TacTex'13: A Champion Adaptive Power Trading Agent

Daniel Urieli Peter Stone

Department of Computer Science
The University of Texas at Austin
{urieli,pstone}@cs.utexas.edu

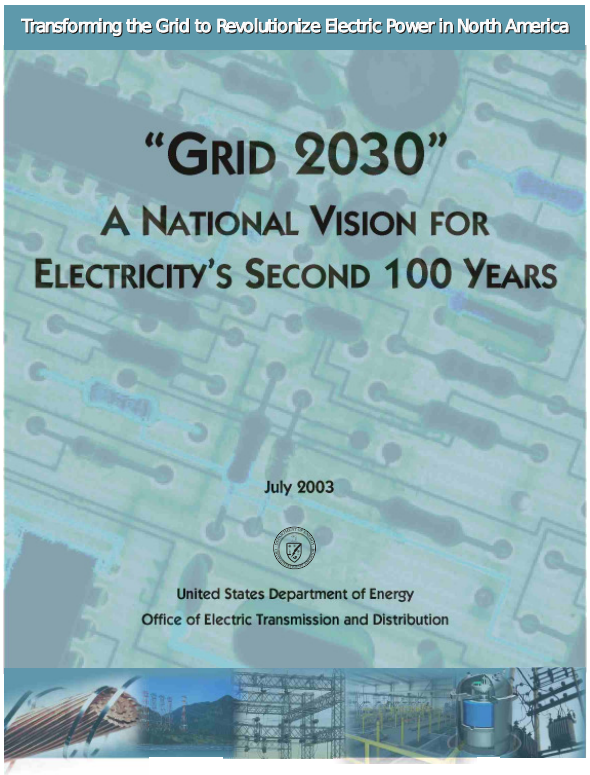
AAAI 2014

UT Austin Learning Agents Research Group

The Smart Grid Vision

- "Grid 2030" - vision for a smart-grid
  - Major challenge: aligning supply-demand in the presence of renewable, intermittent generation
- AI: a main building block
- Smart-grid: new challenges for AI

[Ramchurn et. al 2012]



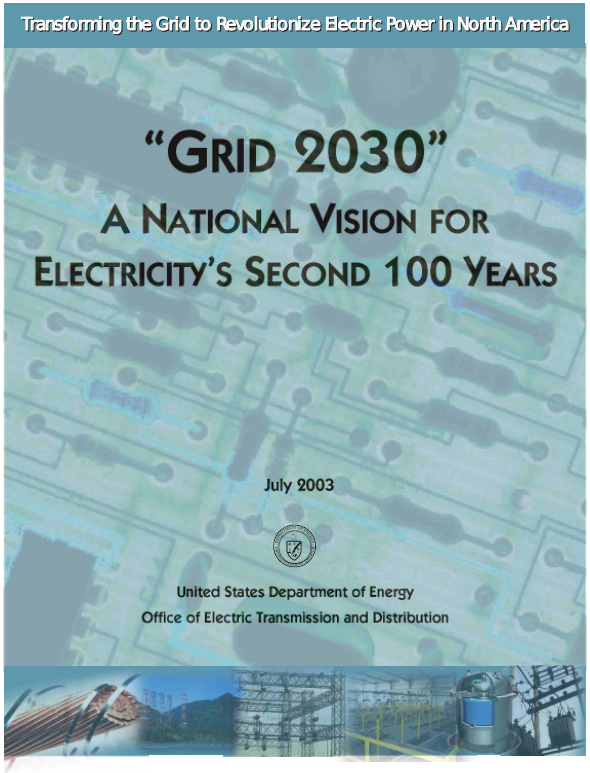
UT Austin Learning Agents Research Group

The Smart Grid Vision

- "Grid 2030" - vision for a smart-grid
  - Major challenge: aligning supply-demand in the presence of renewable, intermittent generation

- AI: a main building block
- Smart-grid: new challenges for AI

[Ramchurn et. al 2012]



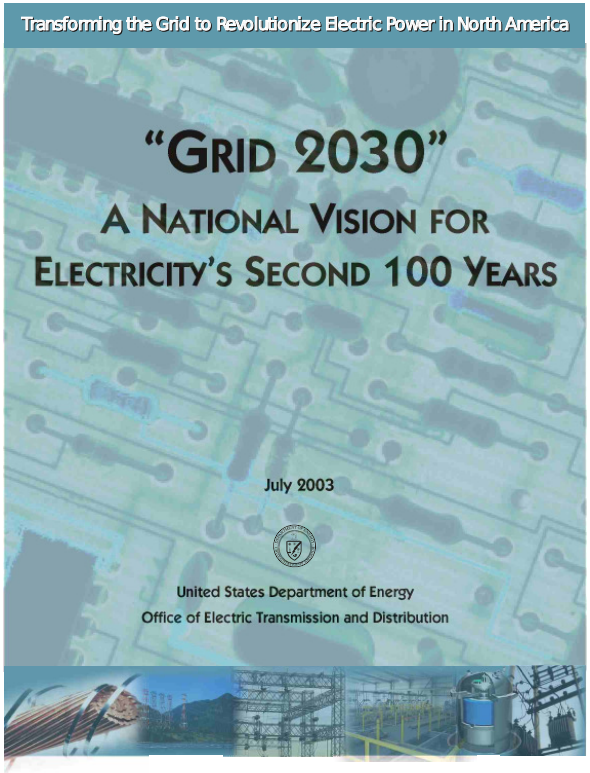
UT Austin Learning Agents Research Group

The Smart Grid Vision

- "Grid 2030" - vision for a smart-grid
  - Major challenge: aligning supply-demand in the presence of renewable, intermittent generation

- AI: a main building block
- Smart-grid: new challenges for AI

[Ramchurn et. al 2012]



UT Austin Learning Agents Research Group

The Power Trading Agent Competition (Power TAC)

- Grid 2030 milestone:

  > "Customer participation in power markets through demand-side management and distributed generation"

- Power TAC (Power Trading Agent Competition)
  - Uses a rich smart grid simulation platform
  - Focuses on retail power markets structure and operation
  - Competitors: autonomous broker agents

UT Austin Learning Agents Research Group

The Power Trading Agent Competition (Power TAC)

- Grid 2030 milestone:

  > "Customer participation in power markets through demand-side management and distributed generation"

- Power TAC (Power Trading Agent Competition)
  - Uses a rich smart grid simulation platform
  - Focuses on retail power markets structure and operation
  - Competitors: autonomous broker agents

UT Austin Learning Agents Research Group

The Power Trading Agent Competition (Power TAC)

- Grid 2030 milestone:

> "Customer participation in power markets through demand-side management and distributed generation"

- Power TAC (Power Trading Agent Competition)
  - Uses a rich smart grid simulation platform
  - Focuses on retail power markets structure and operation
  - Competitors: autonomous broker agents

UT Austin Learning Agents Research Group

Approach

Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Approach

Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Approach

Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Approach

Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Approach

Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Approach

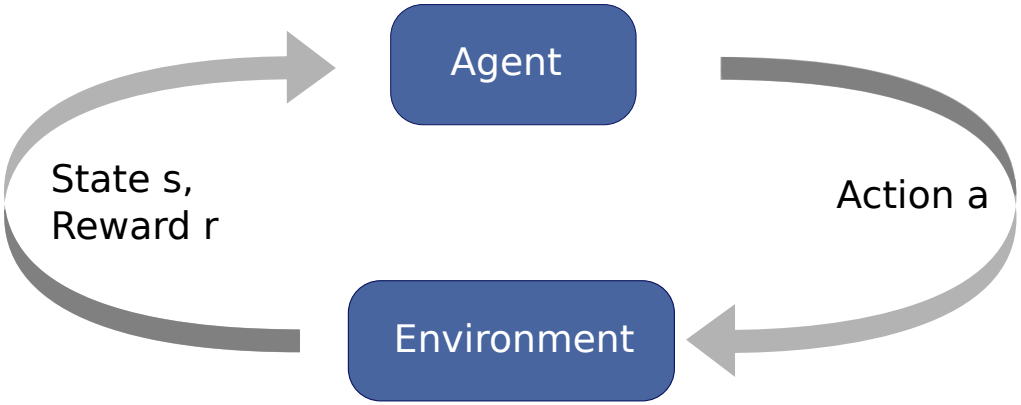
Application domain: autonomous energy-trading

- In this domain:
  - An agent is deployed into an unknown environment
  - The agent is expected to make robust, real-time decisions
  - Environment is realistic $\Longrightarrow$ complex
- To perform robustly, agent need to:
  - Learn
  - Predict
  - Plan
  - Adapt
- A natural approach: Reinforcement Learning

UT Austin Learning Agents Research Group

Reinforcement Learning in the Smart Grid

- Reinforcement Learning (RL):

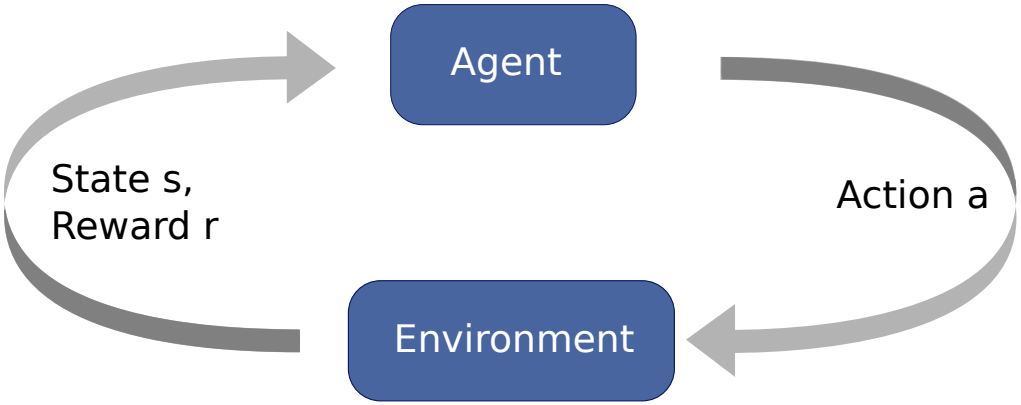


- **Our domains require from an RL agent:**
  - **Sample-efficiency**
  - **Computationally-efficiency**
  - **Handle high-dimensional continuous state**
  - **Handle continuous-actions and/or delayed-actions**
  - **Handle possible non-stationarity**
- Combination that was not addressed by past RL algorithms

UT Austin Learning Agents Research Group

Reinforcement Learning in the Smart Grid

- Reinforcement Learning (RL):

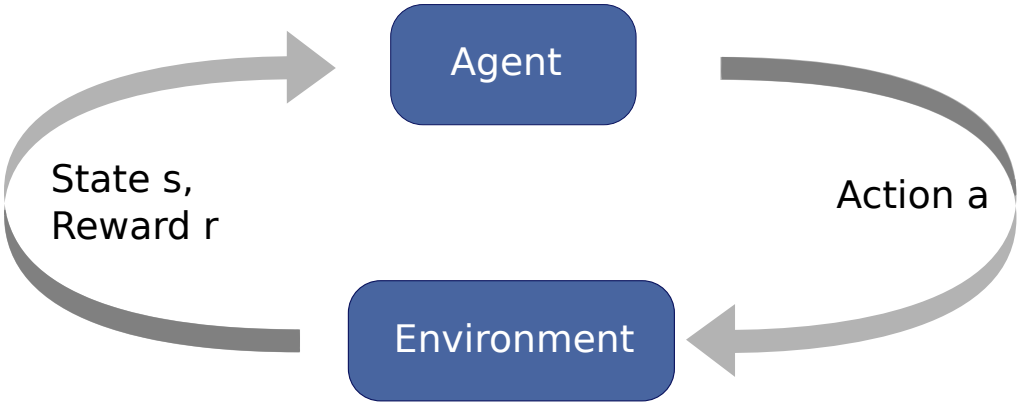


- **Our domains require from an RL agent:**
  - **Sample-efficiency**
  - **Computationally-efficiency**
  - **Handle high-dimensional continuous state**
  - **Handle continuous-actions and/or delayed-actions**
  - **Handle possible non-stationarity**
- Combination that was not addressed by past RL algorithms

UT Austin Learning Agents Research Group

Reinforcement Learning in the Smart Grid

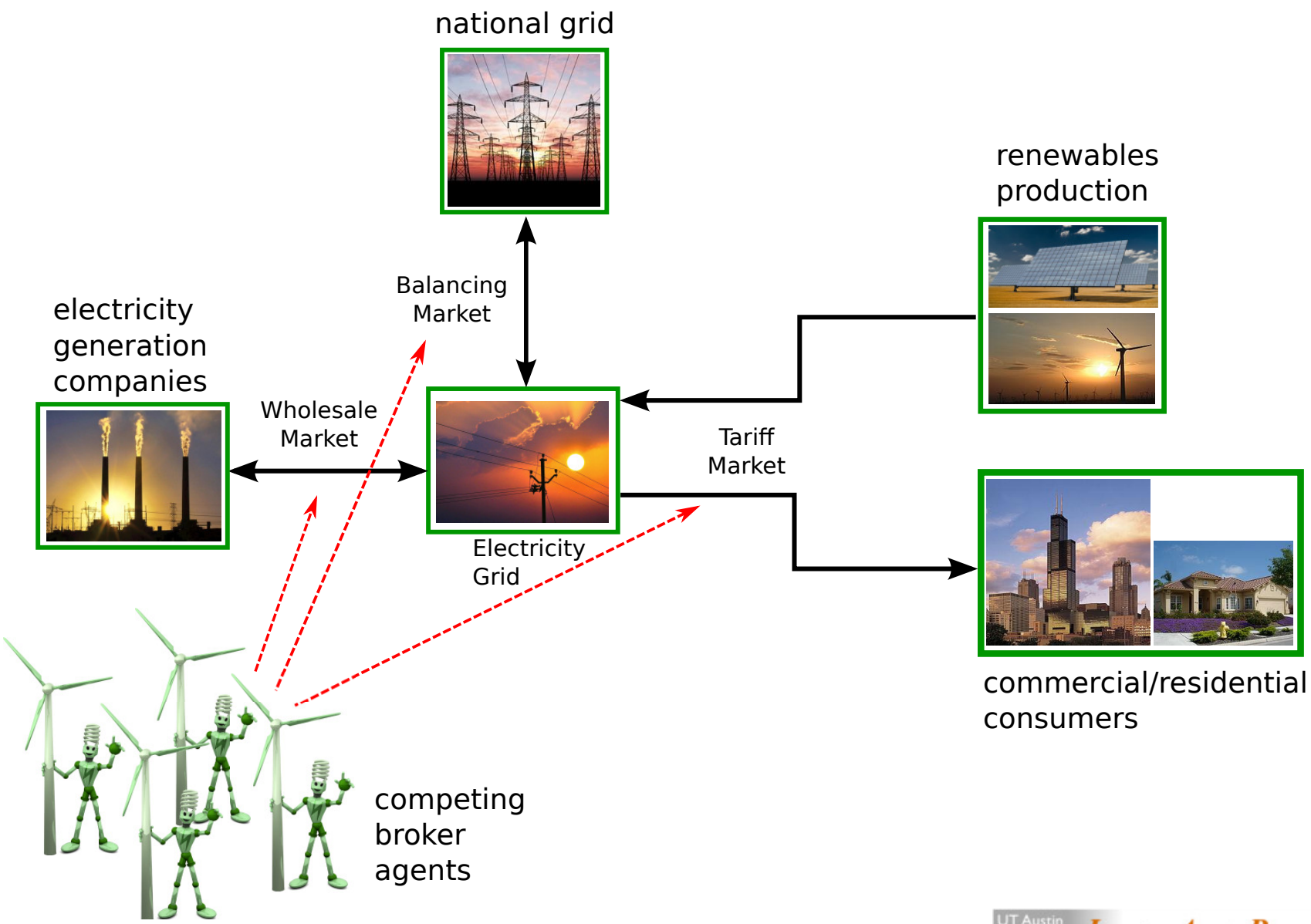
- Reinforcement Learning (RL):



- **Our domains require from an RL agent:**
  - **Sample-efficiency**
  - **Computationally-efficiency**
  - **Handle high-dimensional continuous state**
  - **Handle continuous-actions and/or delayed-actions**
  - **Handle possible non-stationarity**
- Combination that was not addressed by past RL algorithms

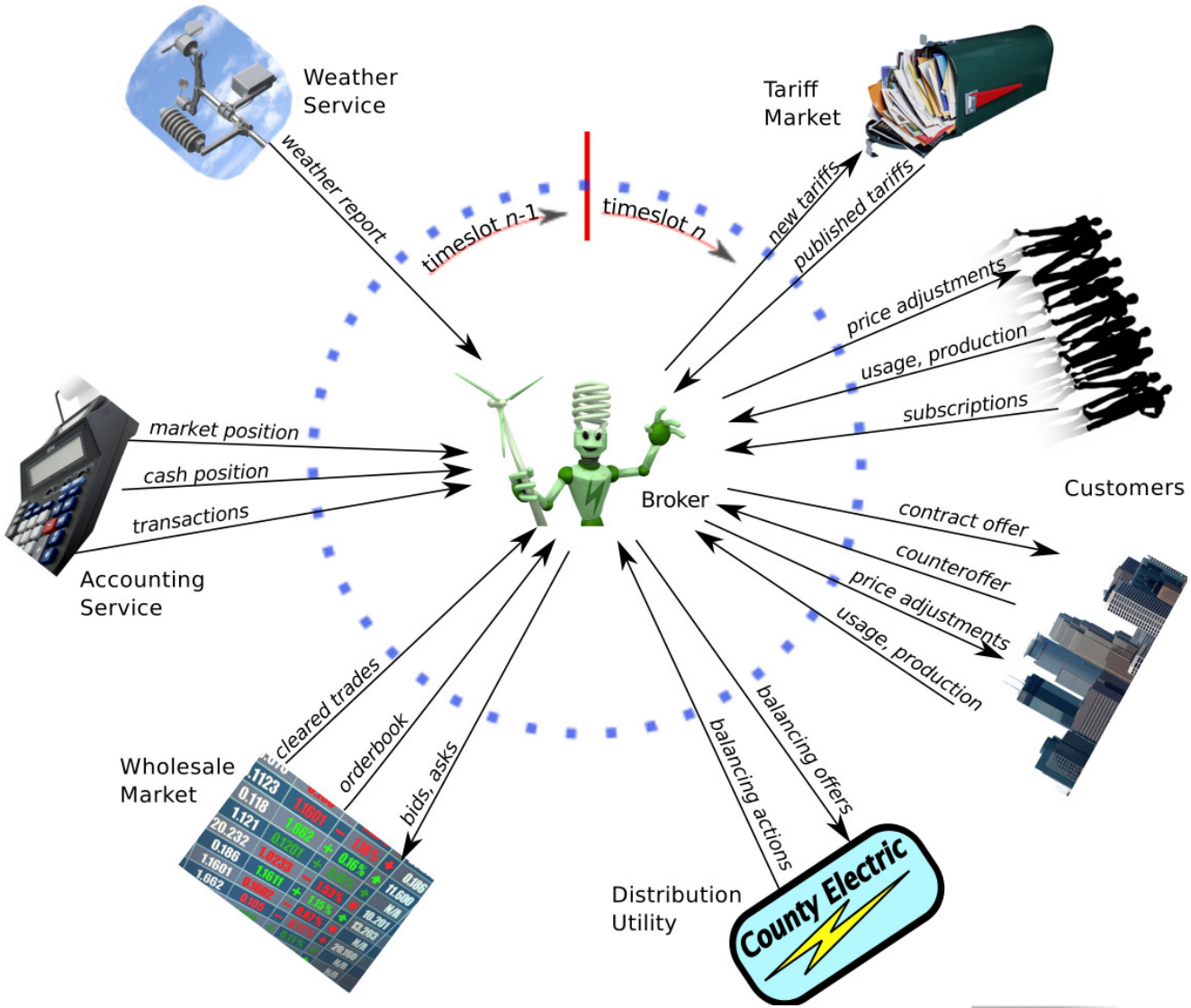
UT Austin Learning Agents Research Group

Power TAC: Game Description



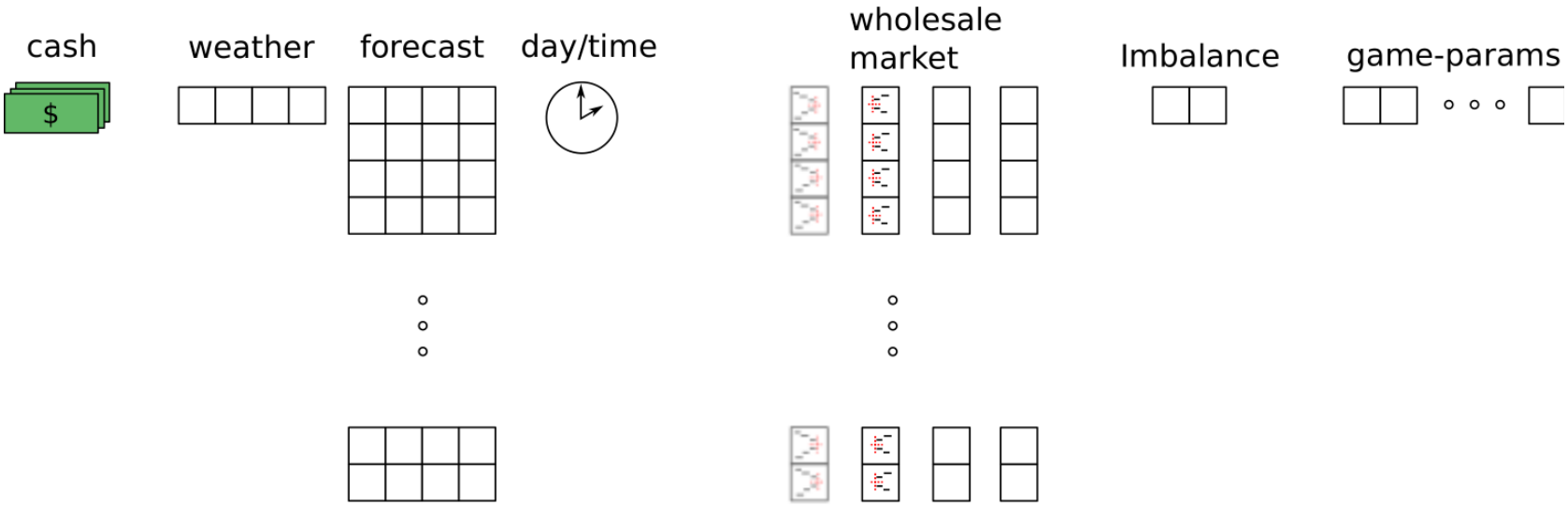
UT Austin Learning Agents Research Group

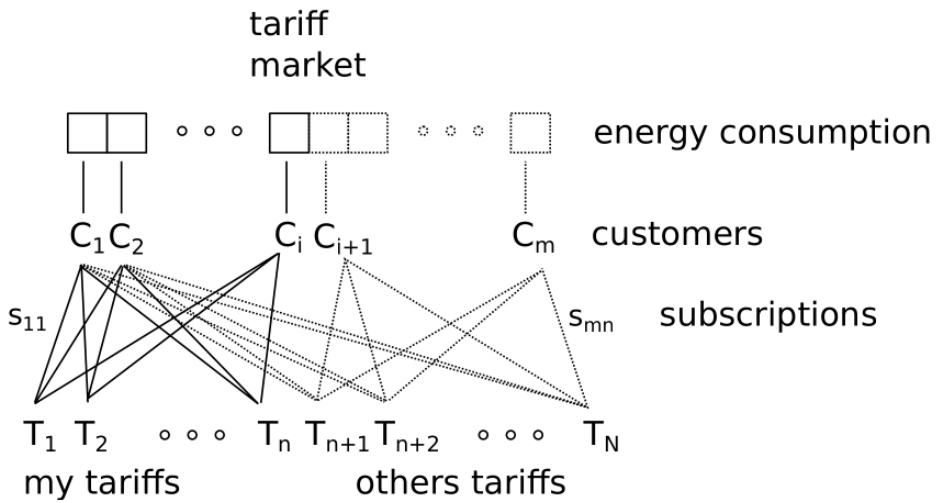
Power TAC: Broker Operation Cycle



UT Austin Learning Agents Research Group

Power TAC Game State





UT Austin Learning Agents Research Group

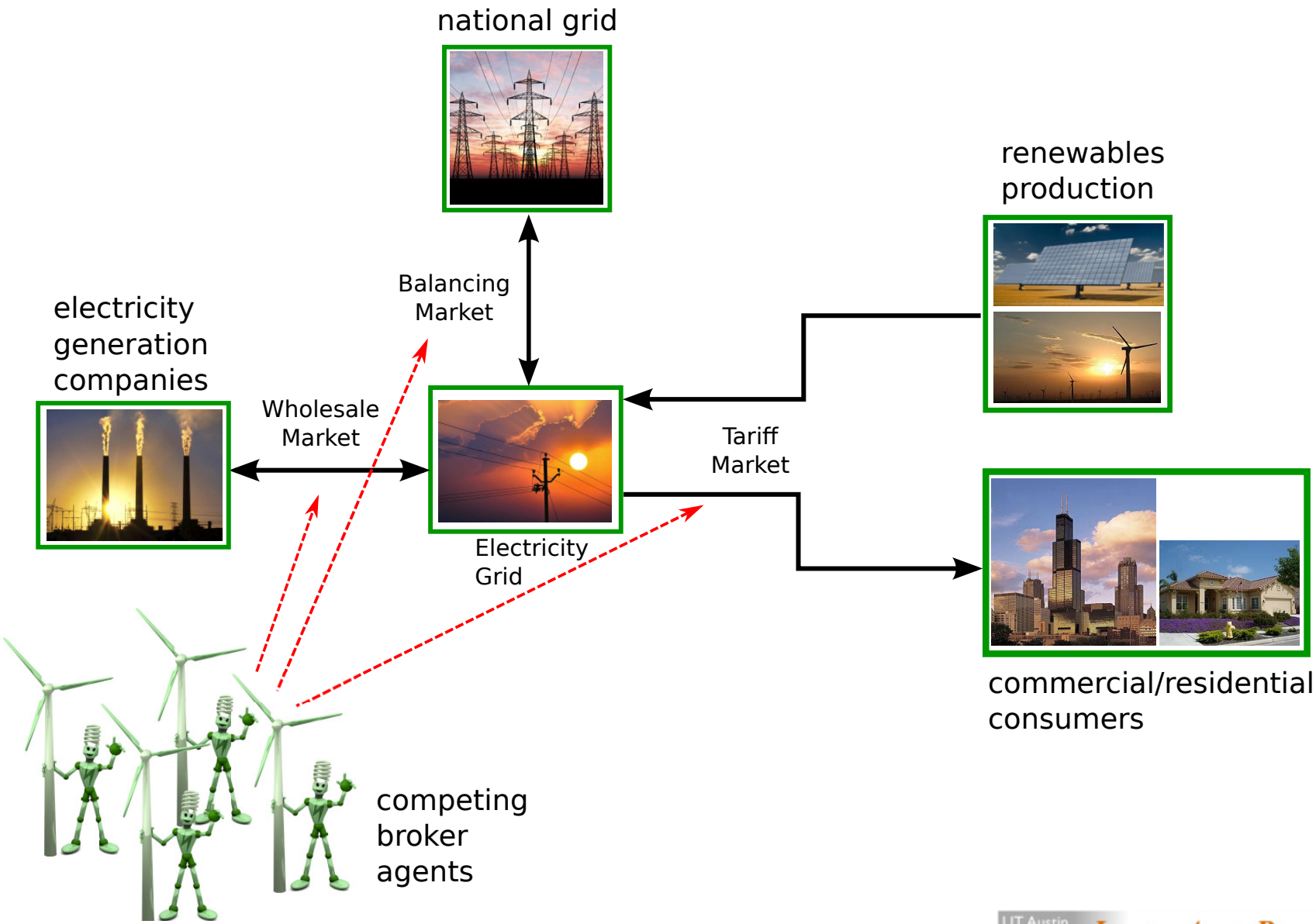
Power TAC 2013 Competition Results

- Our agent, TacTex'13, won the Power TAC 2013 finals:

Broker	7-broker	4-broker	2-broker	Total (not normalized)
TacTex	-705248	13493825	17853189	30641766
cwiBroker	647400	12197772	13476434	26321606
MLLBroker	8533	3305131	9482400	12796064
CrocodileAgent	-361939	1592764	7105236	8336061
AstonTAC	345300	5977354	5484780	11807435
Mertacor	-621040	1279380	4919087	5577427
INAOEBroker02	-76112159	-497131383	-70255037	-643498580

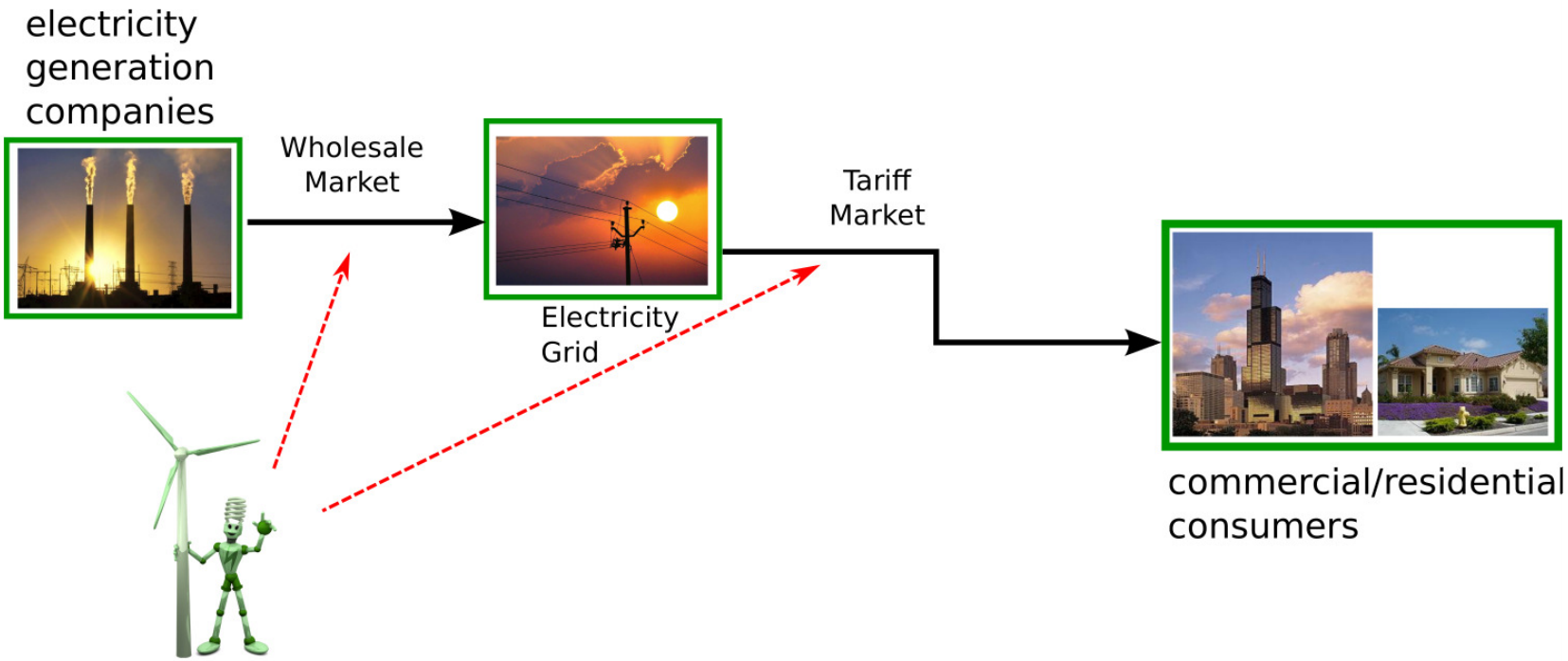
TacTex'13: Approach

TacTex'13: Approach

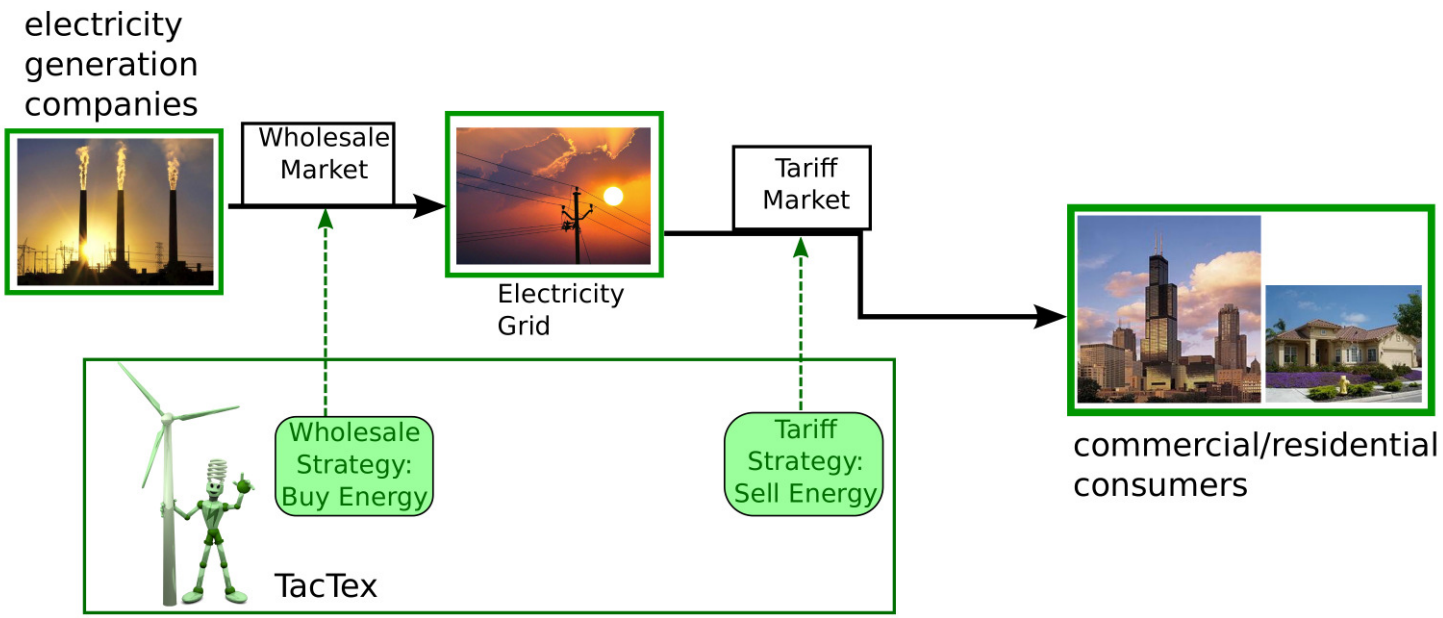


UT Austin Learning Agents Research Group

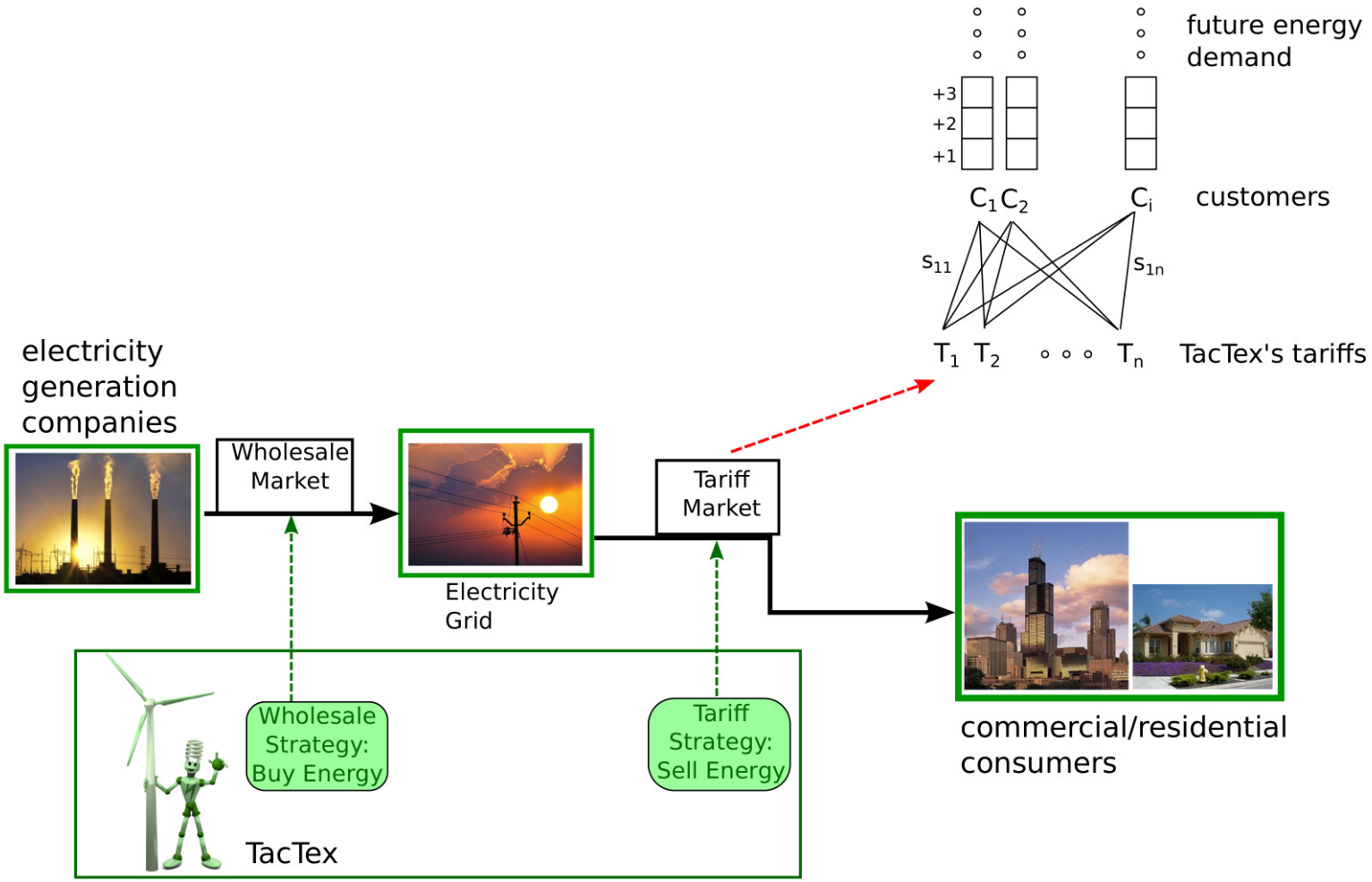
TacTex'13: Approach



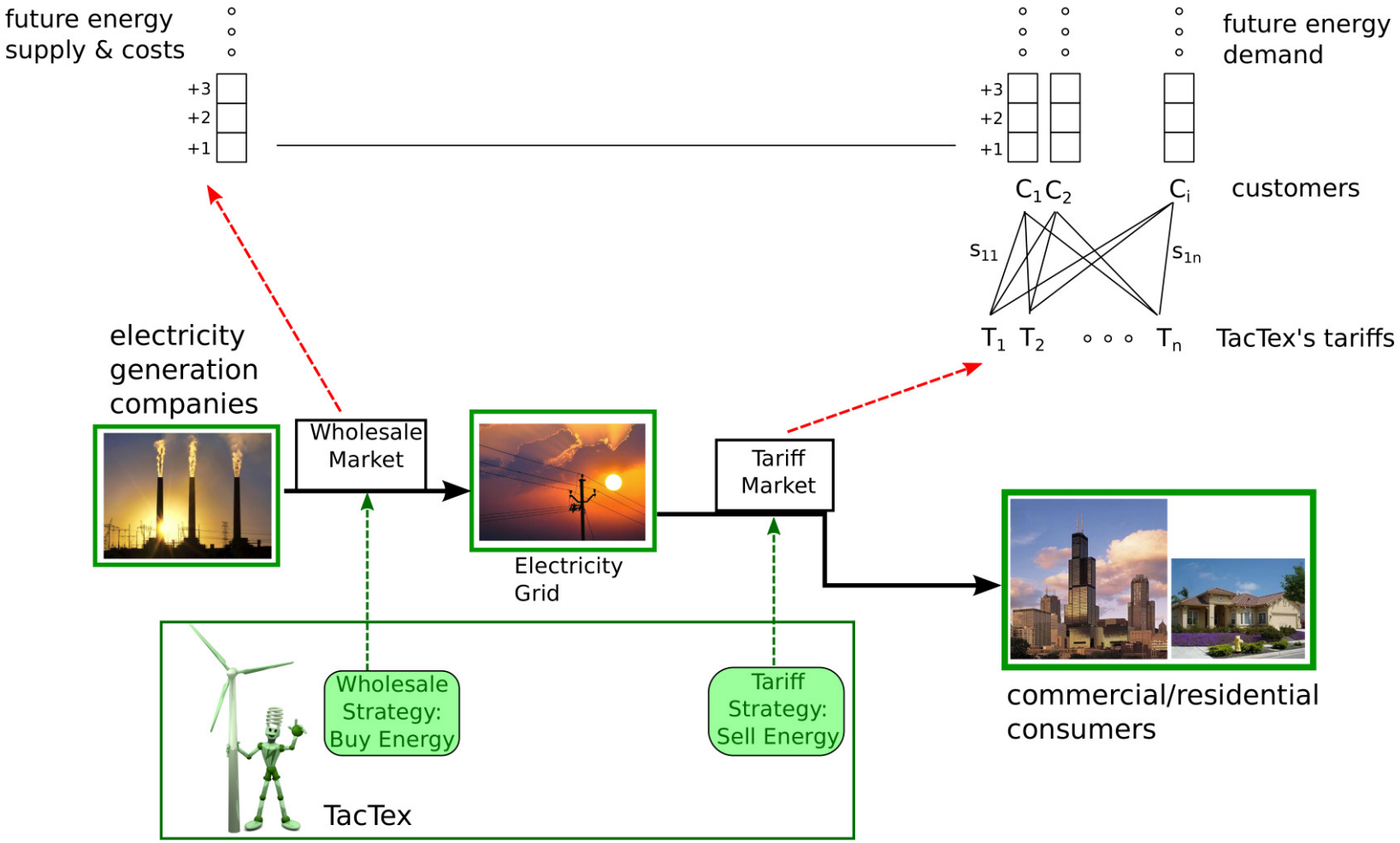
TacTex'13: Approach



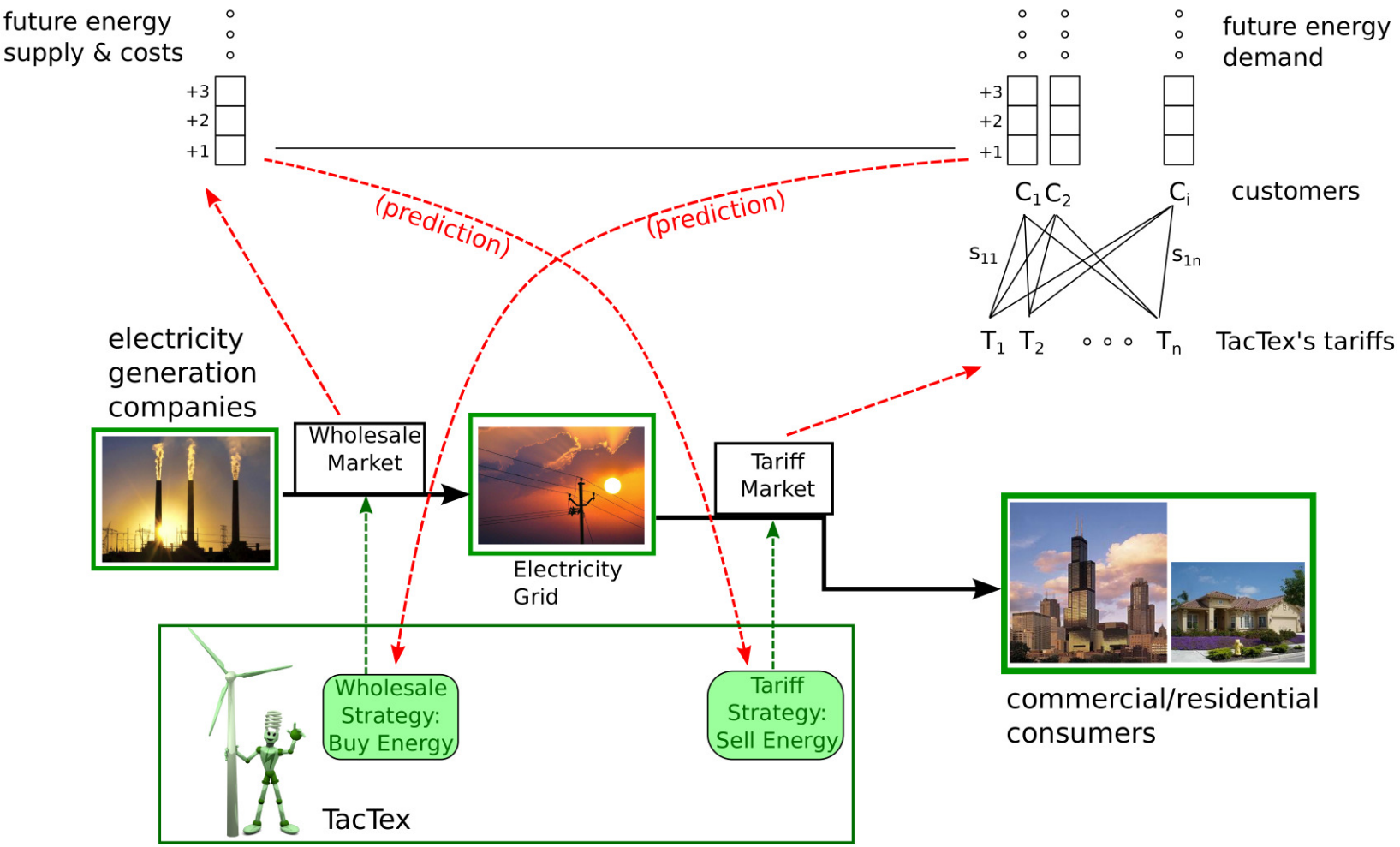
TacTex'13: Approach



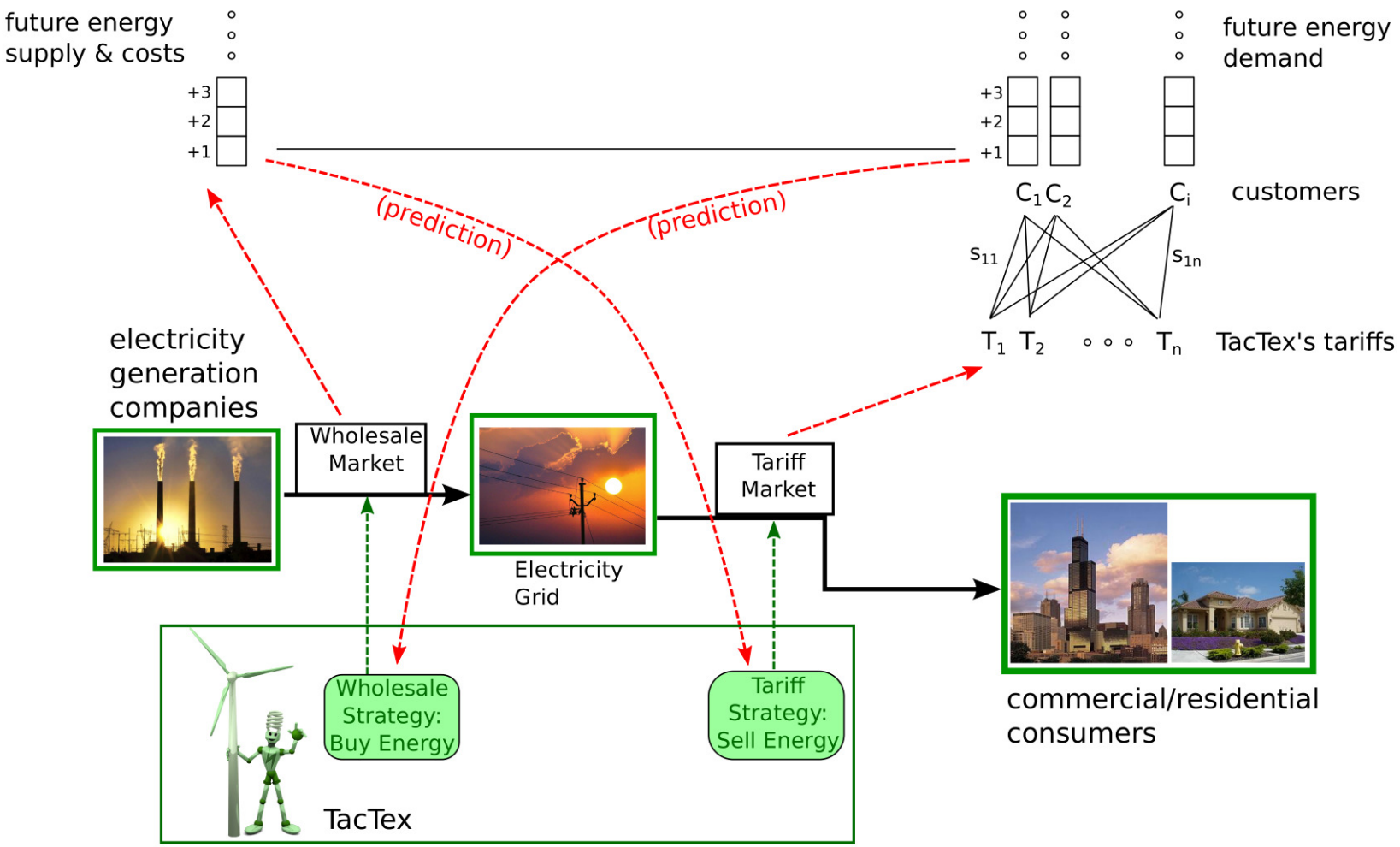
TacTex'13: Approach



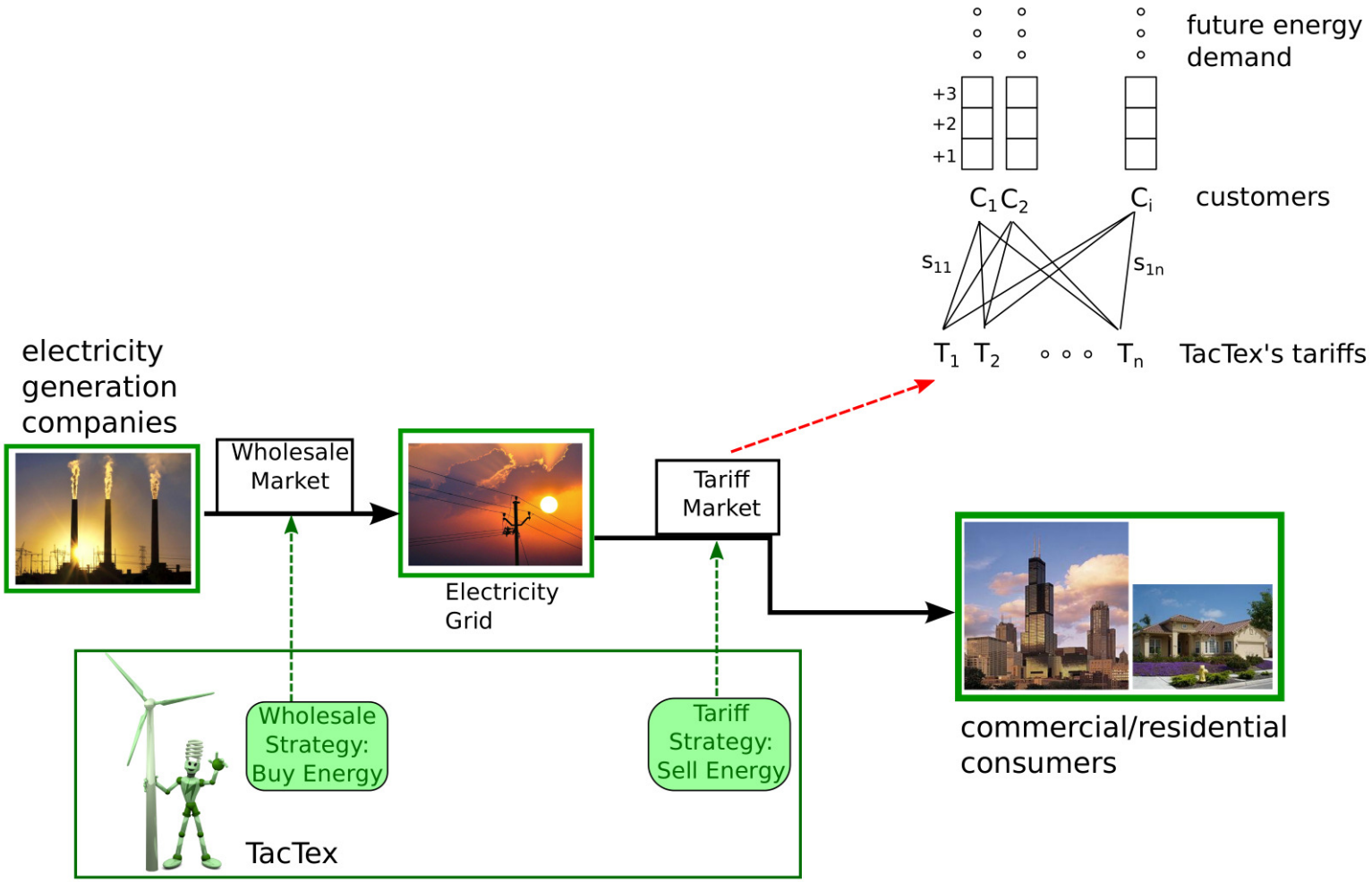
TacTex'13: Approach



TacTex'13: Tariff Market Strategy



TacTex'13: Tariff Market Strategy



Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Decision Making in the Tariff Market

- Available actions: tariff publications
- Tariff: contract for selling/buying energy
  - E.g.: [type=consumption, rates=(rate1, rate2,...), signup-fee=none,...]
- Rate: energy prices per time and/or quantity
  - Rate types: fixed, time-of-use (TOU), real-time (RT)...
  - Fixed: [fixed=true, price=7cent/kWh]
  - TOU: [(time1=Mon-Fri 7am-6pm, price=8cent/kWh), (time2=Sat, ...), ...]
  - RT: [expected/min/max-price=7/5/8 cent/kWh, rate-notice=3 hours...]
- Customers subscribe to tariffs they find attractive
  - Cheap, minimizes inconvenience...

Challenge: what tariffs should a broker publish?

UT Austin Learning Agents Research Group

Tariff Market Approach

- TacTex uses a utility-based approach
  - Optimizes long-term utility (= profits)
- Core computation:
  "if I publish tariff t, how would it affect my long-term utility?"

Tariff Market Approach

- TacTex uses a utility-based approach
  - Optimizes long-term utility (= profits)
- Core computation:
  "if I publish tariff t, how would it affect my long-term utility?"

Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence

- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

UT Austin Learning Agents Research Group

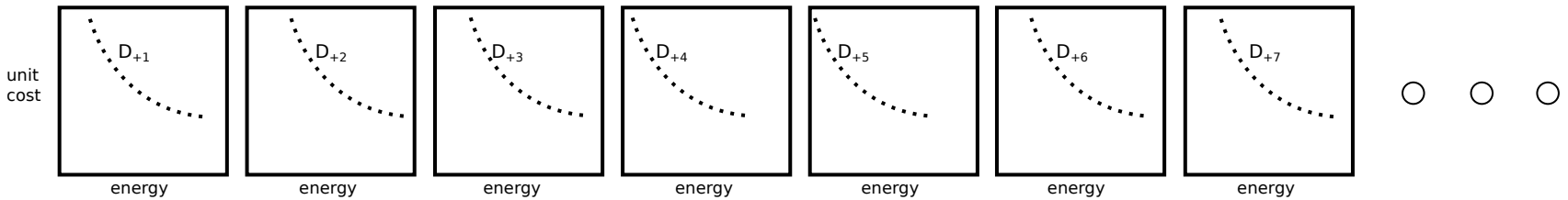
Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence
- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

UT Austin Learning Agents Research Group

Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence

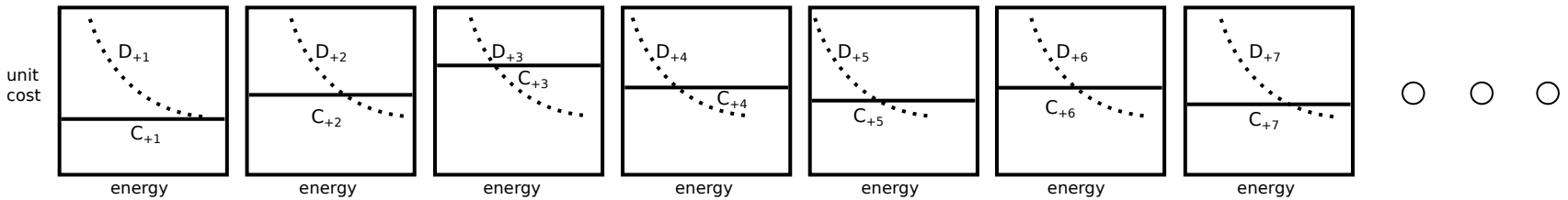


- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

UT Austin Learning Agents Research Group

Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence

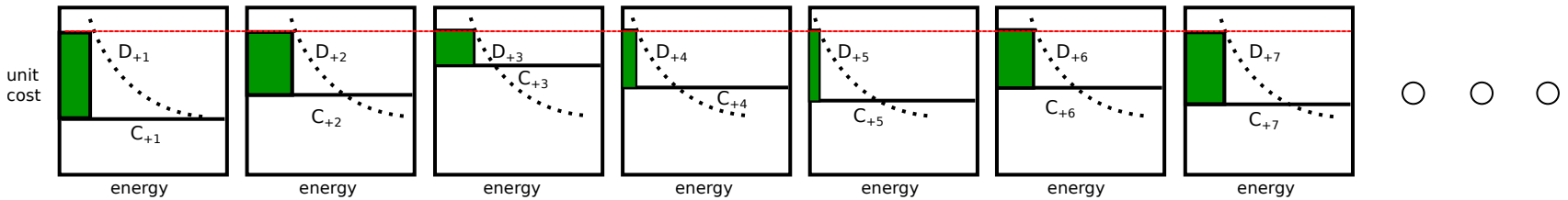


- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

UT Austin Learning Agents Research Group

Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence



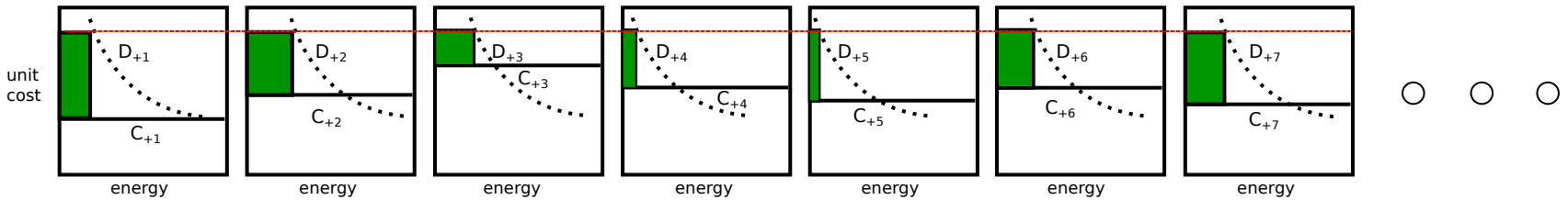
- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

Publish tariff if expected to increase utility

UT Austin Learning Agents Research Group

Tariff Utility Estimation

- Considering only fixed-rate tariffs
  - More attractive to customers
  - Optimizing one future price instead of a sequence

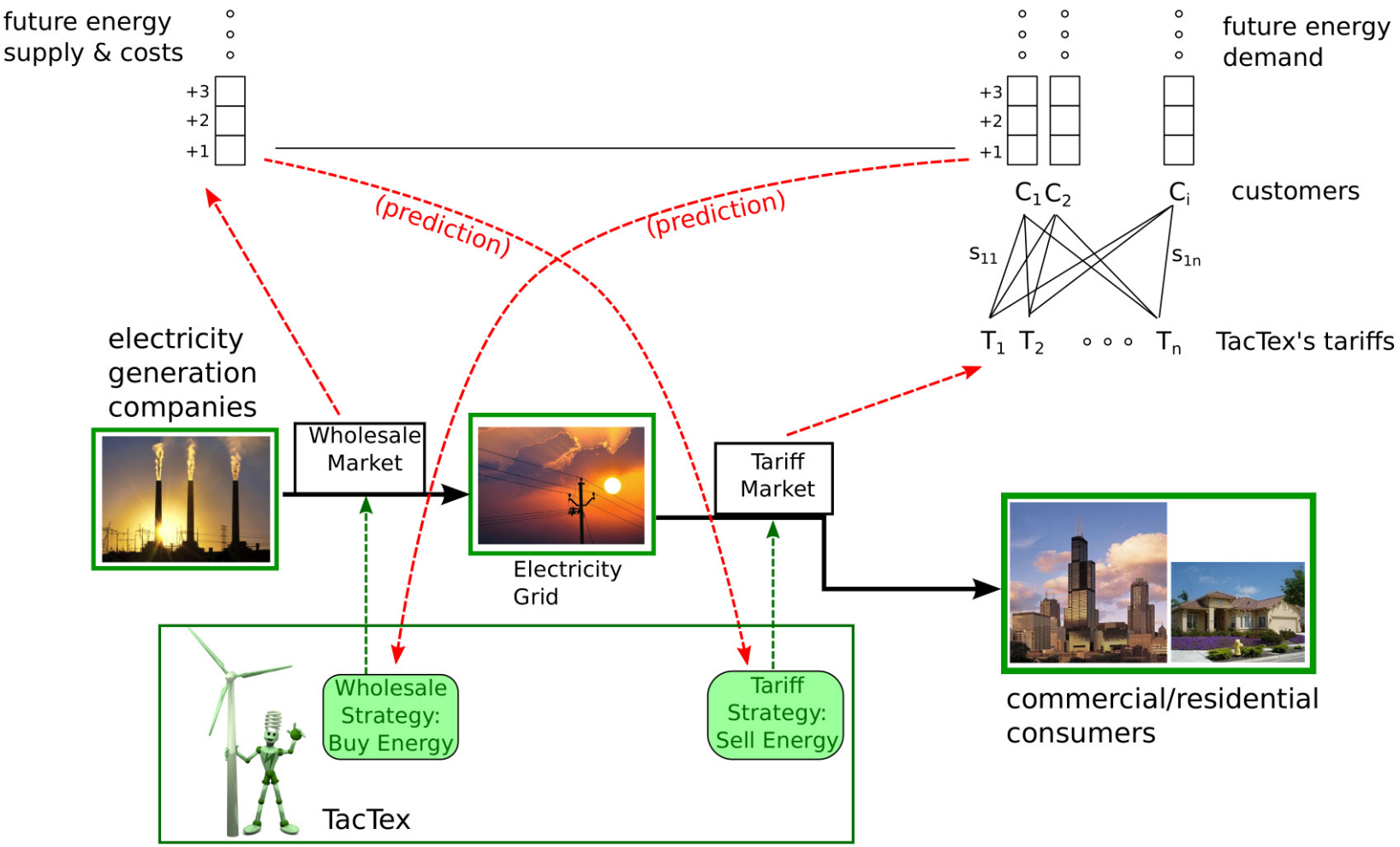


- Estimate future customers demand
- Estimate future wholesale costs
- Select price that maximizes profits

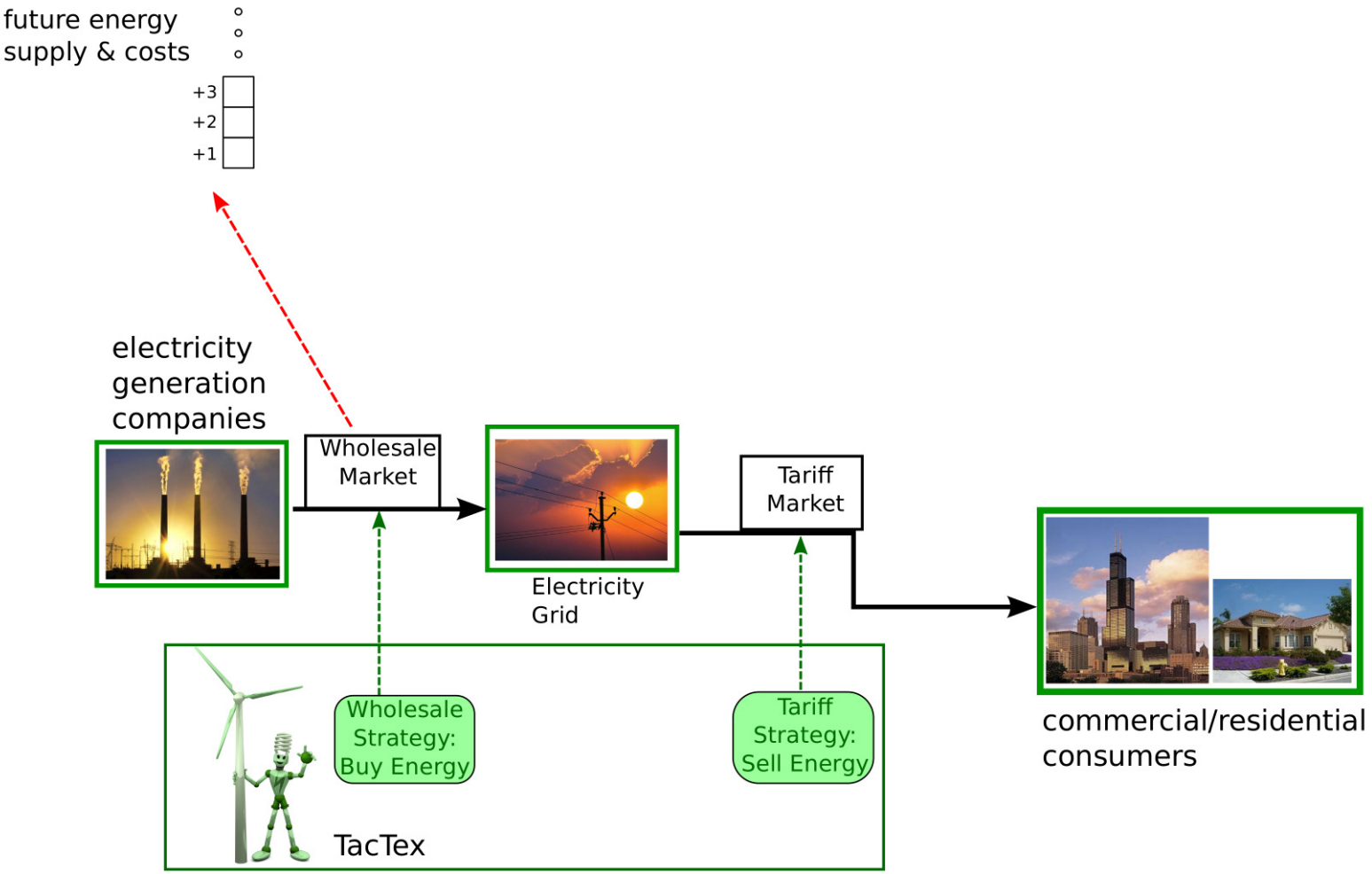
Publish tariff if expected to increase utility

UT Austin Learning Agents Research Group

TacTex'13: Wholesale Market Strategy



TacTex'13: Wholesale Market Strategy



Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:

- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Decision Making in the Wholesale Market

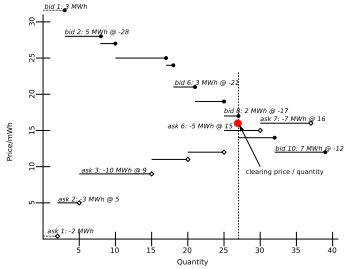
- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:
- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:



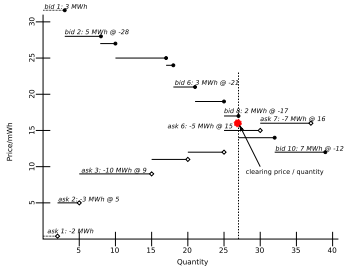
- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:



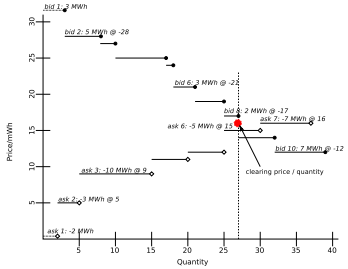
- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:



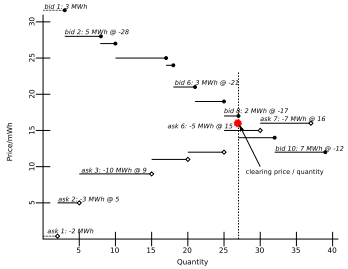
- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Decision Making in the Wholesale Market

- Available actions: bid submissions
  - Bid: [needed-amount=2mWh, limit=25$/mWh, when=5pm]
- Bids cleared in a double auction:



- Day ahead market $\Longrightarrow$ 24 auctions for each timeslot
- Need to:
  - Buy energy cheaply
  - Avoid imbalance costs $\Longrightarrow$ buy all needed energy

Challenge: what bidding strategy to use?

UT Austin Learning Agents Research Group

Wholesale Market Strategy

- Per timeslot: estimate future demand
- Minimize cost for satisfying this demand
- Online RL bidding algorithm:

UT Austin Learning Agents Research Group

Wholesale Market Strategy

- Per timeslot: estimate future demand
- Minimize cost for satisfying this demand
- Online RL bidding algorithm:

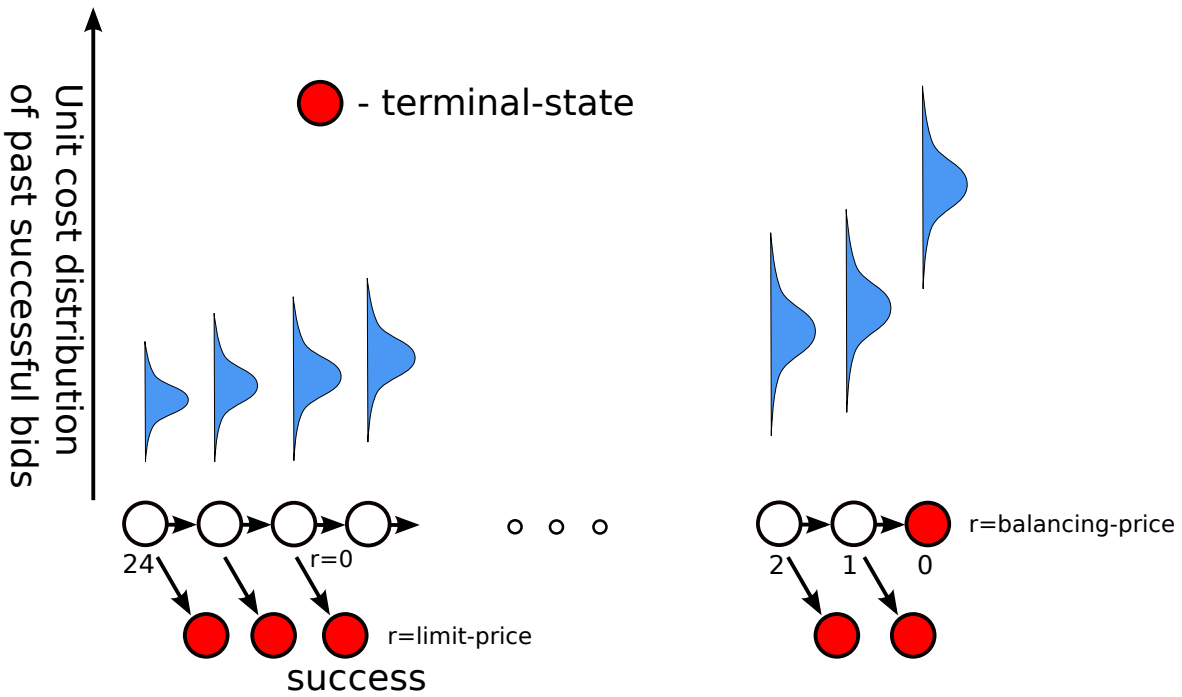
UT Austin Learning Agents Research Group

Wholesale Market Strategy

- Per timeslot: estimate future demand
- Minimize cost for satisfying this demand
- Online RL bidding algorithm:

Wholesale Market Strategy

- Per timeslot: estimate future demand
- Minimize cost for satisfying this demand
- Online RL bidding algorithm:



MDP States:
$\{0, 1, \ldots, 24, \textit{success}\}$

MDP Actions:
limit-price $\in \mathbb{R}$

UT Austin Learning Agents Research Group

Controlled Experiments - Ablation Analysis

- Round-Robin 2-agent tournament between:
  - B: baseline agent
  - U1: adding tariff-market strategy
  - U9_MDP: adding wholesale-market strategy
  - U9_MDP_LWR: adding LWR customer prediction
- Each pair played 200 games with similar conditions

	B	U1	U9_MDP
U9_MDP_LWR	1278.3 (43.2)	708.9 (35.6)	34.2 (23.2)
U9_MDP	966.4 (40.5)	592.6 (22.2)	
U1	547.4 (27.7))		

UT Austin Learning Agents Research Group

Ablation Analysis Using Available Finalist Agents

- 4-agent games using 3 available finalist agents

Broker	Cash
cwiBroker	340.9 (8.4)
Mertacor	-276.2 (40.2)
CrocodileAgent	-287.1 (14.5)
B	**-334.6 (8.0)**

Broker	Cash
cwiBroker	315.4 (9.3)
U1	**135.3 (12.3)**
CrocodileAgent	-372.1 (17.0)
Mertacor	-485.5 (28.1)

Broker	Cash
U9_MDP	**389.9 (13.3)**
cwiBroker	138.3 (8.7)
CrocodileAgent	-333.3 (17.0)
Mertacor	-494.1 (29.6)

Broker	Cash
U9_MDP_LWR	**350.8 (13.3)**
cwiBroker	132.4 (9.0)
CrocodileAgent	-336.9 (17.3)
Mertacor	-566.1 (26.8)

- Tariff and Wholesale strategies improve performance
- LWR customer prediction reduces performance
  - Should relax LWR's extrapolation assumptions?

UT Austin Learning Agents Research Group

Ablation Analysis Using Available Finalist Agents

- 4-agent games using 3 available finalist agents

Broker	Cash
cwiBroker	340.9 (8.4)
Mertacor	-276.2 (40.2)
CrocodileAgent	-287.1 (14.5)
B	**-334.6 (8.0)**

Broker	Cash
cwiBroker	315.4 (9.3)
U1	**135.3 (12.3)**
CrocodileAgent	-372.1 (17.0)
Mertacor	-485.5 (28.1)

Broker	Cash
U9_MDP	**389.9 (13.3)**
cwiBroker	138.3 (8.7)
CrocodileAgent	-333.3 (17.0)
Mertacor	-494.1 (29.6)

Broker	Cash
U9_MDP_LWR	**350.8 (13.3)**
cwiBroker	132.4 (9.0)
CrocodileAgent	-336.9 (17.3)
Mertacor	-566.1 (26.8)

- Tariff and Wholesale strategies improve performance
- LWR customer prediction reduces performance
  - Should relax LWR's extrapolation assumptions?

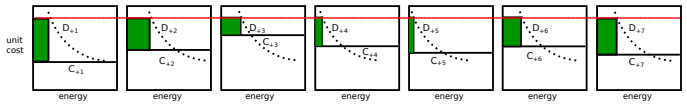
UT Austin Learning Agents Research Group

Related Work: Power Trading Agents

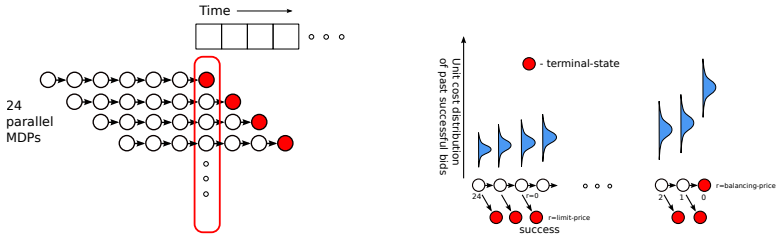
- RL for tariff publications [Peters-2013]
  - Offline preference learning
- Market Bidding MDP [Kuate-2013]
  - Uses a different MDP representation
- Tariff Publication MDP [Reddy-2011]
  - More restrictive setup
- The Power TAC Platform and Competition [Ketter-2013]

Summary

- TacTex'13: utility-optimizing broker agent
- Interdependent optimization problems
  - Utility-maximizing tariff strategy:



  - Online reinforcement learning bidding algorithm:



- Outlook
  - Investigating other tariff, wholesale and balancing strategies
  - Impact on the smart grid and customer behaviors

UT Austin Learning Agents Research Group