Appendix A: Bellman's Principle of Optimality for the Mapping Definition

Gao and Huang (2012) show that Bellman's Principle of optimality does not hold for the delayed, pre-trip, radio or no online information schemes with additive objective functions, and is valid for the perfect online information scheme. The following proposition shows that Bellman's Principle of optimality is valid when the information scheme covers trajectory information and the optimality criterion is minimum expected travel time (METT).

Proposition 1 *If a routing policy is defined as a mapping from state to decision, i.e., $\mu : \{j, t, EV\} \rightarrow k$, and the information scheme Q covers the trajectory information H, then any sub-policy of a routing policy with the METT is also with the METT.*

Proof.

We prove this proposition by contradiction.

Consider the problem of finding the optimal routing policy with METT from node-time pair (j, t) to the destination with an existing trajectory $H = \{(j_0, t_0), (j_1, t_1), \ldots, (j, t)\}$ and information $Q = H$. Let $e_\mu(j, t, EV)$ be the expected travel time of policy μ, $S_\mu(j, t, r)$ the support point travel time of policy μ in support point r, and $P_r(r|EV)$ the conditional probability of support point r given EV. By definition,

$$e_\mu(j, t, EV) = \sum_{r \in EV} S_\mu(j, t, r) \cdot P_r(r|EV) \tag{1}$$

Assume $\mu^\star$ is an optimal routing policy for this problem and the next node $k = \mu^\star(j, t, EV)$. Denote the i-th support point of the conditional marginal distribution of $\widetilde{C}_{jk,t}$ as $\tau^i_{jk,t}$. The corresponding trajectory at node k is H'_i with arrival time $t + \tau^i_{jk,t}$, and the corresponding sub-policy is $\lambda^\star_i$.

Assume by contradiction that $\mu^\star$ has a sub-policy $\lambda^\star_1$ for trajectory H'_1 that is not with METT, and the policy with METT for trajectory H'_1 is λ_1. Therefore we have $e_{\lambda_1}(k, t + \tau^1_{jk,t}, EV'_1) < e_{\lambda^\star_1}(k, t + \tau^1_{jk,t}, EV'_1)$, where EV'_1 is the event collection compatible with H'_1.

We can construct a trajectory-adaptive routing policy μ for trajectory H, such that it shares with $\mu^\star$ all the sub-policies $\lambda^\star_i$, for $i = 2, 3, \ldots,$ except that $\lambda^\star_1$ is replaced by λ_1. It is shown in the following equations that μ has a lower expected travel time than $\mu^\star$ for H. Note that $\tau^i_{jk,t}$ is shortened as τ^i, since it is clear that link (j,k) and time t are under discussion.

$$\begin{aligned}
e_\mu(j,t,EV) &= \sum_{r\in EV} S_\mu(j,t,r)\cdot P_r(r|EV) \\
&= \sum_i P(EV_i'|EV)\cdot(\tau^i + \sum_{r\in EV_i'} S_{\lambda_i}(k,t+\tau^i,r)\cdot P_r(r|EV_i')) \\
&= \sum_i P(EV_i'|EV)\cdot(\tau^i + e_{\lambda_i}(k,t+\tau^i,EV_i')) \\
&< \sum_i P(EV_i'|EV)\cdot(\tau^i + e_{\lambda_i^\star}(k,t+\tau^i,EV_i')) \\
&= e_{\mu^\star}(j,t,EV) \qquad (2)
\end{aligned}$$

Note that in the equations, EV is compatible with H and EV_i' is compatible with H_i. The first equality is by definition. The second equality is a re-arrangement conditional on the travel time on the next link τ^i and the next trajectory H_i'. The third equality calculates the unconditional expected travel time of a sub-policy λ_i. The fourth inequality is due to the contradiction assumption. The last equality can be derived following the same logics in the first three equalities, but in a reverse order.

This contradicts the assumption that $\mu^\star$ is with METT for trajectory H. Thus, all the sub-policies $\lambda_i^\star$ are with METT for the corresponding trajectories H_i'. **Q.E.D.**

The above proposition shows that trajectory information is a sufficient condition for Bellman's Principle of optimality. The following propositions shows that, if the information scheme covers more than the trajectory information, Bellman's Principle of optimality also holds.

Proposition 2 *If a routing policy is defined as a mapping from state to decision, i.e.,* $\mu : \{j,t,EV\} \rightarrow k$*, and the information scheme* Q *covers more than the trajectory information* H*, then any sub-policy of a routing policy with the METT is also with the METT.*

Proof.

When the information scheme Q covers more than the trajectory information H, the corresponding event collection set $EV(Q)$ is contained by the event collection set $EV(H)$ when $Q = H$ (Gao and Huang, 2012). Therefore, all the EV and EV' are the same as or finer than those in the equations in the proof of Proposition 1, and the proof can follow the same logic. **Q.E.D.**

An example can be the perfect information. The information scheme covers more than the trajectory information and Bellman's Principle of optimality is valid.

The next proposition is to show that the trajectory information is not a necessary condition for the validity of Bellman's Principle of optimality.

Proposition 3 *If a routing policy is defined as a mapping from state to decision, i.e.,* $\mu : \{j, t, EV\} \to k$*, and any sub-policy of a routing policy with the METT is also with the METT, then the information scheme* Q *does not necessarily cover the trajectory information* H*.*

Proof.

We prove this proposition by counterexample.

In the example of Figure 1, suppose the information covers only link (O, a) at time 0. Due to the dependencies among link travel times, that piece of information still allows travelers to identify which support point is realized and thus will result in the same optimal routing policies as trajectory information or even perfect information. Therefore, Bellman's Principle is valid, i.e., any sub-policy of an optimal routing policy is also optimal. **Q.E.D.**

Above is a counterexample where the information coverage is part of the trajectory H. Other examples can be constructed with information coverage not as part of the trajectory H, but with or without intersection with H. One example can be the global pre-trip information with departure time $0 : Q_i = A \times \{0\}$. Another example can be an information coverage on only link (O, a) at time 1, assuming that the travel time on link (O, a) at time 1 is 1 for support point C^1 and 2 for support point C^2, respectively.

Appendix B: An Illustrative Example of the Algorithm

In this section, it is illustrated how the algorithm works to find the optimal routing policy from origin O and departure time 0 in Figure 1.

At time $t = 3$ (static and deterministic):

For each node j except D:

Shortest path from node j to D will be calculated and stored in $\chi(j,3)$ as $\mu^\star(j,3)$.

For example, for node b, the shortest path to D is to take node D next and then $STOP$; for node a, the shortest path to D is to take node b next and then take the shortest path from node b to D, which is to take D and then $STOP$; and so on.

Table 1: Pure Policies at Each Node at Time $t = 3$ (Shortest Paths)

Node j	λ in $\chi(j,3)$	$\pi_\lambda(j,3)$	$L_\lambda(j,3,3)$	$S_\lambda(j,3,C^1)$	$S_\lambda(j,3,C^2)$
O	μ^*(O, 3)	a	μ^*(a,3)	2	2
a	μ^*(a, 3)	b	μ^*(b,3)	2	2
b	μ^*(b, 3)	D	STOP	1	1
c	μ^*(c, 3)	D	STOP	1	1

For $t = 2$ down to 0 (stochastic and time-dependent):

For each link (j,k):

find all possible arrival times at node k, t'_i, $1 \leq i \leq l$, assuming l is the total number of possible arrival times at node k

for each combination of indexes $(p_1, p_2, \ldots, p_l) \in \chi(k,t'_1) \times \chi(k,t'_2) \times \ldots \times \chi(k,t'_l)$ do

Construct a new routing policy λ as follows:

$\pi_\lambda(j,t) = k, L_\lambda(j,t,t'_i) = p_i, \forall 1 \leq i \leq l$

Dominance is checked and non-dominated policies will be stored in $\chi(j,t)$.

At time $t = 2$:

For example, for link (b,D), there are two possible arrival time at D: 3 and $2+M$, but since $\chi(D,3)$ and $\chi(D,2+M)$ both contain only one policy $STOP$, link (b,D) will generate one policy to $\chi(b,2)$, which is to take node D next and $STOP$ for both possible arrival times, and its support point travel time is 1 and M for C^1 and C^2, respectively; for link (b,c), there is only one possible arrival time at node c: 4, and $\chi(c,4)$ contains one policy $\mu^\star(c,3)$, thus link (b,c) will generate one policy to $\chi(b,2)$, which is to take node c next and then take node D and $STOP$, and its support point travel time is 3 for both support points. The two policies for node b are both pure and thus stored in $\chi(b,2)$.

At time $t = 1$:

For example, similarly to at time $t = 2$, for $\chi(b,1)$, both links (b,D) and (b,c) will generate one policy, but the one generated by the former will

Table 2: Pure Policies at Each Node at Time $t = 2$

Node j	λ in $\chi(j, 2)$	$\pi_\lambda(j, 2)$	t'_i	$L_\lambda(j, 2, t'_i)$	$S_\lambda(j, 2, C^1)$	$S_\lambda(j, 2, C^2)$
O	λ(O, 2)	a	2	λ(a, 2)	2	2
a	λ(a, 2)	b	3	μ^*(b, 3)	2	2
b	λ_1(b, 2)	D	3	STOP	1	M
			2+M	STOP		
	λ_2(b, 2)	c	4	μ^*(c,3)	3	3
c	λ(c, 2)	D	3	STOP	1	1

be discarded because it is dominated by the latter (comparison of support point travel times: $3 = 3$ for C^1 and $M > 3$ for C^2), and thus $\chi(b, 1)$ will contain only one policy, which is to take node c next and then take node D and $STOP$; for link (a, b), there are two possible arrival times at node b: 2 and 3, $\chi(b, 2)$ contains two pure policies and $\chi(b, 3)$ one, thus link (a, b) will generate two policies to $\chi(a, 1)$: one is to take node b next and take $\lambda_1(b, 2)$ if arriving at node b at time 2 and take $\mu^\star(b, 3)$ if arriving at time 3, whose support point travel time is 2 for C^1 and 3 for C^2, and the other is to take node b next and take $\lambda_2(b, 2)$ if arriving at node b at time 2 and take $\mu^\star(b, 3)$ if arriving at time 3, whose support point travel time is 4 for C^1 and 3 for C^2. The latter is dominated by the former and so discarded, thus $\chi(a, 1)$ contains only one policy.

Note that, when calculating the support point travel times for the policies, the phantom travel times are not included. For example, when calculating the support point travel time for $\lambda(a, 1)$, the next node is $\pi_\lambda(a, 1) = b$, and it has two possible arrival times at node b: 1) for arrival time at 2, the sub-policy is $\lambda_1(b, 2)$, whose travel time in support point C^2, M, is not compatible with the arrival time and thus not included in the calculation of travel time for $\lambda(a, 1)$; and 2) similarly, for arrival time at 3, the sub-policy is $\mu^\star(b, 3)$, whose travel time in support point C^1, 3, is not compatible with the arrival time and thus not included in the calculation of travel time for $\lambda(a, 1)$. The same applies to other policies (e.g., $\lambda(O, 0)$ at time $t = 0$): the phantom travel time of a sub-policy will not contribute to the calculation of the travel time for the policy because it is not compatible with the arrival time.

At time $t = 0$:

For link (a, b), there are two possible arrival times at node b: 1 and 2, $\chi(b, 1)$ contains one policy and $\chi(b, 2)$ contains two, thus link (a, b) will generate two policies to $\chi(a, 0)$: one is to take node b next and take $\lambda(b, 1)$

Table 3: Pure Policies at Each Node at Time $t = 1$

Node j	λ in $\chi(j, 1)$	$\pi_\lambda(j, 1)$	t'_i	$L_\lambda(j, 1, t'_i)$	$S_\lambda(j, 1, C^1)$	$S_\lambda(j, 1, C^2)$
O	λ(O, 1)	a	1	λ(a, 1)	2	3
a	λ(a, 1)	b	2	λ_1(b, 2)	2	3
			3	μ^*(b, 3)		
b	λ(b, 1)	c	3	μ^*(c, 3)	3	3
c	λ(c, 1)	D	2	STOP	1	1

if arriving at time 1 and take $\lambda_1(b, 2)$ if arriving at time 2, whose support point travel time is 4 for C^1 and $2 + M$ for C^2, and the other is to take node b next and take $\lambda(b, 1)$ if arriving at time 1 and take $\lambda_2(b, 2)$ if arriving at time 2, whose support point travel time is 4 for C^1 and 5 for C^2. The former is dominated by the latter and so discarded, thus $\chi(a, 0)$ contains only one policy, denoted as $\lambda(a, 0)$. For link (O, a), there are two possible arrival times at node a: 1 and 0, and, since both $\chi(a, 1)$ and $\chi(a, 0)$ contains one policy, it will generate one policy for $\chi(O, 0)$.

Table 4: Pure Policies at Each Node at Time $t = 0$

Node j	λ in $\chi(j, 0)$	$\pi_\lambda(j, 0)$	t'_i	$L_\lambda(j, 0, t'_i)$	$S_\lambda(j, 0, C^1)$	$S_\lambda(j, 0, C^2)$
O	λ(O, 0)	a	1	λ(a, 1)	3	5
			0	λ(a, 0)		
a	λ(a, 0)	b	1	λ_1(b, 1)	4	5
			2	λ_2(b, 2)		
b	λ(b, 0)	D	1	STOP)	1	1
c	λ(c, 0)	D	2	STOP	1	1

For node-time pair $(O, 0)$, $\chi(O, 0)$ contains only one policy, thus it is optimal.

References

Gao, S. and Huang, H. (2012). Real-time traveler information for optimal adaptive routing in stochastic time-dependent networks, *Transportation Research Part C* **21**(1): 196–213.