorbent cycle - IV

- Cycle $u_0 u_1 \cdots u_{n-1} u_0$, with T_i the extender of $u_i u_{i+1}$, is strictly absorbent if

$$\forall_{\alpha_0 \prec \bullet, \alpha_1 \prec \bullet, \ldots, \alpha_{n-1} \prec \bullet} \ \exists_i \ \alpha_i \prec T_i(\alpha_{i+1})$$



Correctness: forward implication

All cycles of the network strictly absorbent ⇒ Robust correctness, every destination (anycast destinations included)

[Griffin et al.,2002]

[Sobrinho, 2005]

Correctness: backward implication

All cycles of the network strictly absorbent ⇐ Robust correctness, every destination (anycast destinations included)

[Sobrinho, 2016]

Strict absorbency: GR variants

- GR and GRBack
  - – Cycle not formed exclusively by customer links
  - – Cycle not formed exclusively by provider links
- GRPeer+
  - – Cycle not formed exclusively by a mix of customer links and peer+ links
  - – Cycle not formed exclusively by provider links

∎

Outline

6. Survey of applications

Applications - I

- Sibling ASs
  - All routes are shared
  - **Guidelines for correctness**

[Liao et al., 2010]
[Sobrinho, 2016]

- internal BGP (iBGP)
  - Route reflection
  - **Guidelines for correctness and visibility**

[Griffin and Wilfong, 2002]
[Vissichio et al., 2012]

- Deployment of Secure BGP (S-BGP)
  - Security first, second, or last
  - **Efficient computation of stable states**
  - **Analysis of collateral damages**

[Lychev et al., 2013]

Applications - II

- Interconnection of routing instances [Le and Sobrinho, 2014]
  - **Current limitations**
  - **Better performance and reliability**
- Link-state protocols [Sobrinho, 2002] [Sobrinho and Griffin, 2010]
  - Separate computation of optimal paths over a common topology
  - **Conditions for efficient computation, correctness, and optimality**

Applications - III

- Distributed Route Aggregation on the Global Network (DRAGON)
  - **Filtering and aggregation of prefixes while respecting routing policies**
  - Filtering strategy: 49% savings in routing state
  - Filtering and aggregation strategies: 79% savings in routing state

[Sobrinho et al. 2014, www.route-aggregation.net]

Outline

7. Conclusions

Conclusions - I

- Framework to reason about routing protocols
  - Unified view of route-vector protocol behavior
  - Conditions relating local decisions to global behaviors
- Unified view of route-vector protocol behavior
  - Algebra of attributes equipped with an election operation and extension maps

Conclusions - II

- Conditions relating local decisions to global behaviors
  - Strict-absorbency equivalent to robust correctness
  - Isotonicity implies optimality and visibility
  - Next-hop constrains usability
- Practical uses of the framework
  - Analysis of routing behaviors
  - Guidelines for the configuration of routing policies
  - Toward an automated management of routing

ありがとう