Candidate Elimination Algorithm using Version Spaces

1. Initialize G to the set of maximally general hypotheses in H
2. Initialize S to the set of maximally specific hypotheses in H
3. For each training example d, do
    a. If d is a positive example
        i. Remove from G any hypothesis inconsistent with d,
        ii. For each hypothesis s in S that is not consistent with d,
            Remove s from S
            Add to S all minimal generalizations h of s such that h is consistent with d, and some member of G is more general than h
            Remove from S, hypothesis that is more general than another in S
    b. If d is a negative example
        i. Remove from S any hypothesis inconsistent with d
        ii. For each hypothesis g in G that is not consistent with d
            Remove g from G
            Add to G all minimal specializations h of g such that h is consistent with d, and some member of S is more specific than h
            Remove from G any hypothesis that is less general than another in G

Example-1 (Dataset from the text book)

Dataset

sunny,warm,normal,strong,warm,same,Y
sunny,warm,high,strong,warm,same,Y
rainy,cold,high,strong,warm,change,N
sunny,warm,high,strong,cool,change,Y

Solution

```
S[0]: {('0', '0', '0', '0', '0', '0')}
G[0]: {('?', '?', '?', '?', '?', '?')}

Training Instance d: ('sunny', 'warm', 'normal', 'strong', 'warm', 'same', 'Y')
  +ve instance
     G after removing inconsistent hypothesis with d:
        {('?', '?', '?', '?', '?', '?')}
     S generalization
        Consider s:  ('0', '0', '0', '0', '0', '0')
        s after min generalization
           {('sunny', 'warm', 'normal', 'strong', 'warm', 'same')}
        S after retaining s where s is consistent with d & their exist more general g in G
           {('sunny', 'warm', 'normal', 'strong', 'warm', 'same')}
        S after removing less specific hypothesis
           {('sunny', 'warm', 'normal', 'strong', 'warm', 'same')}

S[1]: {('sunny', 'warm', 'normal', 'strong', 'warm', 'same')}
G[1]: {('?', '?', '?', '?', '?', '?')}

Training Instance d: ('sunny', 'warm', 'high', 'strong', 'warm', 'same', 'Y')
  +ve instance
     G after removing inconsistent hypothesis with d:
        {('?', '?', '?', '?', '?', '?')}
     S generalization
        Consider s:  ('sunny', 'warm', 'normal', 'strong', 'warm', 'same')
        s after min generalization
           {('sunny', 'warm', '?', 'strong', 'warm', 'same')}
        S after retaining s where s is consistent with d & their exist more general g in G
           {('sunny', 'warm', '?', 'strong', 'warm', 'same')}
        S after removing less specific hypothesis
           {('sunny', 'warm', '?', 'strong', 'warm', 'same')}

S[2]: {('sunny', 'warm', '?', 'strong', 'warm', 'same')}
G[2]: {('?', '?', '?', '?', '?', '?')}

Training Instance d: ('rainy', 'cold', 'high', 'strong', 'warm', 'change', 'N')
  -ve instance
     S after removing consistent hypothesis with d
        {('sunny', 'warm', '?', 'strong', 'warm', 'same')}
     G specialization
        Consider g:  ('?', '?', '?', '?', '?', '?')
        g after min specialization:
           {('?', '?', '?', '?', 'cool', '?'), ('?', '?', '?', '?', '?', 'same'), ('?', 'warm', '?', '?', '?', '?'),
            ('sunny', '?', '?', '?', '?', '?'), ('?', '?', 'normal', '?', '?', '?')}
        G after retaining g where g is consistent with d & some member of S is more specific than g
           {('?', '?', '?', '?', '?', 'same'), ('?', 'warm', '?', '?', '?', '?'), ('sunny', '?', '?', '?', '?', '?')}
```

G after removing less general hypothesis
{('?', '?', '?', '?', '?', 'same'), ('?', 'warm', '?', '?', '?', '?'), ('sunny', '?', '?', '?', '?', '?')}

S[3]: {('sunny', 'warm', '?', 'strong', 'warm', 'same')}
G[3]: {('?', '?', '?', '?', '?', 'same'), ('?', 'warm', '?', '?', '?', '?'), ('sunny', '?', '?', '?', '?', '?')}

Training Instance d: ('sunny', 'warm', 'high', 'strong', 'cool', 'change', 'Y')
+ve instance
G after removing inconsistent hypothesis with d:
{('?', 'warm', '?', '?', '?', '?'), ('sunny', '?', '?', '?', '?', '?')}
S generalization
Consider s: ('sunny', 'warm', '?', 'strong', 'warm', 'same')
s after min generalization
{('sunny', 'warm', '?', 'strong', '?', '?')}
S after retaining s where s is consistent with d & their exist more general g in G
{('sunny', 'warm', '?', 'strong', '?', '?')}
S after removing less specific hypothesis
{('sunny', 'warm', '?', 'strong', '?', '?')}

S[4]: {('sunny', 'warm', '?', 'strong', '?', '?')}
G[4]: {('?', 'warm', '?', '?', '?', '?'), ('sunny', '?', '?', '?', '?', '?')}

Example-2 (Dataset used in Lab program)

Dataset

big,red,circle,N
small,red,triangle,N
small,red,circle,Y
big,blue,circle,N
small,blue,circle,Y

Solution

S[0]: {('0', '0', '0')}
G[0]: {('?', '?', '?')}

Training Instance d: ('big', 'red', 'circle', 'N')
-ve instance
S after removing consistent hypothesis with d
{('0', '0', '0')}
G specialization
Consider g: ('?', '?', '?')
g after min specialization:
{('?', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', '?')}
G after retaining g where g is consistent with d & some member of S is more specific than g
{('?', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', '?')}
G after removing less general hypothesis
{('?', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', '?')}

S[1]: {('0', '0', '0')}
G[1]: {('?', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', '?')}

Training Instance d: ('small', 'red', 'triangle', 'N')
-ve instance
S after removing consistent hypothesis with d
{('0', '0', '0')}
G specialization
Consider g: ('?', '?', 'triangle')
g after min specialization:
{('big', '?', 'triangle'), ('?', 'blue', 'triangle')}
G after retaining g where g is consistent with d & some member of S is more specific than g
{('?', 'blue', '?'), ('big', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', 'triangle')}
G after removing less general hypothesis
{('?', 'blue', '?'), ('big', '?', 'triangle'), ('small', '?', '?')}
Consider g: ('small', '?', '?')
g after min specialization:
{('small', 'blue', '?'), ('small', '?', 'circle')}
G after retaining g where g is consistent with d & some member of S is more specific than g
{('?', 'blue', '?'), ('small', 'blue', '?'), ('big', '?', 'triangle'), ('small', '?', 'circle')}
G after removing less general hypothesis
{('?', 'blue', '?'), ('big', '?', 'triangle'), ('small', '?', 'circle')}
Consider g: ('?', 'blue', '?')
d is not consistent with g, Do nothing

S[2]: {('0', '0', '0')}
G[2]: {('?', 'blue', '?'), ('big', '?', 'triangle'), ('small', '?', 'circle')}

```
Training Instance d: ('small', 'red', 'circle', 'Y')
  +ve instance
      G after removing inconsistent hypothesis with d:
          {('small', '?', 'circle')}
      S generalization
          Consider s:  ('0', '0', '0')
          s after min generalization
              {('small', 'red', 'circle')}
          S after retaining s where s is consistent with d & their exist more general g in G
              {('small', 'red', 'circle')}
          S after removing less specific hypothesis
              {('small', 'red', 'circle')}

S[3]: {('small', 'red', 'circle')}
G[3]: {('small', '?', 'circle')}

Training Instance d: ('big', 'blue', 'circle', 'N')
  -ve instance
      S after removing consistent hypothesis with d
          {('small', 'red', 'circle')}
      G specialization
          Consider g:  ('small', '?', 'circle')
              d is not consistent with g, Do nothing

S[4]: {('small', 'red', 'circle')}
G[4]: {('small', '?', 'circle')}

Training Instance d: ('small', 'blue', 'circle', 'Y')
  +ve instance
      G after removing inconsistent hypothesis with d:
          {('small', '?', 'circle')}
      S generalization
          Consider s:  ('small', 'red', 'circle')
          s after min generalization
              {('small', '?', 'circle')}
          S after retaining s where s is consistent with d & their exist more general g in G
              {('small', '?', 'circle')}
          S after removing less specific hypothesis
              {('small', '?', 'circle')}

S[5]: {('small', '?', 'circle')}
G[5]: {('small', '?', 'circle')}
```

Example-3 (Datset given in Assignment-1)

Dataset

small,blue,circle,Y
big,red,circle,N
small,red,triangle,N
small,red,circle,Y
big,blue,circle,N

Solution

S[0]: {('0', '0', '0')}
G[0]: {('?', '?', '?')}

Training Instance d: ('small', 'blue', 'circle', 'Y')
  +ve instance
    G after removing inconsistent hypothesis with d:
      {('?', '?', '?')}
    S generalization
      Consider s: ('0', '0', '0')
      s after min generalization
        {('small', 'blue', 'circle')}
      S after retaining s where s is consistent with d & their exist more general g in G
        {('small', 'blue', 'circle')}
      S after removing less specific hypothesis
        {('small', 'blue', 'circle')}

S[1]: {('small', 'blue', 'circle')}
G[1]: {('?', '?', '?')}

Training Instance d: ('big', 'red', 'circle', 'N')
  -ve instance
    S after removing consistent hypothesis with d
      {('small', 'blue', 'circle')}
    G specialization
      Consider g: ('?', '?', '?')
      g after min specialization:
        {('?', '?', 'triangle'), ('small', '?', '?'), ('?', 'blue', '?')}
      G after retaining g where g is consistent with d & some member of S is more specific than g:
        {('small', '?', '?'), ('?', 'blue', '?')}
      G after removing less general than hypothesis
        {('small', '?', '?'), ('?', 'blue', '?')}

S[2]: {('small', 'blue', 'circle')}
G[2]: {('small', '?', '?'), ('?', 'blue', '?')}

Training Instance d: ('small', 'red', 'triangle', 'N')
  -ve instance
    S after removing consistent hypothesis with d
      {('small', 'blue', 'circle')}
    G specialization
      Consider g: ('small', '?', '?')
      g after min specialization:
        {('small', 'blue', '?'), ('small', '?', 'circle')}
      G after retaining g where g is consistent with d, and some member of S is more specific than g:

```
            {('small', 'blue', '?'), ('small', '?', 'circle'), ('?', 'blue', '?')}
        G after removing less general than hypothesis
            {('small', '?', 'circle'), ('?', 'blue', '?')}

          Consider g:  ('?', 'blue', '?')
            d is not consistent with g, Do nothing

S[3]: {('small', 'blue', 'circle')}
G[3]: {('small', '?', 'circle'), ('?', 'blue', '?')}

Training Instance d: ('small', 'red', 'circle', 'Y')
  +ve instance
      G after removing inconsistent hypothesis with d:
          {('small', '?', 'circle')}
      S generalization
          Consider s:  ('small', 'blue', 'circle')
          s after min generalization
              {('small', '?', 'circle')}
          S after retaining s for which their exist g in G such that g is general than s
              {('small', '?', 'circle')}
          S after removing less specific hypothesis
              {('small', '?', 'circle')}

S[4]: {('small', '?', 'circle')}
G[4]: {('small', '?', 'circle')}

Training Instance d: ('big', 'blue', 'circle', 'N')
  -ve instance
      S after removing consistent hypothesis with d
          {('small', '?', 'circle')}
      G specialization
          Consider g:  ('small', '?', 'circle')
              d is not consistent with g, Do nothing

S[5]: {('small', '?', 'circle')}
G[5]: {('small', '?', 'circle')}
```

Example 4: (Dataset from Internet)

Dataset

Japan,Honda,Blue,1980,Economy,Y
Japan,Toyota,Green,1970,Sports,N
Japan,Toyota,Blue,1990,Economy,Y
USA,Chrysler,Red,1980,Economy,N
Japan,Honda,White,1980,Economy,Y
Japan,Toyota,Green,1980,Economy,Y
Japan,Honda,Red,1990,Economy,N

Solution

S[0]: {('0', '0', '0', '0', '0')}
G[0]: {('?', '?', '?', '?', '?')}

Training Instance d: ('Japan', 'Honda', 'Blue', '1980', 'Economy', 'Y')
  +ve instance
    G after removing inconsistent hypothesis with d:
      {('?', '?', '?', '?', '?')}
    S generalization
      Consider s: ('0', '0', '0', '0', '0')
      s after min generalization
        {('Japan', 'Honda', 'Blue', '1980', 'Economy')}
      S after retaining s where s is consistent with d & their exist more general g in G
        {('Japan', 'Honda', 'Blue', '1980', 'Economy')}
      S after removing less specific hypothesis
        {('Japan', 'Honda', 'Blue', '1980', 'Economy')}

S[1]: {('Japan', 'Honda', 'Blue', '1980', 'Economy')}
G[1]: {('?', '?', '?', '?', '?')}

Training Instance d: ('Japan', 'Toyota', 'Green', '1970', 'Sports', 'N')
  -ve instance
    S after removing consistent hypothesis with d
      {('Japan', 'Honda', 'Blue', '1980', 'Economy')}
    G specialization
      Consider g: ('?', '?', '?', '?', '?')
      g after min specialization:
        {('?', '?', '?', '1980', '?'), ('?', '?', '?', '?', 'Economy'), ('?', '?', '?', '1990', '?'), ('?', '?', 'Red', '?', '?'),
         ('?', 'Chrysler', '?', '?', '?'), ('?', 'Honda', '?', '?', '?'), ('?', '?', 'Blue', '?', '?'), ('USA', '?', '?', '?', '?'),
         ('?', '?', 'White', '?', '?')}
      G after retaining g where g is consistent with d & some member of S is more specific than g
        {('?', '?', '?', '1980', '?'), ('?', '?', '?', '?', 'Economy'), ('?', 'Honda', '?', '?', '?'), ('?', '?', 'Blue', '?', '?')}
      G after removing less general hypothesis
        {('?', '?', '?', '1980', '?'), ('?', '?', '?', '?', 'Economy'), ('?', 'Honda', '?', '?', '?'), ('?', '?', 'Blue', '?', '?')}

S[2]: {('Japan', 'Honda', 'Blue', '1980', 'Economy')}
G[2]: {('?', '?', '?', '1980', '?'), ('?', '?', '?', '?', 'Economy'), ('?', 'Honda', '?', '?', '?'), ('?', '?', 'Blue', '?', '?')}

Training Instance d: ('Japan', 'Toyota', 'Blue', '1990', 'Economy', 'Y')
  +ve instance
    G after removing inconsistent hypothesis with d:
      {('?', '?', '?', '?', 'Economy'), ('?', '?', 'Blue', '?', '?')}
    S generalization

```
        Consider s:  ('Japan', 'Honda', 'Blue', '1980', 'Economy')
        s after min generalization
            {('Japan', '?', 'Blue', '?', 'Economy')}
        S after retaining s where s is consistent with d & their exist more general g in G
            {('Japan', '?', 'Blue', '?', 'Economy')}
        S after removing less specific hypothesis
            {('Japan', '?', 'Blue', '?', 'Economy')}

S[3]: {('Japan', '?', 'Blue', '?', 'Economy')}
G[3]: {('?', '?', '?', '?', 'Economy'), ('?', '?', 'Blue', '?', '?')}

Training Instance d: ('USA', 'Chrysler', 'Red', '1980', 'Economy', 'N')
  -ve instance
      S after removing consistent hypothesis with d
          {('Japan', '?', 'Blue', '?', 'Economy')}
      G specialization
          Consider g:  ('?', '?', '?', '?', 'Economy')
          g after min specialization:
              {('?', '?', 'White', '?', 'Economy'), ('?', '?', '?', '1990', 'Economy'), ('?', '?', 'Blue', '?', 'Economy'),
               ('Japan', '?', '?', '?', 'Economy'), ('?', 'Honda', '?', '?', 'Economy'), ('?', '?', '?', '1970', 'Economy'),
               ('?', 'Toyota', '?', '?', 'Economy'), ('?', '?', 'Green', '?', 'Economy')}
          G after retaining g where g is consistent with d & some member of S is more specific than g
              {('?', '?', 'Blue', '?', '?'), ('Japan', '?', '?', '?', 'Economy'), ('?', '?', 'Blue', '?', 'Economy')}
          G after removing less general hypothesis
              {('?', '?', 'Blue', '?', '?'), ('Japan', '?', '?', '?', 'Economy')}
          Consider g:  ('?', '?', 'Blue', '?', '?')
              d is not consistent with g, Do nothing

S[4]: {('Japan', '?', 'Blue', '?', 'Economy')}
G[4]: {('?', '?', 'Blue', '?', '?'), ('Japan', '?', '?', '?', 'Economy')}

Training Instance d: ('Japan', 'Honda', 'White', '1980', 'Economy', 'Y')
  +ve instance
      G after removing inconsistent hypothesis with d:
          {('Japan', '?', '?', '?', 'Economy')}
      S generalization
          Consider s:  ('Japan', '?', 'Blue', '?', 'Economy')
          s after min generalization
              {('Japan', '?', '?', '?', 'Economy')}
          S after retaining s where s is consistent with d & their exist more general g in G
              {('Japan', '?', '?', '?', 'Economy')}
          S after removing less specific hypothesis
              {('Japan', '?', '?', '?', 'Economy')}

S[5]: {('Japan', '?', '?', '?', 'Economy')}
G[5]: {('Japan', '?', '?', '?', 'Economy')}

Training Instance d: ('Japan', 'Toyota', 'Green', '1980', 'Economy', 'Y')
  +ve instance
      G after removing inconsistent hypothesis with d:
          {('Japan', '?', '?', '?', 'Economy')}
      S generalization
          Consider s:  ('Japan', '?', '?', '?', 'Economy')
              d is consistant with s, Do nothing
```

S[6]: {('Japan', '?', '?', '?', 'Economy')}
G[6]: {('Japan', '?', '?', '?', 'Economy')}

Training Instance d: ('Japan', 'Honda', 'Red', '1990', 'Economy', 'N')
  -ve instance
    S after removing consistent hypothesis with d
      Null Set {}
    G specialization
      Consider g: ('Japan', '?', '?', '?', 'Economy')
      g after min specialization:
        {('Japan', '?', 'Blue', '?', 'Economy'), ('Japan', 'Toyota', '?', '?', 'Economy'),
         ('Japan', 'Chrysler', '?', '?', 'Economy'), ('Japan', '?', 'White', '?', 'Economy'),
         ('Japan', '?', 'Green', '?', 'Economy'), ('Japan', '?', '?', '1980', 'Economy'),
         ('Japan', '?', '?', '1970', 'Economy')}
      G after retaining g where g is consistent with d & some member of S is more specific than g
        Null Set {}
      G after removing less general hypothesis
        Null Set {}

S[7]: Null Set {}
G[7]: Null Set {}