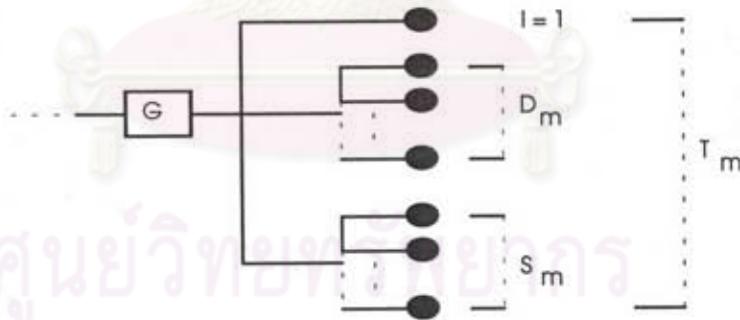




การจำกัดขอบเขตของการแก้ไขปัญหา

การจำกัดขอบเขตของการแก้ไขปัญหาก็เป็นสิ่งที่จำเป็นต้องมี มิฉะนั้นแล้วตัวแรงแสวนจะพยายามทำงานในทุกทางที่เป็นไปได้ ทำให้เสียเวลาในการทำงานมากและบางครั้งอาจทำให้เกิดการทำงานวนซ้ำไม่รู้จบได้ เพราะการทำงานของตัวแรงแสวนเพื่อแก้ปัญหาการผิตไวยากรณ์นั้น ตัวแรงแสวนจะต้องมีการแยกสายโครงสร้างข้อมูล ณ จุดที่ทำการแก้ไขทุกครั้ง ทำให้ขอบเขตของการแก้ไขขยายออกไปเรื่อยๆ

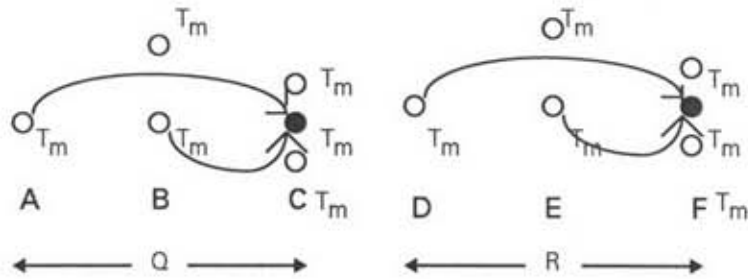
พิจารณาตัวอย่างง่ายๆ ของการแก้ไขการผิตไวยากรณ์ ณ จุดที่ตรวจพบข้อผิดพลาดเท่านั้น การแก้ไขประโยคโดยวิธีนี้ แต่ละครั้งที่มีการแก้ไขประโยค โครงสร้างข้อมูลในสายนั้นมีโอกาสที่จะถูกแยกเป็นจำนวนสูงสุดที่เป็นไปได้ คือ T_m สาย เมื่อ $T_m = I + D_m + S_m$, $I=1$ เป็นสายที่ใช้แก้การผิตไวยากรณ์จากข้อผิดพลาดที่เกิดจากการแทรกข้อมูล, $D_m = S_m$ เป็นจำนวนของชนิดคำที่เป็นไปได้สูงสุด ที่จะใช้ในการแก้ไขประโยคที่เกิดจากข้อผิดพลาดอันเนื่องมาจาก การขาดและการแทนที่องค์ประกอบในประโยคตามลำดับ โดยพิจารณาได้จากตารางแอลอาร์ (จากตัวอย่างในตารางแอลอาร์ที่ให้มาจะมีค่าเป็น 2)



รูปที่ 8.1 แสดงจำนวนโครงสร้างข้อมูล T_m สาย ที่ถูกแยกออกมาเมื่อมีการแก้ไขประโยค

ดังนั้นถ้าในประโยคมีการผิตไวยากรณ์ทั้งหมด n ครั้ง จำนวนของโครงสร้างข้อมูลสูงสุดทั้งหมด ที่มีโอกาสถูกแยกออกมาจะมีค่าเป็น $(T_m)^n$ ซึ่งถ้าใช้ค่าจากตารางแอลอาร์ที่กำหนดให้ก็จะมีความเป็น 5^n

ในกรณีของการแก้ไขไวยากรณ์โดยการแรงแสวนแบบจีแอลอาร์-ไอจีแอลอาร์ ก็จะมีค่าเพื่อขึ้นเป็น สองเท่า คือมีความเป็น $2 \cdot (T_m)^n$ และถ้าเป็นการแรงแสวนโดยวิธีการย้อนรอยกองซ้อนโครงสร้างกราฟโดยทั่วไปแล้วก็จะยิ่งมีค่ามากเพิ่มขึ้นไปอีก โดยอาจพิจารณาคำนวนค่าสูงสุดโดยประมาณ ดังนี้คือ



รูปที่ 8.2 การแยกสายข้อมูลของการแก้ไขประโยคแบบย้อนรอยกองซ้อนโครงสร้างกราฟ

จากรูปที่ 8.2 แสดงการตรวจพบข้อผิดพลาดที่ตำแหน่ง C และ F เราจะประมาณค่าการแยกสายโครงสร้างในกรณีที่มีมากที่สุด โดยกำหนดให้จำนวนสายโครงสร้างที่ถูกแยกออกในแต่ละครั้งเป็น T_m และทำการย้อนรอยกลับมาแก้ไขในทุกทางเลือกที่เป็นไปได้โดยไม่มีข้อจำกัดใดๆ จำนวนทางเลือกทั้งหมดที่เกิดขึ้น จะสามารถคำนวณได้ คือ

การตรวจพบข้อผิดพลาดที่จุด C ตัวแรงส่วนจะย้อนรอยไปแก้ไขประโยคที่จุด A, B และ C ทางเลือกในการแก้ไข คือ

1. แก้ไข 1 จุด มีทางเลือกได้ ${}^3C_1 \cdot T_m$ ทางเลือก
2. แก้ไข 2 จุด มีทางเลือกได้ ${}^3C_2 \cdot (T_m)^2$ ทางเลือก
3. แก้ไข 3 จุด มีทางเลือกได้ ${}^3C_3 \cdot (T_m)^3$ ทางเลือก

รวมเป็นทางเลือกที่ต้องพิจารณา คือ ${}^3C_1 \cdot T_m + {}^3C_2 \cdot (T_m)^2 + {}^3C_3 \cdot (T_m)^3$ ทางเลือก

การตรวจพบข้อผิดพลาดที่จุด F ก็สามารถพิจารณาทางเลือกได้เป็นจำนวนเท่ากัน ดังนั้นทางเลือกที่จะมีได้ในการแก้ไขข้อผิดพลาดในการตรวจพบที่จุด C และ F คือ

$$({}^3C_1 \cdot T_m + {}^3C_2 \cdot (T_m)^2 + {}^3C_3 \cdot (T_m)^3) \times ({}^3C_1 \cdot T_m + {}^3C_2 \cdot (T_m)^2 + {}^3C_3 \cdot (T_m)^3)$$

สามารถสรุปเป็นรูปทั่วไปของจำนวนทางเลือกทั้งหมดได้

$$= ({}^0C_0 \cdot (T_m)^0 + {}^0C_{0,1} \cdot (T_m)^{0,1} + \dots + {}^0C_1 \cdot T_m) \times ({}^RC_R \cdot (T_m)^R + {}^RC_{R,1} \cdot (T_m)^{R,1} + \dots + {}^RC_1 \cdot T_m) \times \dots \\ = [(T_m + 1)^0 - 1] \times [(T_m + 1)^R - 1] \times \dots$$

เมื่อ Q คือช่วงกว้างของการเกิดข้อผิดพลาดที่จุดแรก และ R คือช่วงกว้างของข้อผิดพลาดที่ 2 ถ้ามีข้อผิดพลาดที่จุดอื่นอีกก็จะมีชุดของทางเลือกมากขึ้นในลักษณะเดียวกัน

จะเห็นได้ว่าทางเลือกที่เกิดขึ้นมีเป็นจำนวนมากมาย และการเลือกจุดที่เป็นสาเหตุแท้จริงของการผิดพลาดในกรณีในระดับนี้ก็เป็นที่น่าพอใจ ถ้าจะทำการแก้ไขในทุกทางเลือกก็ไม่เหมาะสมในทางปฏิบัติ ดังนั้นการจำกัดขอบเขตของการแก้ไขประโยคจะช่วยให้การแก้ไขประโยคของตัวแรงส่วนเหมาะสมกับการใช้งานมากขึ้น

การจำกัดขอบเขตของการแก้ไขประโยค สามารถกระทำได้ในหลายทาง เช่น จำกัดจำนวนของการผิดพลาดที่มีในประโยค จำกัดด้วยจำนวนค่าสูงสุดในประโยค ฯลฯ ในการทำวิทยานิพนธ์นี้ เพื่อป้องกันปัญหาจากการใช้หน่วยความจำของเครื่อง PC คอมพิวเตอร์ จึงใช้การจำกัดด้วยการเลือกจุดที่ทำการแก้ไข โดยจะแก้ไขเมื่อมีการตรวจพบข้อผิดพลาด และแก้ไขเฉพาะจุดที่เป็นองค์ประกอบอิสระเท่านั้น ด้วยวิธีนี้จะลดช่วงกว้างของจุดที่ต้องแก้ไขลง ซึ่งก็ขึ้นอยู่กับไวยากรณ์ที่ใช้ (จากตัวอย่างไวยากรณ์ที่ให้สามารถลดจำนวนจุดลงไปประมาณ 1/3) และมีเงื่อนไขในการแก้ไขว่า การแก้ไขที่จุดใดๆ ก็ตามจะต้องไม่ทำให้เกิดข้อผิดพลาดใหม่ตามมาติดๆ กัน (เพื่อป้องกันการแทรกประโยคอย่างต่อเนื่องไม่รู้จบ) ดังนั้นถ้ามีการแก้ไขประโยคในสายโค๊ดที่แก้ลงไป แล้วทำให้เกิดข้อผิดพลาดในจุดใหม่ที่ติดกันตามมา ก็จะหยุดการทำงานในสายโค๊ดตรงนั้นเสีย

อย่างไรก็ตามการจำกัดขอบเขตด้วยเงื่อนไขต่างๆ เหล่านี้อาจจะทำให้การแก้ไขข้อผิดพลาดในบางลักษณะถูกมองข้ามไป ดังนั้นการจำกัดขอบเขตการแก้ไขปัญหาเหล่านี้ จึงขึ้นอยู่กับการทำงานแต่ละอย่าง กลุ่มผู้ใช้งาน และ อุปกรณ์ที่ใช้



ศูนย์วิทยทรัพยากร
จุฬาลงกรณ์มหาวิทยาลัย